BOOTSTRAP ESTIMATES OF CONFIDENCE INTERVALS FOR PARTIAL SURVEY INPUT-OUTPUT MODELS: A FOREST PRODUCTS EXAMPLE

By

Rodney Lee Busby

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Forestry

1991

ABSTRACT

BOOTSTRAP ESTIMATES OF CONFIDENCE INTERVALS FOR PARTIAL SURVEY INPUT-OUTPUT MODELS: A FOREST PRODUCTS EXAMPLE

By

Rodney Lee Busby

Analysts do not commonly provide information about the statistical precision of the input-output multipliers they develop due to the lack of theory relating the variability of the input data with the input-output multipliers generated. Previous work in this area relies upon the incorrect assumptions that the input data are normally distributed and independent of one another.

The bootstrap procedure is a means of estimating the statistical precision of a measure from a single sample of the data that does not rely upon the limiting assumptions of the input data. The idea is to mimic the process of picking many samples from the sample frame by choosing many artificial samples from the original sample data set.

The bootstrap procedure has been found to be effective and was efficiently used to generate measures of precision for multipliers from a partial survey input-output model for Michigan.

Results vary depending upon both the sector and the multiplier examined. The sawmill sector's output multipliers were precisely defined in the analysis. The bootstrap standard deviation is by far the smallest generated. The next best sector is the wood pallet sector followed by the paperboard sector.

Results for the employment multipliers were different than those of the output multipliers. Employment multipliers are very close, with the wood pallet sector providing the best estimate, followed by the paperboard sector, then the sawmill sector.

Results for the income multipliers show that again the reliability of the multipliers were similar, the paperboard sector edged out the wood pallet sector, followed closely by the sawmill sector.

One pattern did emerge, however, the millwork sector had by far the highest standard deviation estimates of the sectors studied. Its average bootstrap standard deviation was two to three times those of the other sectors.

Sample sizes do matter. For example, the standard deviation for one multiplier was reduced by 29 percent when the sample size was increased by 50 percent, it decreased 42 percent when the sample size was doubled.

ACKNOWLEDGEMENTS

I wish to express my gratitude to the many individuals who gave their time and assistance offered to me for the completion of this manuscript. In particular, thanks are due to Dr. Daniel E. Chappelle for his advice and guidance given during the preparation and completion of this report, and to the members of the guidance committee, and others, for their review.

I am also indebted to Bill Siegel, Chief Economist, Southern Forest Experiment Station, for granting the time and providing the opportunity to complete this project. I would also like to thank the many others in the USDA Forest Service that have given me the opportunity, the encouragement, and the support to complete this research.

Finally, I would especially like to thank my dear wife Kathy for all the support she has given me over the years that has made the completion of this project possible.

TABLE OF CONTENTS

Page

LIST OF TABLES v

LIST OF FIGURES viii

CHAPTER 1: INTRODUCTION 1

Objective 1

Hypothesis 2

CHAPTER 2: REVIEW OF LITERATURE 4

Input-Output Analysis 4

Aggregation 13

The Reconciliation Debate 15

Stochastic Input-Output Analysis 16

CHAPTER 3: METHODS 35

Data Sources 35

The Partial-Survey Input-Output Model 42

Bootstrap Confidence Intervals and Probability Functions . 47

Computer Hardware and Software Used 52

Tests Conducted 53

CHAPTER 4: RESULTS 57

Determination of the Number of Bootstrap Replications . . 57

Bootstrap Estimates of Confidence Intervals 61

Determination of the Effects of Sample Size 75

Use of Study Results 84

Hypothesis Evaluation 92

Page

CHAPTER 5: SUMMARY AND CONCLUSIONS 96

Future Research 103

LIST OF REFERENCES 104

APPENDIX A 107

APPENDIX B 131

LIST OF TABLES

Table		Page
1	Hypothetical distributions of firm's direct coefficients	33
2	Sector designations in the Michigan input-output model .	37
3	Michigan survey responses to individual questions	38
4	Average sales, employment, and income for selected sectors	40
5	Employment to output ratios for selected sectors	41
6	Income to output ratios for selected sectors	42
7	Sample size for the original data, the full, three-quarters, and the one-half distribution used to examine the effects of sample size on the distribution of multipliers	56
8	Effects of the number of bootstrap replications on the sawmill sector Type II output multiplier	58
9	Effects of the number of bootstrap replications on the millwork sector Type II output multiplier	59
10	Effects of the number of bootstrap replications on the wood pallet sector Type II output multiplier	60
11	Effects of the number of bootstrap replications on the paperboard sector Type II output multiplier	60
12	Mean Type I and II output, employment, and income multipliers and bootstrap standard deviations for selected sectors	70
13	Summary of mean Type II output multiplier values and bootstrap standard deviations for selected sectors using the full, three-quarters, and one-half sample-size distributions	78

Table		Page

14	Summary of mean Type II employment multiplier values and bootstrap standard deviations for selected sectors using the full, three-quarters, and one-half sample-size distributions	81
15	Summary of mean Type II income multiplier values and bootstrap standard deviations for selected sectors using the full, three-quarters, and one-half sample-size distributions	82
16	Total output change of the Michigan economy due to a hypothetical increase in final demand for output of sawmill, millwork, wood pallet, and paperboard sectors	85
17	Total income change of the Michigan economy due to a hypothetical increase in final demand for output of sawmill, millwork, wood pallet, and paperboard sectors	87
18	Total employment change of the Michigan economy due to a hypothetical increase in final demand for output of sawmill, millwork, wood pallet, and paperboard sectors	89
19	A hypothetical example where the average total output change for the sawmill, millwork, wood pallet, and paperboard sectors are equal	91
A1	Mean Type II output multiplier value, bootstrap standard deviation, and multiplier values for the sawmill sector, for 10 bootstrap trials using the full, three-quarters, and one-half samples	107
A2	Mean Type II output multiplier value, bootstrap standard deviation, and multiplier values for the millwork sector, for 10 bootstrap trials using the full, three-quarters, and one-half samples	108
A3	Mean Type II output multiplier value, bootstrap standard deviation, and multiplier values for the pallet sector, for 10 bootstrap trials using the full, three-quarters, and one-half samples	109
A4	Mean Type II output multiplier value, bootstrap standard deviation, and multiplier values for the paperboard sector, for 10 bootstrap trials using the full, three-quarters, and one-half samples	110
A5	Mean Type II employment multiplier value, bootstrap standard deviation, and multiplier values for the sawmill sector, for 10 bootstrap trials using the full, three-quarters, and one-half samples	111

Table		Page

A6 Mean Type II employment multiplier value, bootstrap standard deviation, and multiplier values for the millwork sector, for 10 bootstrap trials using the full, three-quarters, and one-half samples 112

A7 Mean Type II employment multiplier value, bootstrap standard deviation, and multiplier values for the pallet sector, for 10 bootstrap trials using the full, three-quarters, and one-half samples 113

A8 Mean Type II employment multiplier value, bootstrap standard deviation, and multiplier values for the paperboard sector, for 10 bootstrap trials using the full, three-quarters, and one-half samples 114

A9 Mean Type II income multiplier value, bootstrap standard deviation, and multiplier values for the sawmill sector, for 10 bootstrap trials using the full, three-quarters, and one-half samples 115

A10 Mean Type II income multiplier value, bootstrap standard deviation, and multiplier values for the millwork sector, for 10 bootstrap trials using the full, three-quarters, and one-half samples 116

A11 Mean Type II income multiplier value, bootstrap standard deviation, and multiplier values for the pallet sector, for 10 bootstrap trials using the full, three-quarters, and one-half samples 117

A12 Mean Type II income multiplier value, bootstrap standard deviation, and multiplier values for the paperboard sector, for 10 bootstrap trials using the full, three-quarters, and one-half samples 118

LIST OF FIGURES

Figure		Page
1	Type II output multipliers based upon 600 bootstrap replications	62
2	Type II employment multipliers based upon 600 bootstrap replications	63
3	Type II income multipliers based upon 600 bootstrap replications	64
4	Type I output multipliers based upon 600 bootstrap replications	65
5	Type I employment multipliers based upon 600 bootstrap replications	66
6	Type I income multipliers based upon 600 bootstrap replications	67
7	Type II multiplier 90 percent confidence intervals generated using 600 bootstrap replications	74
8	Type I multiplier 90 percent confidence intervals generated using 600 bootstrap replications	76
A1	Type II output multiplier 90 percent confidence intervals of the sawmill sector generated using 600 bootstrap replications, and 10 trials	119
A2	Type II output multiplier 90 percent confidence intervals of the millwork sector generated using 600 bootstrap replications, and 10 trials	120
A3	Type II output multiplier 90 percent confidence intervals of the wood pallet sector generated using 600 bootstrap replications, and 10 trials	121
A4	Type II output multiplier 90 percent confidence intervals of the paperboard sector generated using 600 bootstrap replications, and 10 trials	122

Figure Page

A5 Type II employment multiplier 90 percent confidence intervals of the sawmill sector generated using 600 bootstrap replications, and 10 trials 123

A6 Type II employment multiplier 90 percent confidence intervals of the millwork sector generated using 600 bootstrap replications, and 10 trials 124

A7 Type II employment multiplier 90 percent confidence intervals of the wood pallet sector generated using 600 bootstrap replications, and 10 trials 125

A8 Type II employment multiplier 90 percent confidence intervals of the paperboard sector generated using 600 bootstrap replications, and 10 trials 126

A9 Type II income multiplier 90 percent confidence intervals of the sawmill sector generated using 600 bootstrap replications, and 10 trials 127

A10 Type II income multiplier 90 percent confidence intervals of the millwork sector generated using 600 bootstrap replications, and 10 trials 128

A11 Type II income multiplier 90 percent confidence intervals of the wood pallet sector generated using 600 bootstrap replications, and 10 trials 129

A12 Type II income multiplier 90 percent confidence intervals of the paperboard sector generated using 600 bootstrap replications, and 10 trials 130

CHAPTER 1

INTRODUCTION

Despite the fact that the parameters of input-output models are not known with certainty, the results from most input-output analyses are provided without any information regarding their reliability. This deficiency has been due to a lack of theory linking errors or variability of the input data to the input-output model generated. Fortunately, recent decreases in the cost of computing has spawned development of computational intensive techniques that use brute computing power to overcome the limiting assumptions required by existing theory.

Objective

The primary objective of this report is to test the effectiveness of using one such technique, the bootstrap procedure, to generate estimates of the statistical precision of input-output multipliers using a partial-survey input-output model. This study will use bootstrap methods to estimate average multiplier values, standard deviations, and confidence intervals for each input-output multiplier generated. A recent partial-survey input-output model for Michigan (Chappelle et al., 1986) will be used in the analysis.

Hypothesis

The hypothesis of this report is that the bootstrap procedure can be effectively and efficiently used to generate confidence intervals and other measures of precision for input-output multipliers.

The evaluation criteria used in this analysis are: (1) the assumptions of the model used to generate the results must be valid; (2) the results must be useful to analysts; (3) the results must be useful to decision-makers; and (4) the approach must be cost effective.

The second chapter contains a review of the literature in this area. First the chapter defines input-output analysis, and describes how each multiplier is typically calculated. The chapter also contains a discussion of three areas that bear upon the problem: (1) aggregation, (2) reconciliation, and (3) stochastic input-output analysis.

The third chapter contains the methods used to conduct the analysis. The chapter starts with a discussion of the data sources used in the analysis. The partial-survey input-output model is discussed next followed by a discussion of the bootstrap confidence intervals and probability functions. This chapter also describes the three tests conducted in this analysis.

The fourth chapter contains the results of the analysis, and contains an example of how the results of the analysis may be used. This chapter also contains an examination of the evaluation criteria and a decision as to whether the hypothesis is accepted or rejected.

The summary and conclusions of the analysis are in the last chapter.

CHAPTER 2

REVIEW OF LITERATURE

Input-Output Analysis[1]

Input-output analysis studies the economic interdependencies within or between nations or regions. Input-output analysis uses a model generated from a flow based accounting system to conduct these analyses. The information is gathered into input-output accounts that provide a quantitative representation of the national, regional, or interregional economy. The information needed for these input-output accounts may be collected using a variety of techniques. Richardson (1985) classifies the data collection techniques into four broad categories: (1) conversion of national coefficients; (2) short cuts; (3) hybrids; and (4) "pure" survey techniques.

There are a variety of techniques to adjust the national input-output model and use the results in regional analysis. Short cuts are techniques to estimate regional input-output multipliers without producing full regional coefficient matrices. Hybrids, or mixed data based models, make the use of some survey information or

---

[1] For a more detailed look at input-output analysis see Miller and Blair (1985), Richardson (1972), or Miernyk (1965).

other reliable information in addition to adjusted national coefficients. The last category is the "pure" survey model.

Jensen and MacDonald (1982:page 34) conclude that the consensus among economists is that "there is no substitute for a good survey-based input-output table." The problem is the cost and time it takes to construct such a model. Richardson (1985: page 618) calls the pure survey based model "an extinct animal on the grounds of time and cost..."

Jensen and MacDonald (1982:page 38) declare that:

> ...the future of regional input-output lies in the development of 'hybrid' tables, i.e. tables which seek to combine the advantages of the expected accuracy of the survey table and the relative speed of construction of the non-survey table.

Richardson (1985:624) agrees, calling the hybrid model "...the wave of the future." This study will use a hybrid input-output model.

As mentioned above, input-output accounts contain flow information that include all of the monetary transactions occurring in a region or between regions for a certain period of time, typically a year. These input-output accounts divide information into three categories: processing sectors, final demand, and the payments or value added sector.

Processing sectors are the intermediate sectors in the economy. These accounts contain information about the purchases and sales among and between firms in the region, rather than final sales, or purchases of inputs from factors of production such as owners of the land, capital, or labor. Final demand accounts contain information about the final sales of goods in the region. The accounts typically divide the information into a number of categories including personal consumption expenditures, exports, and government purchases. The payments or value added accounts contain information about imports, tax payments, wage payments, and capital payments.

Input-output accounts measure flow data. Not all of the information important to a region can be transformed into flow information. Stock information like the land, natural resources, buildings, factories, inventories, and human capital are vitally important to understanding the regional economy. Despite their importance, stock amounts are not captured in most input-output models. Changes in the stocks levels may be incorporated into the model. For example, capital investment or net changes in inventory level may be included as a types of final demand.

The transactions matrix summarizes the information contained in the input-output accounts. The purchases and sales among the institutions inside and outside a region are listed. The transactions table details purchases from: (1) other firms in the region (z), (2) wage payments (L), (3) imports (M), and (4) other value added (V). The total amount of purchases are called total gross outlay (X).

The transactions table also summarizes sales information. This includes sales to: (1) other firms in the region (again z); (2) personal consumption (C), and (3) other final demand (Y). Total sales are called "total gross output" and since total gross output is assumed to be equal to total gross outlay, total gross output is also referred to as X.

Sales to other firms in the region are called interindustry sales. Other final demand (Y) includes governmental purchases, investment, net change in inventory levels, and export sales. Other value added (V) includes indirect business taxes, and capital payments.

The transactions of establishments are aggregated into "sectors" in the transactions table. Equation 1 displays a "n" row and column transactions matrix (T). The sales of each sector may be read by reading across the row. The purchases of each sector may be observed by reading down the column.

$$T = \begin{bmatrix} z_{11} & + \cdots + & z_{1j} & + \cdots + & z_{1n} & + & C_1 & + & Y_1 & = & X_1 \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ z_{i1} & + \cdots + & z_{ij} & + \cdots + & z_{in} & + & C_i & + & Y_i & = & X_i \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ z_{n1} & + \cdots + & z_{nj} & + \cdots + & z_{nn} & + & C_n & + & Y_n & = & X_n \\ L_1 & + \cdots + & L_j & + \cdots + & L_n & + & L_c & + & L_Y & = & L \\ V_1 & + \cdots + & V_j & + \cdots + & V_n & + & V_c & + & V_Y & = & V \\ M_1 & + \cdots + & M_j & + \cdots + & M_n & + & M_c & + & M_Y & = & M \\ X_1 & + \cdots + & X_j & + \cdots + & X_n & + & C & + & Y & = & X \end{bmatrix}, \qquad (1)$$

where all variables have been previously defined.

The transaction matrix may be converted into a useful model by first estimating the technical coefficients matrix (A-matrix) for the region. These technical coefficients show the production function or recipe for producing a dollars worth of output. A fundamental assumption in input-output analysis is that the interindustry flows from sector i to j depend entirely and exclusively on the total output of sector j. The technical coefficients (direct coefficients) may be computed as follows (Miller and Blair, 1985: page 11):

$$a_{ij} = \frac{z_{ij}}{X_j},$$
$$l_i = \frac{L_i}{X_j}, \text{ and} \tag{2}$$
$$c_i = \frac{C_i}{C},$$

where the a_{ij} are the technical coefficients, the l_i are the per unit payments to labor, and the c_i are industry specific distribution of personal consumption expenditures, and the remaining variables are as previously defined.

The production function that is assumed to exist in input-output analysis is as follows:

$$X_j = \left(\frac{z_{1j}}{a_{1j}}, \frac{z_{2j}}{a_{2j}}, \ldots, \frac{z_{mj}}{a_{mj}}\right), \tag{3}$$

the production function uses fixed proportions of inputs to produce a given level output. Additional input of a single commodity would not be sufficient to produce an increase in output. All inputs would have to be increased in the proportion represented in Equation 3 to increase output.

Assuming that consumption expenditures are combined with other final demand expenditures, Equation 2 may be solved for z_{ij} and substituted into Equation 1, the result is Equation 4 below:

$$
\begin{array}{l}
a_{i1}X_i + \cdots + a_{ij}X_i + \cdots + a_{in}X_i + Y_i = X_i \\
\vdots \qquad\qquad \vdots \qquad\qquad \vdots \qquad \vdots \qquad \vdots \\
a_{i1}X_i + \cdots + a_{ij}X_i + \cdots + a_{in}X_i + Y_i = X_i \\
\vdots \qquad\qquad \vdots \qquad\qquad \vdots \qquad \vdots \qquad \vdots \\
a_{n1}X_n + \cdots + a_{nj}X_n + \cdots + a_{nn}X_n + Y_n = X_n.
\end{array} \tag{4}
$$

In more compact matrix notation and solving for X, the set of equations in 4 are the following:

$$X = AX + Y, \tag{5}$$

where A refers to a "n by n" matrix of technical coefficients, X is a "n by 1" vector of total gross outputs, and finally, Y is a "n by 1" vector of final demands. The A matrix or technical coefficients are vital to all input-output calculations, the matrix is as follows:

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{bmatrix}. \tag{6}$$

Equation 5 may be manipulated as follows:

$$X = (I - A)^{-1}Y, \tag{7}$$

where $(I - A)^{-1}$ is called the Leontief inverse or the table of total requirements. In the case of a demand driven input-output model, changes in final demands (Y_i) are predicted, and the Leontief inverse is used to the results are used with Equation 7 to predict new levels of total gross output (X_i). Several assumption are made when using this relationship, they are: (1) no changes in relative prices of the products in the region, (2) no new industrial sectors appear in the region, (3) no change in the technology for producing the product. The Leontief inverse matrix is sometimes called the B matrix and the term may be substituted into Equation 7 as follows:

$$X = BY. \tag{8}$$

The Leontief inverse is a matrix where each column represents the total impact of each industry j. The matrix may be represented as:

$$B = \begin{bmatrix} b_{11} & \cdots & b_{1j} & \cdots & b_{1n} \\ \vdots & & \vdots & & \vdots \\ b_{i1} & \cdots & b_{ij} & \cdots & b_{in} \\ \vdots & & \vdots & & \vdots \\ b_{n1} & \cdots & b_{nj} & \cdots & b_{nn} \end{bmatrix}, \tag{9}$$

where b_{ij} represents the impact on industry i of an unit change in final demand of industry j.

Input-output multipliers are a very important tool used in regional analysis. There are three industry-specific types of input-output multipliers commonly used, they are output, income, and employment multipliers. There are other multipliers that are possible to construct but these multipliers are not as commonly used. Output, income and employment multipliers measure the total impact on output, income, and employment given an initial expenditure change. The multipliers are specific to the industry and vary depending upon the degree of model closure. For example, the A matrix in Equation 6 does not include either the labor income coefficients (l_i) or the personal consumption coefficients (c_i) which are called the household sector. Therefore the calculated B matrix does not measure the effect of increasing income to the workers in a region, and its effect on total output. Output, income, and employment multipliers that are calculated without the household sector in the A matrix are called Type I multipliers. The Type I output multiplier for sector j (OUT_{Ij}) is simply the sum of the j^{th} column of the B matrix:

$$OUT_{Ij} = \sum_{i=1}^{n} b_{ij}. \tag{10}$$

The Type I income multiplier may also be calculated using the information in the B matrix. The multiplier is defined as the ratio of the total income change to the direct income change resulting from a change in final demand. The Type I refers again to the fact that household sector is not incorporated into the model, i.e. the household sector is taken as part of Y, and is fixed exogenously. The Type I income multiplier for sector j (INC_{Ij}) is calculated as follows:

$$INC_{Ij} = \sum_{i=1}^{n} \frac{b_{ij} l_i}{l_j}. \tag{11}$$

The employment to output ratio for each sector j (π_j) is needed to calculate the remaining type of multiplier, the employment multiplier. The employment to output ratio is simply defined as the ratio of total employment in the sector to total output:

$$\pi_j = \frac{E_j}{X_j}, \tag{12}$$

where E_j is the total employment in sector j.

The Type I employment multiplier for sector j (EMP_{Ij}) is defined similarly to the income multiplier, that is, EMP_{Ij} is defined as the ratio of total to direct employment change resulting from a change in final demand. Again, since the household sector is not enclosed in the A matrix, the effects of respending income of households in response to changes in income in the economy are not captured in this multiplier. EMP_{Ij} is calculated as:

$$EMP_{Ij} = \sum_{i=1}^{n} \frac{b_{ij}\pi_i}{\pi_j}. \tag{13}$$

The input-output model may be built incorporating the household sector into the model. The labor input, and personal consumption coefficients are added to the A matrix and the resulting matrix is as follows:

$$A^* = \begin{bmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} & c_1 \\ \vdots & & \vdots & & \vdots & \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} & c_i \\ \vdots & & \vdots & & \vdots & \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} & c_n \\ l_1 & \cdots & l_i & \cdots & l_n & l_c \end{bmatrix}. \tag{14}$$

Equation 5 may be manipulated using this augmented A matrix (A^*) and the Leontief inverse calculated. The resulting matrix may be referred to as the B^* matrix, as follows:

$$X = (I - A^*)^{-1}Y, \text{ or} \\ X = \quad B^*Y. \tag{15}$$

Finally, the Type II output, income, and employment multipliers, OUT_{IIj}, INC_{IIj}, and EMP_{IIj}, respectively, may be calculated using the B^* matrix:

$$OUT_{IIj} = \sum_{i=1}^{n} b^*_{ij}, \\ INC_{IIj} = \sum_{i=1}^{n} \frac{b^*_{ij} l_i}{l_j}, \text{ and} \\ EMP_{IIj} = \sum_{i=1}^{n} \frac{b^*_{ij} \pi_i}{\pi_j}, \tag{16}$$

where b^*_{ij} refer to the entries of the Leontief inverse matrix that included the induced effects of the household sector.

Aggregation

Aggregation affects the results of input-output analysis. A sector is an aggregation of the establishments in a region. A completely disaggregated input-output model could, theoretically, identify each individual in the economy! Analysts aggregate to make the modeling task feasible, to provide confidentiality to individual firms in the region, and still provide a meaningful model of the economy.

While Miller and Blair describe the theoretical input-output production function in Equation 3 above; Karaska (1968: page 215) describes the production function actually estimated:

> ...each industrial sector as defined must be some combination of production functions. Thus coefficients in the input-output matrix are not 'real' technological coefficients, but are averages of production functions from many different firms.

Aggregation in the directly impacted sector has more effect on economic impact analysis results than aggregation in other sectors. Katz and Burford (1981: page 54) write that:

> ...we have shown that the major problem of aggregation is not a loss of accuracy from aggregation in industries other than the one of interest, but rather, the loss of detail for the specific industry or firm of interest ... In general it can be concluded that accurate information is needed for the specific industry or firm of interest but the multipliers are little affected by aggregation or error for other industries in the matrix.

The seriousness of the aggregation problem is also a function of the variability in the production function and purchasing patterns of firms in the sector. Aggregation of similar firms causes no harm. Aggregation of dissimilar firms in sectors of direct interest causes loss of precision of the economic impact analysis.

Aggregation of dissimilar firms increases the variability of the regional direct coefficients. Karaska (1968: page 223) concludes that "...variation increased with the level of aggregation."

The establishments in a region may be aggregated in accord with any sectoring scheme that might meet the needs of the decision-maker and the data constraints that exist. Disclosure rules forbid the release of individual firm survey data. Generally, data from at least three firms must be aggregated to prevent inadvertent disclosure of individual establishment data.

The aggregation problem and the problem of estimation of confidence intervals around input-output coefficients can be separated. The problem may be studied for any proposed aggregation scheme. In the interest of clarification of the confidence interval problem, the existing aggregation scheme in Chappelle et al. (1986) will not be altered.

The Reconciliation Debate

Chappelle et al. (1986) asked each of the establishments surveyed to detail their purchases and sales. The results were two estimated transactions matrices, one based upon purchases and one based upon sales. Reconciliation techniques were used to force the two estimates to agree. Chappelle et al. (1986) used a reconciliation technique suggested by Miernyk et al. (1970).

In total, three transactions matrices were proposed: (1) the purchases; (2) the reconciled; and (3) the sales transactions matrices. Income multipliers were estimated using each of the three transactions matrices.

The purchases and sales income multipliers ranged from 9.7 percent below to 25.5 percent above the reconciled matrix (Chappelle et al., 1986: page 15).

To focus study on the confidence interval question, this study will use only the purchases transaction matrix. The result will be a set of confidence intervals for the purchase-based transactions matrix. A similar study could be made for a sales-based transactions matrix or the reconciled matrix. The key requirement in order to make the process work for the reconciled matrix would be the use of some automatic decision rule that would make the reconciliation process automatic. Procedures are available to accomplish this and are discussed in Miller and Blair (1985). The reconciliation problem limits the usefulness of the confidence interval estimation technique used here but the problems can and will be separated to focus attention on the primary goal of this paper.

Stochastic Input-Output Analysis

One important question in input-output data collection efforts is--what are the important data items to collect? Should the analyst collect data on regional technical coefficients, regional purchase coefficients, household consumption coefficients, and/or labor input coefficients? Can national input-output coefficients be used instead of regional data for some data items?

Analysts may make the distinction between regional technical coefficients and regional purchase coefficients. Regional technical coefficients are the technical requirements for producing a product, no matter where produced. Regional purchase coefficients (RPCs) are the proportion of the technical requirements that are purchased locally. For example, suppose that for each dollar of output of the sawmill sector, the sawmill sector purchases twenty five cents from the logging sector. In a particular region, sixty five percent of such purchases are done locally. The technical coefficient would be .25, the RPC would be .65. The regional direct coefficient that details the sawmill purchases from the local logging sector would be .1625 (.25 times .65). The household consumption coefficients list purchases of households from each sector in the region. The labor input coefficients are the amount of each dollar spent on labor in each sector.

Stevens and Trainer (1978) conclude that first priority should be given to collection of regional purchase coefficient data. In fact they state that:

> It is possible that system errors due to A-matrix errors are likely to be so small that it would be difficult ever again to justify constructing a regional table based entirely on survey data (Stevens and Trainer, 1978: page 28).

Park et al. (1981) studied the impact of four types of error: (1) technical coefficient error; (2) household consumption coefficient error; (3) labor input coefficient error; and (4) regional purchase coefficient error. Random, additive errors of 10, 20, 30 and 40 percent were introduced to a survey based matrix for each type of error. The random error assumed to have a distribution with a mean of zero and a standard deviation of one half the error percentage (10, 20, 30 and 40 percent, respectively). Park et al. (1981:page 335) concluded that:

> Our results reconfirm the main finding of the ST [Stevens and Trainer, 1976] and other studies in this area that errors in sectoral output and multipliers calculated from the nonsurvey I-O table tend to be far more sensitive to errors in the regional purchase coefficients than to errors in the technical coefficients. Moreover, the effect of errors in the technical coefficient matrix is surprisingly negligible...

In contrast, Garhart (1985:page 364) concluded that:

> Rather than completely neglecting technical coefficients on the grounds that their inaccuracy will do little or no harm to multiplier accuracy, the regional analyst should recognize that technology can vary across regions, just as regional trade patterns can vary. Depending on the nature of the error and the type of multiplier being considered, errors in the A matrix can cause even greater errors in

> multipliers than can r.p.c. [regional purchase coefficient] errors.

If Garhart is correct, all sources of error may potentially be important. An analyst interested in the effects of error on multiplier accuracy cannot focus on any single element to the exclusion of other sources of possible error. Note that for extractive industries, such as forestry, technology may differ greatly from region to region.

Park et al., Stevens and Trainer, and Garhart's analyses were made by introducing random errors into a pre-existing "A" matrices. Sampling variability was ignored.

Gerking developed a notion of a stochastic input-output model, he (1976a: page 1) had protested that:

> ... the problem of calculating standard errors for the parameter estimates in a static, open input-output model has been virtually ignored. In fact, the need for calculating standard errors had seldom been recognized even though input-output models are often implemented from sample rather than census data.

Gerking defines the input-output production function in stochastic terms as the following:

$$X_j = \left(\frac{z_{1j}}{a_{1j}}, \frac{z_{2j}}{a_{2j}}, \ldots, \frac{z_{mj}}{a_{mj}}\right) + u_j. \tag{17}$$

Equation 17 is the same as Equation 3 except an error term u_j has been added.

Gerking (1976a:page 22) uses the following assumption, not typically made in input-output analysis:

> To implement the cross-sectional approach to estimating the technical coefficients, a new assumption is required which is typically not made in input-output analysis. In

> particular, all firms in each sector must have the same production function.

The stochastic properties that are assigned to input-output models are made by assuming that the intersectoral flow and total output variables are subject to "measurement error." Gerking admits that this assumption is a strong one. He argues that (Gerking, 1976a:page 22):

> ...it should not be judged in terms of its lack of attention to reality. Instead, it ought to be judged according to the value of the results which it makes possible.

Gerking (1976a) examined five techniques to estimate the regional technical coefficients from survey data. They are: (1) ratio estimation, (2) ordinary least squares (OLS), (3) two-stage least squares (2SLS); (4) Wald-Barlett (WB); and (5) Durbin (DM).

Ratio estimation was defined in Equation 2 above. OLS and 2SLS are parametric regression techniques; WB and DM are nonparametric regression techniques.

Ratio estimation gave biased but consistent estimates of regional technical coefficients (Gerking, 1976a: page 24). Gerking claims the resulting input-output model is deterministic since (page 21):

> ...only one observation can be obtained on the ratio Z_{ij}/X_j from one set of cross-sectional data. In other words, after the technical coefficients have been determined there is no remaining information from which their standard errors may be calculated.

Gerking also claims that OLS will yield biased and inconsistent results. 2SLS does yield consistent estimators when applied to cross-sectional data to obtain: (1) estimates of the technical coefficients, and (2) estimates of the standard error for those estimates. Hanseman (1982), however, says that the 2SLS are biased.

The technical coefficients may be estimated by the following 2SLS model (Gerking, 1976a: page 27):

$$
\begin{aligned}
X_j^{(r)} &= \sum_{i=1}^{m} z_{ij}^{(r)} + RV_j^{(r)} + WSj^{(r)} + PG_j^{(r)} \\
z_{1j}^{(r)} &= a_{1j}X_j^{(r)} + \theta_{1j}^{(r)} \\
z_{2j}^{(r)} &= a_{2j}X_j^{(r)} + \theta_{2j}^{(r)} \\
&\vdots \\
z_{mj}^{(r)} &= a_{mj}X_j^{(r)} + \theta_{mj}^{(r)} \\
RV_j^{(r)} &= a_{m+1,j}X_j^{(r)} + \theta_{m+1,j}^{(r)},
\end{aligned}
\qquad (18)
$$

where $X_j^{(r)}$ is the total output of firm r in sector j. $RV_j^{(r)}$ is a symbol for the residual value added. $WS_j^{(r)}$ and $PG_j^{(r)}$ are firm r's wages and taxes paid, these items are assumed to be known with certainty. The term $z_{ij}^{(r)}$ are firm r's purchases from sector i, note that firm r is classified in sector j. Finally, $\theta({}_{ij}{}^{(r)})$ are the error terms. Note that the final equation states that the term for residual value added is a stochastic variable. Note that Gerking uses a different classification scheme than the one presented in Equation 1 above, Gerking divides value added into wages, taxes, and residual valued added. Equation 1 divided value added into wages, imports and other value added.

A key limitation of Gerking's approach is the requirement to view each firm's contribution to the industry's technology coefficients as being the same. Each firm's contribution to the industry's coefficients should be weighted by firm output. If the sawmill sector in a region consists of six sawmills: (1) one a world class sawmill that produced 99 percent of total production, and (2) five small portable one-man sawmills that together produce only 1 percent of total output. It would be folly to fit an equation using each firm data equally in the analysis.

Gerking (1976a:page 28) reports that the 2SLS estimator for the $\hat{a}_{ij}$ can be simplified as follows:

$$\hat{a}_{ij}(1) = \frac{X_j^T Q_j (Q_j^T Q_j)^{-1} Q_j^T Z_{ij}}{X_j^T Q_j (Q_j^T Q_j)^{-1} Q_j^T X_j}, \tag{19}$$

where X_j and Z_{ij} are n_j x 1 vectors containing the sample information on total gross output and interindustry purchases. Q_j is an n_j x 2 matrix composed of the firm specific wage and tax payments. The superscript T refers to a transformed matrix.

Gerking (1976a:page 28) indicates that the asymptotic variance of the coefficient $\hat{a}_{ij}(1)$ is as follows:

$$\hat{\sigma}^2_{a_{ij}(1)} = \frac{\sigma^2_{e_{ij}}}{X_j^T Q_j (Q_j^T Q_j)^{-1} Q_j^T X_{ij}}, \tag{20}$$

Hanseman and Gustafson (1981:page 469) reformulated Gerking's 2SLS model. They point out that:

> Actual inputs are always in fixed ratios to output. On top of this relation we add measurement errors ... the equation for $X_j^{(r)}$ in Gerking's system [Equation 18] ... leads one to believe that the value of $X_j^{(r)}$ depends on the errors in all the input variables. Actually, ... $RV_j^{(r)}$ is the variable measured as a residual...

Modifying the 2SLS equations allows the estimator to be simplified from Gerking's Equation 19, to Equation 21 below (Hanseman and Gustafson, 1981, page 470):

$$a_{ij} = \frac{Q_j^T Z_{ij}}{Q_j^T X_j}. \tag{21}$$

Hanseman used a simulation study to examine how each of the five estimation techniques worked under "small sample" situations rather than the looking at the "asymptotic" properties of the estimators. The five methods were the ratio, OLS, 2SLS, WB, and DM. Hanseman (1982: page 1433-1434) concludes that under conditions of very heteroskedastic errors the ratio estimator yields the best estimators, while under less heteroskedastic and homoskedastic conditions, OLS performed best.

Hanseman (1982: page 1434) concludes that: "Although the ratio estimator seems to perform well, its sampling distribution is unknown and hence confidence intervals cannot be constructed." This paper will argue that analysts can use a simulation technique to take hypothetical samples from the population, and the distribution of the hypothetical samples could be used to form confidence intervals.

Brown and Giarratani (1979: page 621) criticized Gerking (1976a, 1976b) in the following three areas:

> ... (1) the nature of the distribution of stochastic disturbances has not been adequately explored, (2) the unique nature of regional input-output models makes application of stochastic techniques particularly difficult, if not impossible, (3) the estimator of major interest in Gerking's article produces parameter values that are not constrained to satisfy input-output identities ... the finite sample distribution does not possess moments.

Input-output analysts may stratify firms in a sector, producing sub-strata that may be homogeneous and exhibit constant variance but this would be unlikely across a sector, according to Brown and Giarratani. The authors also say that they must know about the non-sampled establishments in their sample frame and use this information to weight the establishments in the sample. Finally, adjustments of the coefficients for trade and transportation ratios, secondary products (etc.) cause the authors to conclude (Brown and Giarratani, 1979: page 622):

> Each of these adjustments will affect the distribution of stochastic errors. We should not, on this account, expect constant variance across establishments nor should we expect the distribution of errors in any sector is independent of that in all other sectors.

The second issue discussed by Brown and Giarratani was the unique nature of regional input-output models. The region's "direct" coefficients are a mixture of "technical" and "trade" coefficients. Brown and Giarratani (1979: page 622) say:

> It is conceivable that all establishments share a Leontief production function. It is not possible in a space-economy that they share common regional input-output coefficients. Each firm, depending on its location in the region, will require a different mix of domestic and imported inputs. The problem may be seen as denying the validity of stochastic methods that assume constant parameters across establishments, or, perhaps, as a problem in 'spatial' heteroscedasticity.

A third issue made by Brown and Giarratani is that stochastic estimators must be meaningful in small-sample situations if they are to be useful in input-output analysis. Note that Hanseman (1982) examined the behavior of several estimators in a hypothetical small sample situation.

Brown and Giarratani also noted that Gerking's model did not constrain the coefficients to conform with the assumptions used for input-output analysis. The sum of the direct, import, and value added coefficients were not constrained to be equal to one. Gerking (1979a) later corrected this deficiency.

Finally, Brown and Giarratani (1979:page 623) say that since:

> For the case of a structural equation with two included endogenous variables—precisely the case examined by Gerking—... moments of the finite-sample distribution for the TSLS estimator exist only up to the number of overidentifying restrictions ... It follows that none of the moments of this distribution exist. One may make parameter estimates, but associated tests of significance are simply not meaningful.

Gerking (1979b: page 625) responds that "among available estimators for a certain equation, one without moments may be most suitable." Gerking accepted the fact that Brown and Giarratani made a valid point about the lack of moments, but reject their conclusion that the technique "fails". That is just an undesirable feature of such estimators.

Miernyk also criticized Gerking's two assumptions: (1) that all firms in each sector had similar production functions; and (2) differences in reported production functions were due to measurement errors. Miernyk (1979: page 37) said:

> My associates and I knew that the establishments in many sector samples would not have identical production functions. In some cases their production functions were not even remotely alike. This is the aggregation problem in its rawest form. To the best of my knowledge no one has devised any technique, stochastic or deterministic, for dealing with this problem. To make the tables comprehensive we were forced to aggregate unlike establishments. We knew, therefore, that there were more than 'random errors' in our transactions data, and the coefficients derived from them.... To the best of my knowledge no one has devised a mathematical or econometric technique for dealing with this problem.
> ... I see no point in further belaboring the question: Are input-output models deterministic or stochastic? My own view--which I think is the conventional one--is that they are deterministic, although they are anything but error free.

Miernyk does recognize that sampling error is a major problem in survey based input-output analysis. Miernyk tested the calculated coefficients saying (Miernyk, 1976: page 54):

> There is no way of measuring difference between the 'true' coefficients and those calculated from survey data, since the true coefficients must remain forever unknown. But we can measure the representativeness of sectoral samples...

To test his samples, Miernyk et al. (1970: page 3) compared each sector's average earnings with a second sample. If there were no significant difference between the two samples, Miernyk felt that the original sample was representative of the population values.

Miernyk overstates his case against the probabilistic input-output model. Surely an adequately large sample would give better information, barring measurement errors, than a small sample. Miernyk uses both sampled and supplemental information to construct the final transaction table. Miernyk (1976: page 53) reports:

> In constructing the West Virginia input-output tables, we relied heavily on the judgement of a large number of industry specialists, state officials, and others who were asked to check our 'first round' transactions table. A number of changes were made on the basis of their intimate knowledge of specific sectors.
>
> The construction of an input-output model is anything but a mechanical process. One does not collect a huge volume of data, run it through a computer, minimize the variances, and then say 'this is it.' A closer analogy would be that of putting together a large and very complicated puzzle.

This author concedes that additional information other than the sampled information can and will be used to construct a transactions table. Survey data are important, however. This paper will focus on the estimation of confidence intervals from a partial survey input-output model omitting the influence of expert judgement since confidence limits must be set around the sample means. Note that a sample would be unnecessary if expert judgement alone were sufficient to construct the transactions tables but still input-output tables are constructed with sampled data.

As mentioned above, Miernyk does make the point that the "true" value of the input-output coefficients must remain forever unknown. Since the true value of the coefficients are unknown, measurements of accuracy, being defined as differences between the "true" and measured values of the input-output coefficients or multipliers, cannot be accomplished. What is measured is the precision of the estimates, that is, the degree to which multiple measurements of a coefficients or multiplier correspond to one another. A technique may yield very precise measurements of a multiplier but be very inaccurate. Several authors use accuracy since they assume that one particular coefficient or multiplier to be the true value, but as Miernyk points out, that is unknowable. Because of this distinction, the term precision not accuracy, which is commonly used in the literature, will be used in this report.

Gerking and Pleeter (1977) estimated the optimal sample size for calculating the direct coefficients of an input-output model using two stage least squares. Two objective functions were analyzed: (1) a minimum variance table of regional coefficients and (2) a minimum variance forecast of total output. The solution to both problems require estimates of coefficient variance and covariance which must be obtained from previous studies or estimated by two-stage sampling. The minimum variance forecast of total output depends upon "unknown parameters including the levels of final demand...and the values assumed by the regional coefficients (Gerking and Pleeter, 1977: page 74)."

Quandt (1958) gave two rationales why regional direct coefficients can be assumed to be probabilistic. First, he states that measured factor proportions (Equation 3 above) may in fact vary since different production processes may be involved producing the same product, firm expansion paths are not straight lines through the origin, or different firms have different production functions. Second, the data may be gathered by sampling techniques and the estimates of the coefficients would be subject to sampling error. Quandt showed that the standard deviation of the solution can be approximated with a high degree of accuracy if the distribution of input coefficients are available and are relatively small.

Quandt (1959) used a simulation technique that looked at the distribution of coefficients of a solution matrix (Leontief inverse) given assumed errors in the direct matrix. Quandt (1959: page 304) concluded:

> ... (1) that the skewness of the errors in the Leontief matrices tends to be transmitted to the solution and (2) that the lognormal distribution provides a fairly adequate description of the distribution of the solution, irrespective of the distribution of the original errors.

Evans (1954) found that positive errors in a direct coefficients matrix lead to positive errors in the Leontief inverse matrix. Negative errors in the direct coefficients matrix lead to negative errors in the Leontief inverse matrix. If positive errors in some matrix elements are offset by negative errors in other matrix elements, the errors will be somewhat compensating, although the degree of such compensation is unknowable.

Ives (1977) studied the sampling variability in the direct coefficients, sampling variability in the inverse and solution elements, and estimated "quasi" confidence regions for I-O solutions.

Ives used an analytic technique to estimate the distributions. Ives' first step was to estimate a upper and lower bound of the sum of the direct coefficient matrix, which he called $_{T}a$. The direct coefficient matrix associated with the maximum and minimum $_{T}a$ or $\hat{A}$ were used to find the Leontief inverses. Those two estimates were then used with the forecasted new levels of final demand to generate two estimates of total sales, X_U and X_L, which represented an upper and lower estimate of total sales. These values represent quasi-confidence intervals for the estimated total output values.

McCamley et al. (1973) approximated the variance and standard errors of employment multipliers for a survey based input-output model. McCamley et al. (1973: page 83) reasoned that:

> ...Studies of this type ordinarily use information obtained from a sample of firms in each sector to develop transactions tables and subsequent results. The sample of firms is usually selected on a probability basis. Thus if the procedure had been repeated (or a different set of random numbers had been used to draw the sample) a different sample of firms and thus a somewhat different transaction table would have been obtained.

McCamley et al. first estimated the variance matrix associated with the transactions matrix. This matrix was estimated using: (1) each firm's distribution of sales and employment, and (2) each county's distribution of sales and employment. The employment information was used to weigh the results since employment control totals were used to expand the sample totals to regional totals. The result was an estimate of the variance matrix for the transactions matrix.

McCamley et al. reasoned that the task of obtaining estimates of multiplier variances would be formidable since multipliers are nonlinear functions of the transactions table elements. They took advantage of the fact that if x is a random vector then it is possible to approximate the variance of the statistic f(x) by:

$$V(f) = \left(\frac{\partial f}{\partial x}\right)' V(x) \left(\frac{\partial f}{\partial x}\right), \tag{22}$$

where V(x) is the variance matrix of x and ($\partial f/\partial x$) is the derivative of f with respect to x evaluated at the mean of x. The statistic f(x) is the multiplier to be estimated, and x is a vector representation of the transactions table. The elements of the $\partial f/\partial x$ vector are thus the partial derivatives of the multipliers with respect to transactions table elements. The partial derivatives were estimated by:

$$\frac{\partial f}{\partial x} = \frac{b_{ij}}{R_j} q_i, \tag{23}$$

where b_{jh} is the jh^{th} element of the Leontief inverse, R_j is the total output in the base period of the j^{th} sector, and q_i is the i^{th} endogenous sector multiplier.

The variance of the h^{th} endogenous sector multiplier is given by:

$$V(q_h) = \sum_{i=1}^{G} \sum_{m=1}^{G} \sum_{j=1}^{G} \sum_{k=1}^{G} \frac{b_{jh}}{R_j} q_i S_{jk}^{im} q_m \frac{b_{kh}}{R_k}, \tag{24}$$

where S_{jk}^{im} is the covariance of the j^{th} element in the i^{th} row and the k^{th} element in the m^{th} row of the transactions matrix, and there are G endogenous sectors.

Since McCamley's study used data in which each firm interviewed supplied information only about its own employment and sales, the covariance term (S_{jk}^{im}) is zero unless i=m. This leads to the final formula used, it is:

$$V(q_h) = \sum_{j=1}^{G}\sum_{k=1}^{G} \frac{b_{jh}}{R_j}\left\{\sum_{i=1}^{G} q_i^2 S_{jk}^{ii}\right\}\frac{b_{kh}}{R_k}. \tag{25}$$

West (1986) estimated the probability density function of input-output multipliers under the assumption of normality of the regional direct coefficients. The probability density function (pdf) for the k^{th} observed multiplier is (West, 1986: page 364-365):

$$f(y)_k = \frac{A + By}{\sqrt{2\pi}\,[A+2By+Cy^2]^{1.5}} \cdot \exp\left\{-\frac{1}{2}\frac{y^2}{(A+2By+Cy^2)}\right\}, \tag{26}$$
$$-\infty < y < \infty$$

where the value is determined by the following three parameters:

$$A = \sum_{i,j=1}^{n} (b_{jk}M_i\sigma_{ij})^2,$$
$$B = \sum_{i,j=1}^{n} b_{jk}M_i b_{ji}\sigma_{ij}^2, \text{ and} \tag{27}$$
$$C = \sum_{i,j=1}^{n} (b_{ji}\sigma_{ij})^2,$$

where σ_{ij} are the standard errors of each element of the regional direct coefficient (A) matrix, b_{ij} are elements of the Leontief inverse, and M_i is the observed multiplier, whether it be an output, employment, or income multiplier.

West next approximated the mean and variance of y. The problem remains of estimating confidence intervals. It turns out that if the function ($AC-B^2$) was close to zero, the distribution of y closely approximates the normal distribution. West made the assumption that the error associated with assuming that ($AC-B^2$) was close to zero was negligible and he then estimated the ($1-\alpha$) confidence intervals by:

$$M_i - z_{\alpha/2}A/(\sqrt{A}+z_{\alpha/2}B) < M_i^* < M_i+z_{\alpha/2}A/(\sqrt{A}+z_{\alpha/2}B), \tag{28}$$

where $z_{\alpha/2}$ is the critical value of the confidence interval and M_i^* is the true multiplier.

West (1986:page 370) concludes by saying that:

> This study ... is subject to a number of limitations, primarily surrounding the original assumptions on the distribution of the input coefficients. This is one area where greater empirical research is needed; it is quite possible that an alternative input coefficient distribution would be more realistic...

In fact West (1986: page 364) says "This [the normality assumption of the original data] is one aspect where complete lack of prior information prevails." That is, the distributions may not be normal.

Jackson (1986) argues that input-output coefficients are probabilistic since firms in a region have different "industrial, institutional and locational factors" that affect them. Jackson reasons that each coefficient in an input-output model should not be a point estimate but they should be considered a probability density function (pdf). Multipliers with confidence intervals may be obtained by simulation. Each direct coefficient (a_{ij}) is drawn randomly from its particular pdf, all multipliers are then calculated in the usual way. Both steps are repeated a large number of times; the resulting distribution of multipliers are used to develop confidence intervals.

Both West (1986) and Jackson (1986) assume that each direct coefficient is independent. West (1986: page 364) says:

> ... the available evidence suggests that the cost of increased complexity and data requirements of a more general model, providing that an appropriate multivariate distribution of coefficients can be formulated, outweighs the resultant improvement in accuracy.

Jackson (1986: page 522) claims that "The independence assumption in the full pdf formulation is actually less restrictive than in the conventional formulations." The conventional formulation assumes that a point estimate will be made for each direct coefficient, whereas Jackson draws from a probability distribution to estimate each direct coefficient.

West may be correct. The additional effort of discovering the analytic solution of the confidence interval problem without the independence assumption may not be worth it, but analysts may use simulation to examine the problem without the expense of discovering the analytic solution.

Jackson may underestimate the problem of the independence assumption. A hypothetical example may help explain the problem with Jackson's approach. Table 1 contains the hypothetical regional production function data for three firms in the same sector. We may assume that the firms are of identical size, therefore, the average

Table 1. Hypothetical distributions of firm's direct coefficients.

Purchases From	Firm A	Firm B	Firm C	Firm D	Average
Sector 1	.10	.40	.00	.15	.1625
Sector 2	.45	.00	.15	.30	.2250
Sector 3	.00	.10	.15	.00	.0625
Total direct purchases	.55	.50	.30	.45	.4500
Other purchases	.45	.50	.70	.55	.5500

coefficient may be used to represent the sector's direct coefficients. Jackson would maintain the entire distribution of individual a_{ij}'s in the analysis. The individual a_{ij} differ since each firm may have differing production technologies or regional purchasing patterns, for example.

Jackson would calculate multipliers by randomly and independently choosing individual a_{ij} coefficients from their distributions until the direct coefficient matrix is filled and the multipliers are calculated in the usual way. The process is repeated many times and the resulting distribution of multipliers may be used to form confidence intervals around the average multiplier. The sum of the firm's direct purchases from the region would vary from .0000 to 1.0000 by the random selection of a_{ij} from the distributions represented in Table 1. Multipliers calculated from most random drawings from each coefficient distributions would not represent any particular firm's production function nor the industry average production function. An estimate of .0000 is far below the regional purchases of .3000 of firm C. An estimate of 1.0000 is far above firm

A's regional interindustry purchases (of .5500). Such estimates are simply not representative of any firm's regional purchases.

The individual coefficients are not independent either; there is a relationship between the values of the direct coefficients. A firm may be located near a source of supply outside of the region, the sum of that establishment's direct coefficients would reflect that location. Another firm may be more labor intensive than another; the firm's pattern of purchases would reflect that fact.

A more accurate representation of the confidence interval surrounding the multiplier would be to keep the firm's direct coefficients intact, and randomly draw firms into the simulation. Each sample would have a consistent set of direct coefficients, that is, the individual firm's average production function. If one were to select a number of samples of direct coefficients in the region, one could simulate all of the different possible combinations of individual firm's responses to the change in final demand. One or a number of firms in the sector may expand (or contract) their output. By keeping the firm level detail intact, one could develop fairly good confidence intervals in order to estimate the impact of the stimulus.

CHAPTER 3

METHODS

The cost of developing a full primary data input-output model are tremendous. In response to this cost, the idea of using a partial survey input-output model was developed (Richardson, 1985). A partial survey or hybrid input-output model, as mentioned above, is a model that has the data for certain sectors estimated by survey data and the data for other sectors estimated by non-survey methods. The combination of the survey and non-survey methods to gather data forms a hybrid, not fully a survey model, and not fully a non-survey model. The cost of gathering the survey data is, of course, much higher than developing the model by non-survey methods.

Partial survey models use survey data for key sectors in the economy, and non-survey data for all other sectors. The definition of key sectors depends upon the interests of decision-makers. Unfortunately, non-survey techniques do not provide estimates of the precision of their estimates and therefore cannot be used in the estimation of confidence intervals.

Data Sources

The primary source of data for this analysis is the partial survey input-output model constructed for Michigan (Chappelle et al., 1986).

Table 2 contains the sector designations used in the study. Part of the model coefficients were gathered using non-survey techniques using data from the Michigan Energy Administration (1980) input-output model (the MEA model). Chappelle et al. replaced the three forestry sectors in the MEA model with survey data for 10 forest product industries (sectors 9-18) and three roundwood producers (sectors 6-8).

The data gathered for the 13 sectors included: (1) total sales, (2) average employment, (3) payroll, (4) previous year's inventory, (5) current inventory, and (6) distribution of purchases of inputs by sector and the percentage of such purchases made within the state. Additional information was gathered in the survey but not used in this analysis. Details on the methods of the survey are found in Heinen (1982), a copy of the questionnaire used is found in Appendix B.

This study will use these survey data for four of the 38 sectors, listed. They are: (1) sawmills and planing mills (sawmills); (2) millwork, flooring, and structural members (millwork); (3) wood pallets and skids (wood pallets); and (4) paperboard containers and boxes (paperboard). These sectors were selected because: (1) survey data had been collected for them, (2) the sectors are representative of the forest product industries in Michigan, and (3) adequate responses for these sectors allowed the tests to be conducted without violation of the confidentially of the data. The rest of the data was provided by the Chappelle et al. (1986) model. The remainder of this report will identify these four sectors as the survey data. The rest of the sectors will be known as the non-survey sectors, whatever the origin of the data. All analyses were conducted using these samples.

Table 2. Sector designations in the Michigan input-output model.

Sector	Sector description	SIC codes
1	Livestock; other ag. prod	01,02,07
2	Metals, minerals, crude petrol., nat. gas	10-14
3	Construction	Part 138,152-179
4	Meats;dairy;pres. food; grains; bak.; bev.	20
5	Textiles and apparel	22,23
6	National forests	--
7	State forests	--
8	Other stumpage sellers	0811
9	Logging contractors	2411
10	Sawmills and planing mills	2421
11	Millwork, flooring, structural members	2426,2431,2439
12	Wood furniture and fixtures	2434,2511,2512,2517 2521,2531,2541
13	Wood pallets and skids	2448
14	Veneer and plywood; other lumber and wood products	2429,2435,2436 2441,2449,2451 2452,2491,2499
15	Int. pulp and paper or paperbd. mills mills, particleboard	2492,2611,2621
16	Paper mills, ex. build. pap. mills; build. pap. & build. board mills	2621,2631,2661
17	Paperboard containers and boxes	265
18	Other pap. prod., conv. pap. & pap prod.	264
19	Printing and publishing	27
20	Chemicals; plastics; drugs; allied products	28
21	Petroleum refining	29
22	Rubber and leather products	30,31
23	Stone, clay, glass, and concrete products	32
24	Primary metal industries	33
25	Fab. met. prod., ex. mach. & trans. equip.	34
26	Machinery	35[excl.355,356, 358,359],36
27	Transportation equipment	37
28	Misc. manufacturing	38,39 & all 24, 25 & 26 not above
29	Transportation and communication	40-42,44-48
30	Electrical and gas utilities	491-492,Pt. 493
31	Water and sanitary service	Pt. 493, 494-497
32	Wholesale and retail trade	501-599
33	Finance, insurance and real estate	601-679
34	Other services	08,355-6,358-9, 701-899 (ex. 88)
35	Government enterprises	
36	Households	88
37	Change in capital & inventory, exports, gov.	
38	Imports and value added	

Table 3 contains the responses to certain survey questions for the four sectors analyzed in this study. The table lists the number of firms responding to: (1) the questionnaire in general, (2) the total sales question, (3) the total employment question, and (4) the distribution of purchases question.

Table 3. Michigan survey responses to individual questions.

Sector	Total responses	Total sales	Total employment	Distribution of purchases
	(number)			
Sawmills and planing mills	78	55	51	42
Millwork, flooring and structural members	39	29	28	23
Wood pallets and skids	52	45	45	40
Paperboard containers and boxes	26	24	24	23

Survey data were checked for completion, consistency between payments to households and indicated payroll, and reasonableness of purchases based upon product mix. Hansen et al. (1953) described the process as follows:

> ...Often the results of compilations, either from samples or censuses or from other sources, are in error due to various causes. The errors have arisen from blunders in compilation, errors in interviewing, or other sources, and obviously the results should be checked thoroughly ... verification procedures and the correction of errors should be carried through...

The question on distribution of purchases had the lowest response (Table 3); the process of checking on completion, consistency, and reasonableness reduced the numbers of useful responses even further. If, for example, a firm said that it had 12 employees but recorded nothing for payroll, the questionnaire may be dropped if there is not enough information to edit the data. The total number of responses remaining are:

sawmills and planing mills	29 responses,
millwork, flooring and structural members	19 responses,
wood pallets and skids	35 responses, and
paperboard containers and boxes	18 responses.

All additional calculations for these four sectors were made from these edited samples.

Note that this editing process is not unique to this study. In any practical survey, the data will be checked and edited, this step is meant to add realism to the project. Note also that this editing process was in addition to any editing that may have been done to the survey forms by Chappelle et al. (1986).

Table 4 shows the average sales, employment and income for the four sectors studied in detail for this analysis. Income is defined in this analysis as payroll income, it does not include proprietors income. The paperboard containers and boxes sector had the highest average sales, employment and income of the four sectors. The lowest average sales, employment, and income came in the wood pallet and skids sector. The data in each sector were quite variable. Sales of firms in the sawmill sector, for example, ranged from one sixteenth of average to several times the average sales.

Table 4. Average sales, employment, and income for selected sectors.

Sector	Number of survey responses	Average sales	Average employ-ment	Average income
		($)	(jobs)	($)
Sawmills and planing mills	29	763,000	15.9	175,000
Millwork, flooring and structural members	19	2,280,000	32.2	490,000
Wood pallets and skids	35	602,000	15.6	131,000
Paperboard containers and boxes	18	5,283,573	50.3	1,087,000

The employment to output ratio (π_j) is defined as the number of jobs provided for each one thousand dollars of output. Jobs refer to employees of the firm, not firm owners who may provide a large part of the labor in small firms. Table 5 shows the weighted average employment to output ratios, and their ranges for each of the 4 sectors. The weighted average employment to output ratio is simply a summation of sampled firm employment divided by sampled firm output. The range in the wood pallets and skids sector is by far the largest examined, although the firm that reported the lowest ratio was an outlier from the rest of the firms in that sector.

Table 5. Employment to output ratios for surveyed sectors.

Sector	Number of survey responses	Range of employment to output ratios	Weighted average employment to output ratios
		------ (jobs/$1000 output) ------	
Sawmills and planing mills	29	.01165 to .08658	.02089
Millwork, flooring and structural members	19	.00703 to .04213	.01412
Wood pallets and skids	35	.00543 to .13060	.02605
Paperboard containers and boxes	18	.00484 to .01794	.00954

The income to output ratio (l_j) is defined as total amount of payroll divided by total output. Table 6 list the range of l_j for the firms in the analysis. The weighted average income to output ratios were weighted by output. It is interesting to note, that although the range of individual firms income to output ratios vary tremendously, the averages for the four sectors are very similar. The range is only from .20567 for the paperboard sector to .22941 for the sawmills and planing mills sector.

Table 6. Income to output ratios for selected sectors.

Sector	Number of survey responses	Range of income to output ratios	Weighted average income to output ratios
		-------- (income/output) --------	
Sawmills and planing mills	29	.11114 to .68264	.22941
Millwork, flooring and structural members	19	.10315 to .50227	.21497
Wood pallets and skids	35	.02082 to .52000	.21830
Paperboard containers and boxes	18	.09550 to .36000	.20567

The Partial Survey Input-Output Model

As mentioned above, there are two sources of data for the partial survey or hybrid input-output model. The first is the survey data for the four sectors modeled in this analysis (h). The information for the remaining sectors were assumed to be fixed in this analysis, they are the survey and non-survey data developed by Chappelle et al. (1986) for their partial survey-based input-output study. The basic model for the transactions data for this hybrid model is presented as Equation 29. The transaction matrix (T_h) is divided into two parts, the sectors where survey data was used (h) and the remaining sectors where the Chappelle's model was used. This analysis uses purchases data only, therefore only column data for the four sectors have the distinctive h. Equation 1 is modified by specifying the four sectors (h) as follows:

$$T_h = \begin{bmatrix} z_{11} + \cdots + z_{1h-1} + z_{1h} + z_{1h+1} + \cdots + z_{1n} + C_1 + Y_1 = X_1 \\ \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\ z_{i1} + \cdots + z_{ih-1} + z_{ih} + z_{ih+1} + \cdots + z_{in} + C_i + Y_i = X_i \\ \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\ z_{n1} + \cdots + z_{nh-1} + z_{nh} + z_{nh+1} + \cdots + z_{nn} + C_n + Y_n = X_n \\ L_1 + \cdots + L_{h-1} + L_h + L_{h+1} + \cdots + L_n + L_C + L_Y = L \\ V_1 + \cdots + V_{h-1} + V_h + V_{h+1} + \cdots + V_n + V_C + V_Y = V \\ M_1 + \cdots + M_{h-1} + M_h + M_{h+1} + \cdots + M_n + M_C + M_Y = M \\ X_1 + \cdots + X_{h-1} + X_h + X_{h+1} + \cdots + X_n + C + Y = X \end{bmatrix}, \tag{29}$$

where h refer to the sectors where survey data is used.

The technical matrix (A_h) without the household sector is calculated next following the same line of reasoning as used to find the similar matrix in Equation 6 above. Equation 30 is the same as Equation 6 except the source of the data for the generation of the coefficients is specified in Equation 30. Again the survey sectors are identified with an h:

$$A_h = \begin{bmatrix} a_{11} & \cdots & a_{1h-1} & a_{1h} & a_{1h+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i1} & \cdots & a_{ih-1} & a_{ih} & a_{ih+1} & \cdots & a_{in} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nh-1} & a_{nh} & a_{nh+1} & \cdots & a_{nn} \end{bmatrix}. \tag{30}$$

where all variables have been previously defined.

Next we have the A matrix with the household sector (A_h^*) is calculated. The sector specific ratio of employee payroll to total output (l_h) is the major contribution of the surveyed data. The remaining information is provided by the Chappelle et al. (1986) model. The A matrix becomes:

$$A_h^* = \begin{bmatrix} a_{11} & \cdots & a_{1h-1} & a_{1h} & a_{1h+1} & \cdots & a_{1n} & c_1 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\ a_{i1} & \cdots & a_{ih-1} & a_{ih} & a_{ih+1} & \cdots & a_{in} & c_i \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & \cdots & a_{nh-1} & a_{nh} & a_{nh+1} & \cdots & a_{nn} & c_n \\ l_1 & \cdots & l_{h-1} & l_h & l_{h+1} & \cdots & l_n & l_c \end{bmatrix}. \tag{31}$$

Sector specific labor/output ratios (π) are also needed to complete the analysis. Again there are two sources of data for the calculation of the labor/output ratios: (1) survey data signified by sector h, and (2) the remaining data which uses Chappelle et al. (1986) model. They are:

$$\pi_h = [\pi_1 \ \cdots \ \pi_{h-1} \ \pi_h \ \pi_{h+1} \ \cdots \ \pi_{1n}]. \tag{32}$$

In summary, the survey data (S_h) needed for this partial survey input-output model is the following:

$$S_h = \begin{bmatrix} a_{1h} \\ \vdots \\ a_{ih} \\ \vdots \\ a_{nh} \\ l_h \\ \pi_h \end{bmatrix}. \tag{33}$$

The survey data used to estimate S_h consists of individual establishment's responses that must be transformed to an estimate of the sector's coefficients. The first step taken was to estimate total firm sales (X_f) as reported sales in Question 1 of the survey (X_r)

plus ending inventory (I_e) less beginning inventory (I_b) as reported in Question 5 of the survey:

$$X_f = X_r + I_e - I_b. \tag{34}$$

The second step of the process was to transform the firm's purchases data in the survey (Question 10) which is expressed in two percentage terms, sector specific purchases as a percentage of total sales (z_{ip}) and sector specific purchases from Michigan industries (M_{im}) into sector specific sales (z_{if}). The transformation was as follows:

$$z_{if} = \frac{(X_f)\,(z_{ip})\,(M_{ix})}{10000}, \tag{35}$$

where the 10,000 is in the equation to cancel the effect of the two percentage terms in the model.

The third step in this process was to attach the firm's total sales data (X_f) to the matrix. The distribution of sales data does not add to total sales because it represents only interindustry transactions, and not wages (L) or other forms of valued added (V) or imports (M). Payroll data (P_f) as recorded in Question 4 of the survey was also attached to the data set. Note that although Question 10 asked for wages as a percentage of total sales, that information was not used since the Question 4 information was assumed to be more reliable since wage information was asked directly. Data for number of employees in each firm (E_f) was also attached to the firm's data set. In addition, there were some minor modifications of sector definitions between the questionnaire and the final sectors used in the analysis, that was accomplished in this step.

The fourth step was to estimate the total of the sample's sales (X_s), distribution of purchases (z_{is}), employment (E_s), and payroll (P_s). This was accomplished by simply summing the survey data for each sector surveyed, as follows.

$$
\begin{aligned}
z_{is} &= \sum_{f=1}^{K} z_{if}, \\
X_s &= \sum_{f=1}^{K} X_f, \\
E_s &= \sum_{f=1}^{K} E_f, \\
P_s &= \sum_{f=1}^{K} P_f,
\end{aligned}
\tag{36}
$$

where there are K firms in the sample. Note, that the effect of this transformation is to weigh the responses by firm sales. A sampled firm with a low sales will have proportionately less impact on the final estimates of S_h than a firm with large sales.

The fifth step in the estimation of sector's data in S_h is to calculate the A matrix elements (a_{is}), the payroll/output ratio (l_s), and the labor/output ratio (π_s) for the sector (s). The transformation is:

$$
\begin{aligned}
a_{is} &= \frac{z_{is}}{X_s}, \\
l_s &= \frac{P_s}{X_s}, \\
\pi_s &= \frac{E_s}{X_s}.
\end{aligned}
\tag{37}
$$

A problem may occur at this point, the sum of the A matrix elements either with or without the payroll/output ratio may sum to more than

1.00 since there was no restrictions placed upon the coefficient values. If the problem exists, there would be a key violation made in the assumptions behind input-output analysis. That problem was tracked, but it was found not to be significant in this analysis.

Equation 37 contains the information required for each surveyed sector (S_h). This surveyed information is used with the non-survey information to generate the Leontief inverse matrices (B and B^*) which are calculated both without the household sector as a part of the A matrix (Equation 30) and with the household sector as a part of the A matrix (Equation 31) using the same technique that Equations 6 and 14 were transformed to the appropriate Leontief inverse matrixes in Equations 7, 8, 9, and 15. Next, Type I and Type II output, income, and employment multipliers are developed for the survey sectors (h) as follows:

$$\begin{aligned} OUT_{Ih} &= \sum_{i=1}^{n} b_{ih}, & OUT_{IIh} &= \sum_{i=1}^{n} b^*_{ih}, \\ INC_{Ih} &= \sum_{i=1}^{n} \frac{b_{ih} l_i}{l_h}, & INC_{IIh} &= \sum_{i=1}^{n} \frac{b^*_{ih} l_i}{l_h}, \\ EMP_{Ih} &= \sum_{i=1}^{n} \frac{b_{ih} \pi_i}{\pi_h}, \text{ and } & EMP_{IIh} &= \sum_{i=1}^{n} \frac{b^*_{ih} \pi_h}{\pi_h}. \end{aligned} \tag{38}$$

Bootstrap Confidence Intervals and Probability Functions

Diaconis and Efron (1983: page 116) discuss the "bootstrap" method. The bootstrap method is a calculation intensive technique that may be used to analyze a data set. As Diaconis and Efron say:

> The payoff for such intensive computation is freedom from two limiting factors that have dominated statistical theory since its beginnings: the assumption that the data

> conform to a bell-shaped curve and the need to focus on statistical measures whose theoretical properties can be analyzed mathematically.
>
> ... Experience has shown that Gaussian theory works quite well even when the Gaussian distribution is only roughly approximated by the data ... For sets of data that do not satisfy the Gaussian assumptions, however, the results of statistical methods based on such assumptions are obviously less reliable. Computer-intensive methods can solve most problems without assuming that the data have a Gaussian distribution.

The bootstrap method would seem to present a valid way of evaluating data. We do not need to make restrictive assumptions about the data prior to using this method. We can generate non-parametric confidence intervals about many input-output coefficients. The result will be distributions for each sector's multipliers. The unique advantage of this approach will provide an estimate of empirical sampling error, not hypothetical error as studied in Stevens and Trainer (1978), Park et al. (1981), and Garhart (1985).

The bootstrap procedure is a means of estimating the statistical precision of a measure from a single sample of the data. The idea, developed by Efron (1979,1982), is to mimic the process of picking many samples from the sample frame by choosing many artificial samples from the original sample data set. Diaconis and Efron (1983: page 120) say:

> The samples are generated from the data in the original sample. The name bootstrap, which is derived from the old saying about pulling yourself up by your own bootstraps, reflects the fact that one available sample gives rise to many others.

The bootstrap procedure does not always generate a picture of the true population values. Diaconis and Efron (1983: page 122) say:

> ... the good properties of the bootstrap are good average properties. Like any other statistical procedure, the bootstrap will give misleading answers for a small

> percentage of the possible samples.
> ... The bootstrap does not always guarantee a true picture of the statistical accuracy of a sample estimate. What has been proved is that the bootstrap gives a good picture of the accuracy of the estimate most of the time. There are always a few samples for which the bootstrap does not work, and one cannot in advance which they are. The limitation is not so much a failure of the bootstrap procedure as a restatement of the conditions of uncertainty under which all statistical analyses must proceed.

In other words, the bootstrap procedure will not work when the original sample does not represent values of the population. By chance, the original sample drawn may be totally unrepresentative of the population. The probability of that occurring, of course, depends upon the variability of the population and the sample size drawn.

The bootstrap procedure was used to estimate the survey data (S_h) needs of the partial survey input-output model in Equation 33 above. There were 4 sectors where survey data was used. The following are the steps used in the analysis which have been adapted from Efron (1979).

1. The first step is to construct the empirical distribution of sampled firms for each of the four sectors. A data set exists with sample sizes of m, n, o, and p, respectively, from the four sectors surveyed. Each sampled firm is an independent observation taken from a sample from all firms available in the sector, after editing. We observe:

$$\begin{aligned} &G_1, G_2, \cdots, G_m \sim G, \\ &H_1, H_2, \cdots, H_n \sim H, \\ &I_1, I_2, \cdots, I_o \sim I, \text{ and} \\ &J_1, J_2, \cdots, J_p \sim J. \end{aligned} \tag{39}$$

2. The second step was to develop an empirical distribution for each sampled sector. This was done by giving each firm an equal chance of being selected. The assumption being is that these distributions will represent the population of each sector. The empirical distributions are as follows:

$$\begin{aligned}
&\hat{G}\text{: mass } \frac{1}{m} \text{ on each observed data point } G_g,\\
&\hat{H}\text{: mass } \frac{1}{n} \text{ on each observed data point } H_h,\\
&\hat{I}\text{: mass } \frac{1}{o} \text{ on each observed data point } I_i, \text{ and}\\
&\hat{J}\text{: mass } \frac{1}{p} \text{ on each observed data point } J_p,
\end{aligned} \tag{40}$$

where:

$g = 1, 2, \cdots, m,$
$h = 1, 2, \cdots, n,$
$i = 1, 2, \cdots, o,$ and
$p = 1, 2, \cdots, p.$

3. The third step is to draw the bootstrap sample from each distribution. That is draw, using independent random sampling with replacement, the following: (1) $\{g_1^*, g_2^*, \ldots g_m^*\}$ from $\hat{G}$; $\{h_1^*, h_2^*, \ldots h_n^*\}$ from $\hat{H}$; $\{i_1^*, i_2^*, \ldots i_o^*\}$ from $\hat{I}$; and finally $\{j_1^*, j_2^*, \ldots j_p^*\}$ from $\hat{J}$. A typical bootstrap sample for industry G might be 2 samples from firm 1, no sample from firm 2, three samples from firm 4, etc. until the sample size of m is reached. The assumption is that each bootstrap sample would be similar to one collected from the population for each sector.

4. Compute the bootstrap replication (R), that is compute the desired input-output coefficients using the bootstrap sample drawn in step 3. Several steps are required to

transform the sampled data into the form required in Equation 33 above (S_h). The steps required are detailed in Equations 34 through 37 above, where Equation 38 contains the desired input-output coefficients.

5. Repeat steps 3 and 4 some large number (B) of times obtaining B independent replications and B estimates of each multiplier listed in Equation 38. The result is a bootstrap estimate of the distribution of multipliers gathered by Monte Carlo simulation. A question immediately surfaces, how large is B? Unfortunately, there is no immediate answer to that question. A test, described in the section below, was used to answer that question.

6. Calculate selected points in the bootstrap distribution. This paper will estimate the 5th, 10th, 25th, 50th, 75th, 90th and 95th percentile points (t) of the cumulative frequency distribution (CFD) of each multiplier distribution E. Given the bootstrap distribution of E calculated in step 5, it is easy to compute the CFD as follows:

$$CFD_{(t)} = \frac{\#(\hat{E} \leq t)}{B}. \tag{41}$$

7. Nonparametric confidence intervals may be estimated using the percentiles calculated in step 6 by setting:

$$E \in [\hat{E}(\alpha), \hat{E}(1-\alpha)], \tag{42}$$

For example, 90 percent confidence intervals may be estimated by $E \in [\hat{E}(5), \hat{E}(95)]$.

8. Bootstrap estimates of the standard deviation of each distribution may be made in a two step process. The first step is to let $[a^*, b^*]$ be the central 68.26% interval of the $\hat{E}^*$ values:

$$\frac{\#\{\hat{E}^* < a^*\}}{B} = .1587, \quad \frac{\#\{\hat{E}^* < b^*\}}{B} = .8413. \tag{43}$$

The bootstrap estimate of the standard deviation ($\hat{\sigma}^{(b)}$) is defined as follows:

$$\hat{\sigma}^{(b)} = \frac{b^* - a^*}{2}. \tag{44}$$

The preceding eight steps were used to conduct the three tests that follow.

Computer Hardware and Software Used

All computer programs used to generate the results of this analysis were original programs using the SAS programming language version 6, using SAS/IML for all matrix calculations. The computer used was a IBM PS/2 Model 80.

Tests Conducted

There were three specific tests conducted in this analysis. The first test examined the number of bootstrap replications (B) needed to estimate the bootstrap distributions. The second test estimated the bootstrap multiplier distributions, from which bootstrap confidence intervals and standard deviations were calculated. The third and final test was to examine the effect small sample size has on the distribution of multipliers.

The literature concerning the bootstrap technique mentions that a large number of bootstrap replications (B) are necessary to define the distributions. The exact number required depends upon variability of the data used to generate the samples. Examples in the literature suggest anywhere from 100 to 1600 replication may be necessary to define each distribution. A key question was--when do the addition of replications fail to change the distribution? Not having prior knowledge of reasonable estimates of the number of replications required, it was decided to test a range of replications to examine the effects on the multiplier distributions. The entire bootstrap procedure as listed above was used with two modifications: (1) the number of replications varied from 200 through 1600 in 200 step increments; and (2) only the Type II output multipliers were calculated for each replication. The range of replications was large in order to capture the upper range of possibilities represented in the literature. The lower value of the range, 200 replications, was chosen since it was felt that it represented a minimum number needed to define the distributions of multipliers. The Type II output multiplier was selected since it is probably the most important of all multipliers used and since output changes are the driving force of all multipliers. That is, the Leontief inverse in output terms is transformed to produce income and employment multipliers.

The second test is the core of the analysis presented here, it is the actual generation of the bootstrap multiplier distributions and confidence intervals for each surveyed sector and each multiplier examined. The eight steps outlined above were used. The number of replications (B) in step 5 was determined by the results of the previous test. Each surveyed firm represents a sample from its sector. A bootstrap sample, as described above, was drawn randomly and with replacement from the pool of samples from each sector. The number of samples drawn in each sector was the same as the number of sampled firms used in the analysis. Once a bootstrap sample was drawn, all of the multipliers were calculated and the process repeated, until the desired number of replications was reached.

The final test conducted in this analysis is to examine the effects of sample size on the process. It would be ideal to have the entire population of firms in each sector upon which to draw samples of various size. The results could then be used to study the impact of sample size on the multipliers generated. These data do not exist, however, the only sample data that exists are the sample data originally collected. The following test was used to examine the effects of sample size on the results given that the only data available are the original sampled data. The smaller the sample size, the more chance that any one sample could be unrepresentative of the population. Therefore, ten samples were drawn for each sample size.

The procedure to test the effects of sample size is as follows:

1. Select the sample size m, n, o, and p for the four sectors to be examined. The bootstrap procedure above uses all samples collected, therefore this step was not listed above.
2. Select the m, n, o, and p samples for this trial for each sector. Each sample is drawn using independent random sampling with replacement from the entire pool of samples originally selected. This new sample is used to generate all multipliers for this trial.
3. Generate the distributions of multipliers using the bootstrap procedure as listed above with the exception that the sample drawn in step 2 is used exclusively to make all calculations.
4. Repeat steps 1 through 3 for each combination of sample size (step 1) and number of trials (step 2) desired.

The original data base, after editing, had 29 samples from the sawmill sector, 19 samples from the millwork sector, 35 from the wood pallet sector, and 18 from the paperboard sector. There are a number of ways to test for the effects of sample size on the results, the test shown here is to simulate the effects of two lower levels of effort in obtaining samples. In order to simulate the effect of "somewhat reduced" effort to gather samples and a "greatly reduced" effort to gather samples, an arbitrary three-quarters and one-half sample size was taken from each sector. The three-quarters sample size was to simulate the effects of a "somewhat reduced" effort and the one-half sample size was assumed to represent a "greatly reduced" effort to gather samples. Since the original number of samples represent a sample from the original population without replacement, it was felt that in order to compare these results, a third sampling number would have to be calculated, that is the "full" distribution. The "full" distribution contains the same number of samples as the original data base but step 1 was gathered with replacement from the original data set. The three distributions are: (1) full, (2) three-quarters, and (3) one-half; the sample size for each sector in each distribution are given in Table 7.

Table 7. Sample size for the original data, the full, three-quarters, and the one-half distribution used to examine the effects of sample size on the distribution of multipliers.

		Sector sample size		
Sector	Original data	full	three-quarters	one-half
sawmills	29	29	22	15
millwork	19	19	14	10
wood pallets	35	35	26	18
paperboard	18	18	14	9

The number of trials was arbitrarily set at 10 each.

CHAPTER 4

RESULTS

Determination of the Number of Bootstrap Replications

The first test conducted was the determination of the number of bootstrap replications required. As mentioned above, calculation of the bootstrap distributions require repeated calculation of parameters. The number of replications are referred to as "large." A key question is--when do the addition of additional replications fail to change the distribution of multipliers? The Type II output multiplier was calculated for four sectors examined in the study. Eight different sample sizes were used ranging from 200 to 1600 replications.

Table 8 shows the results for the sawmill sector. The average mean for all replications for the Type II output multiplier was 2.760 with a range of 2.756 with 400 trials to a high of 2.761 with 200 and 800 trials. The multiplier values at the 5th percentile show a range from 2.653 at 1400 replications through 2.674 at 200 replications. The results show that for the sawmill sector, there is no significant variability related to the number of bootstrap replications. The results for the 200 replications captured the essence of the distribution, the additional replications were not needed.

Table 8. Effects of the number of bootstrap replications on the sawmill sector Type II output multiplier.

Replications		Value of coefficient at the following percentile						
(B)	Mean	5	10	25	50	75	90	95
200	2.761	2.674	2.694	2.723	2.763	2.801	2.830	2.846
400	2.756	2.661	2.687	2.718	2.755	2.794	2.833	2.854
600	2.760	2.670	2.688	2.720	2.755	2.800	2.838	2.862
800	2.761	2.655	2.682	2.722	2.764	2.801	2.837	2.866
1000	2.760	2.667	2.687	2.722	2.756	2.800	2.838	2.865
1200	2.760	2.664	2.685	2.722	2.761	2.799	2.833	2.858
1400	2.759	2.653	2.678	2.718	2.759	2.803	2.839	2.858
1600	2.759	2.660	2.684	2.719	2.758	2.797	2.837	2.860
avg. mean	2.760							

Table 9 show the results for the millwork sector. The range in Type II output coefficients was from 2.262 at 200 replications to 2.307 at 600 replications. The average value of the mean for all replications was 2.295. The 5th percentile multiplier was lowest at 1.862 which was the result of 400 replications, it was highest at 1.909 at 200 replications. The 95th percentile value ranged from a low of 2.774 at 200 replications to 2.910 at 800 replications. The millwork sector has a larger range of multiplier values, but again there were no detectable patterns in the data that showed that 200 replications were not sufficient to capture the essence of the distribution.

Table 9. Effects of the number of bootstrap replications on the millwork sector Type II output multiplier.

Replications (B)	Mean	Value of coefficient at the following percentile 5	10	25	50	75	90	95
200	2.262	1.909	1.932	2.076	2.215	2.441	2.642	2.774
400	2.291	1.862	1.928	2.065	2.240	2.472	2.735	2.849
600	2.307	1.885	1.944	2.079	2.242	2.521	2.780	2.882
800	2.298	1.877	1.948	2.072	2.241	2.484	2.791	2.910
1000	2.302	1.886	1.949	2.064	2.254	2.498	2.759	2.868
1200	2.302	1.889	1.951	2.071	2.257	2.492	2.765	2.877
1400	2.304	1.883	1.937	2.055	2.245	2.504	2.800	2.896
1600	2.295	1.887	1.951	2.071	2.251	2.469	2.745	2.870
avg. mean	2.295							

Table 10 show the Type II output multiplier for the wood pallet sector. The mean values range from 3.063 at both 1400 and 1600 replications to 3.070 at 1200 replications. The multiplier values at the 5th percentile show a range from 2.873 at both 600 and 1400 replications through 2.884 at 400 replications. Again, there is no pattern in the data to suggest that the additional effort of replications beyond 200 added anything to the analysis.

Table 10. Effects of the number of bootstrap replications on the wood pallet sector Type II output multiplier.

Replications (B)	Mean	5	10	25	50	75	90	95
		Value of coefficient at the following percentile						
200	3.069	2.880	2.918	2.981	3.059	3.153	3.234	3.277
400	3.065	2.884	2.915	2.981	3.061	3.145	3.229	3.290
600	3.064	2.873	2.908	2.977	3.060	3.142	3.217	3.263
800	3.064	2.876	2.919	2.987	3.059	3.138	3.211	3.254
1000	3.065	2.874	2.917	2.992	3.062	3.141	3.219	3.264
1200	3.070	2.876	2.198	2.989	3.065	3.149	3.223	3.265
1400	3.063	2.873	2.909	2.985	3.061	3.139	3.212	3.253
1600	3.063	2.879	2.918	2.980	3.060	3.139	3.214	3.263
avg. mean	3.065							

Finally, Table 11 shows results for the paperboard sector. The mean Type II output multiplier ranged from 2.046 at 200 replications to 2.075 at 800 replications. The average value was 2.066. Again, there was no detectable pattern in the data to suggest that 200 replications was too few to capture the essence of the distribution.

Table 11. Effects of the number of bootstrap replications on the paperboard sector Type II output multiplier.

Replications (B)	Mean	5	10	25	50	75	90	95
		Value of coefficient at the following percentile						
200	2.046	1.804	1.842	1.923	2.013	2.134	2.276	2.365
400	2.070	1.830	1.864	1.952	2.041	2.146	2.304	2.390
600	2.066	1.823	1.854	1.939	2.042	2.149	2.313	2.423
800	2.075	1.820	1.867	1.941	2.041	2.177	2.328	2.431
1000	2.071	1.825	1.866	1.946	2.045	2.162	2.315	2.398
1200	2.066	1.834	1.877	1.948	2.042	2.159	2.295	2.397
1400	2.068	1.825	1.863	1.942	2.045	2.172	2.305	2.404
1600	2.063	1.818	1.856	1.937	2.038	2.163	2.294	2.398
avg. mean	2.066							

In all four cases, 200 trials were sufficient to portray the distribution of the multipliers. But in order to be conservative, 600 replications (B) were used in all analyzes in this report. The initial estimate of 200 trials was tripled since not all multipliers were represented in this test and it is unknown whether 200 replications would have been sufficient if other multipliers were chosen in the test.

Bootstrap Estimates of Confidence Intervals

The second test conducted in this analysis was the actual calculation of all the multipliers for the four survey sectors used in the analysis. Both Type I and II output, employment, and income multipliers were generated for each sector sampled. On the basis of the results of the previous section, 600 bootstrap replications were used to define each multiplier. This information is the heart of this analysis.

Figures 1 through 6 contain histograms of the results of this test, each based on 600 replications. Figures 1 through 3 contain the information for the Type II output, employment, and income multipliers, respectively. Figures 4 through 6 contain the information for the Type I output, employment, and income multipliers. Each figure contains a separate histogram for the four sectors examined. Note that all Type II multipliers have been printed using the same scale to allow for each comparisons between the histograms, as have all of the Type I multipliers. Also note that the Type II multipliers are presented first since they are probably the most

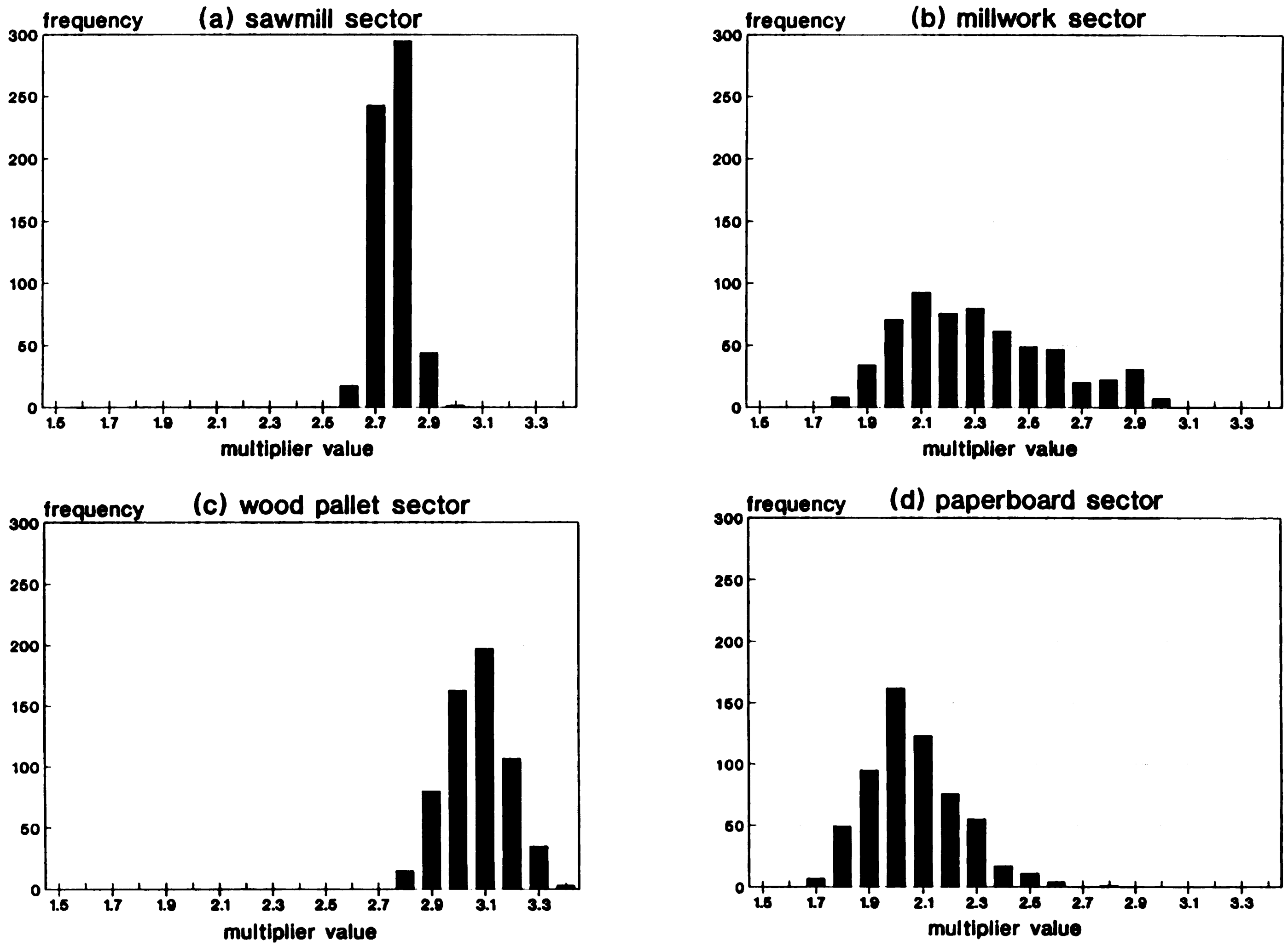

Figure 1. Type II output multipliers based upon 600 bootstrap replications.

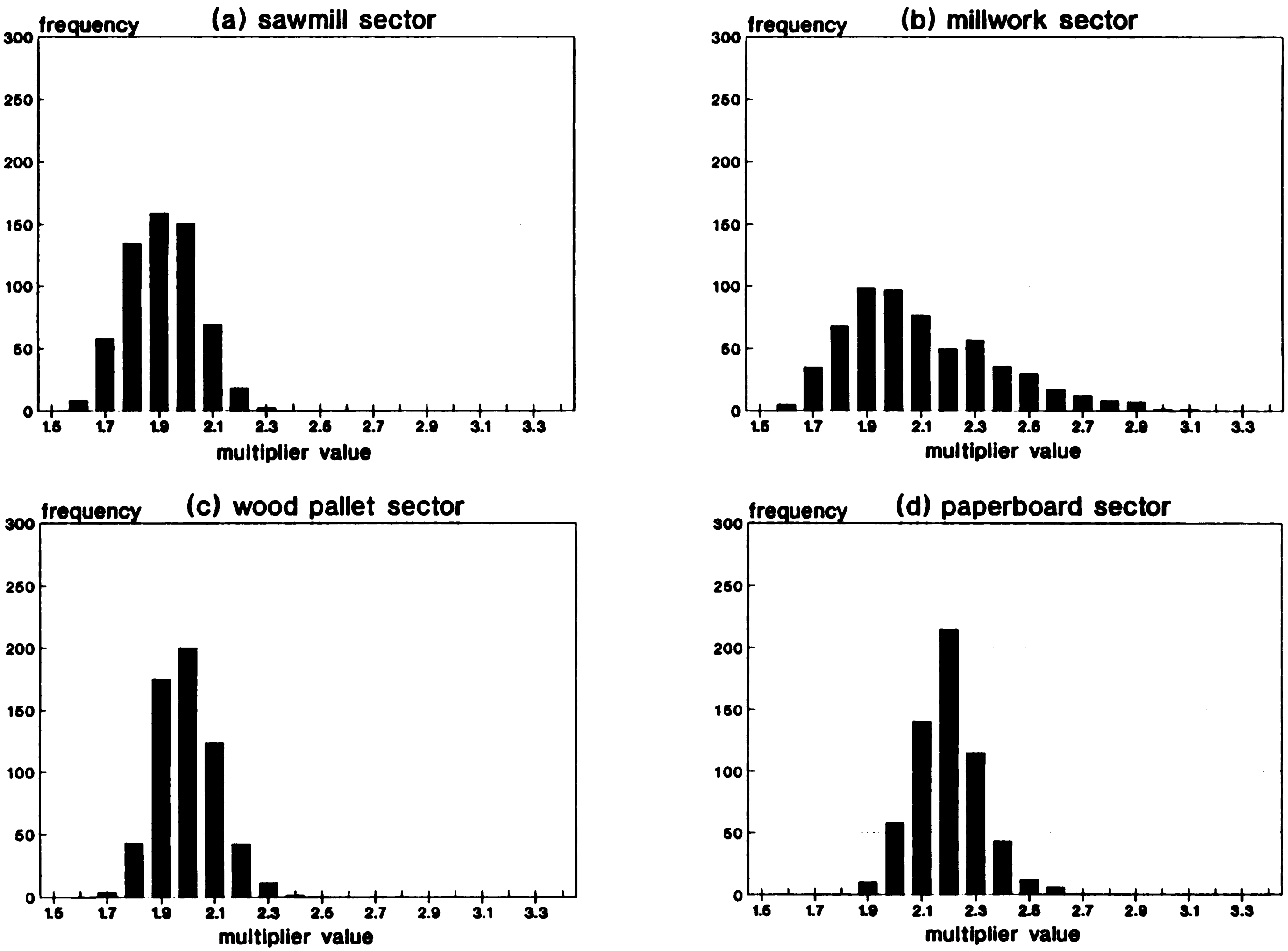

Figure 2. Type II employment multipliers based upon 600 bootstrap replications.

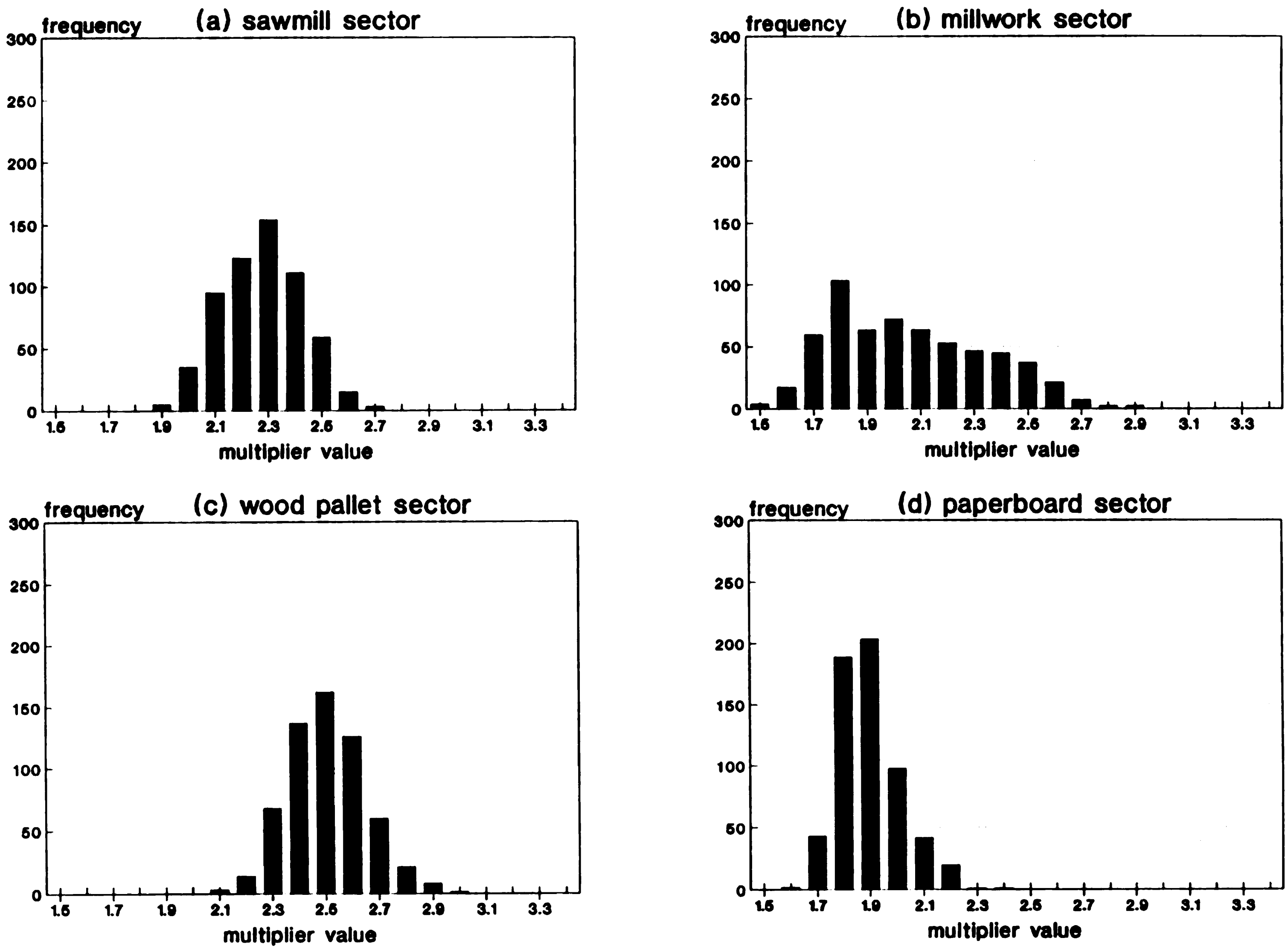

Figure 3. Type II income multipliers based upon 600 bootstrap replications.

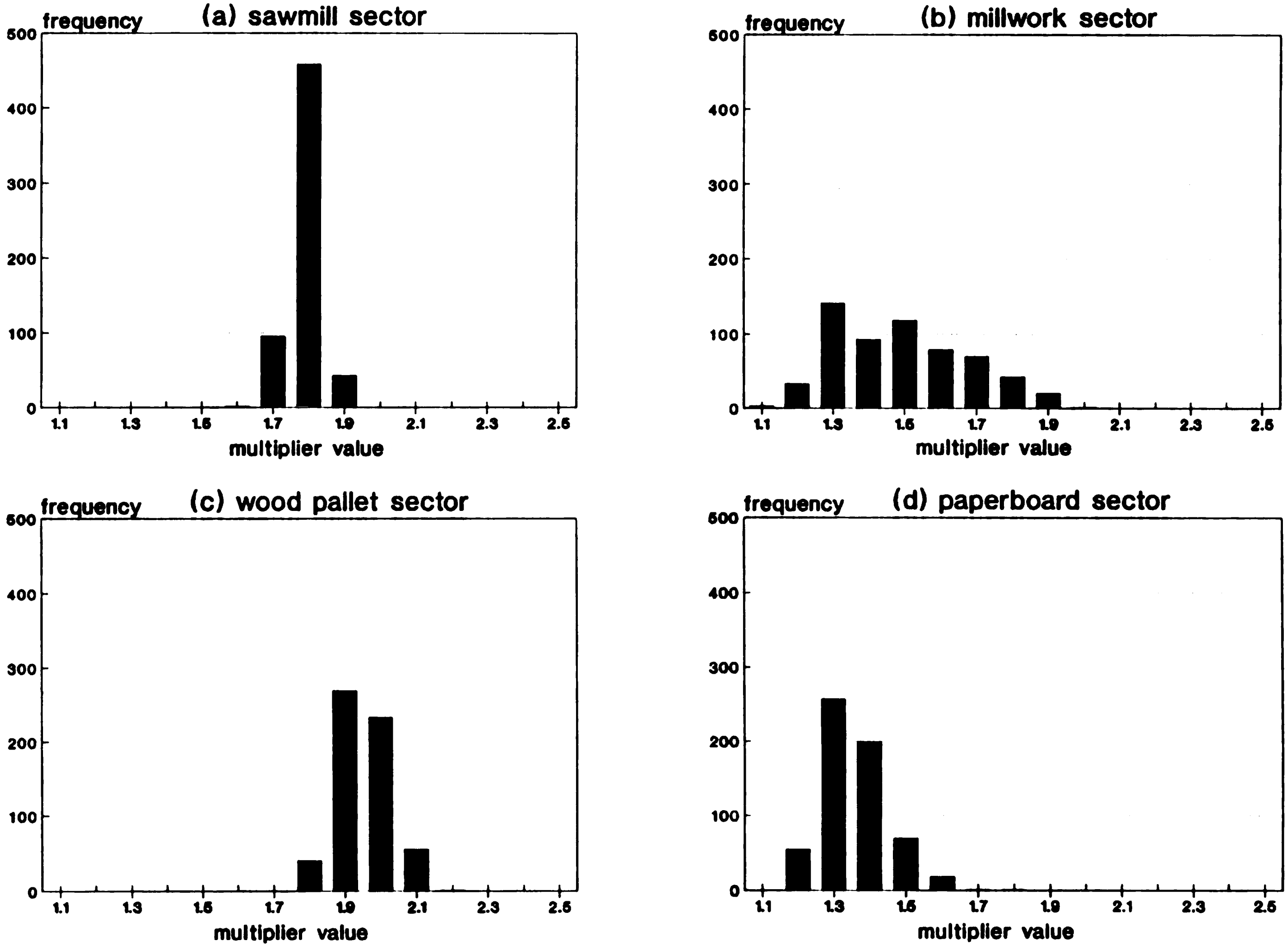

Figure 4. Type I output multiplers based upon 600 bootstrap replications.

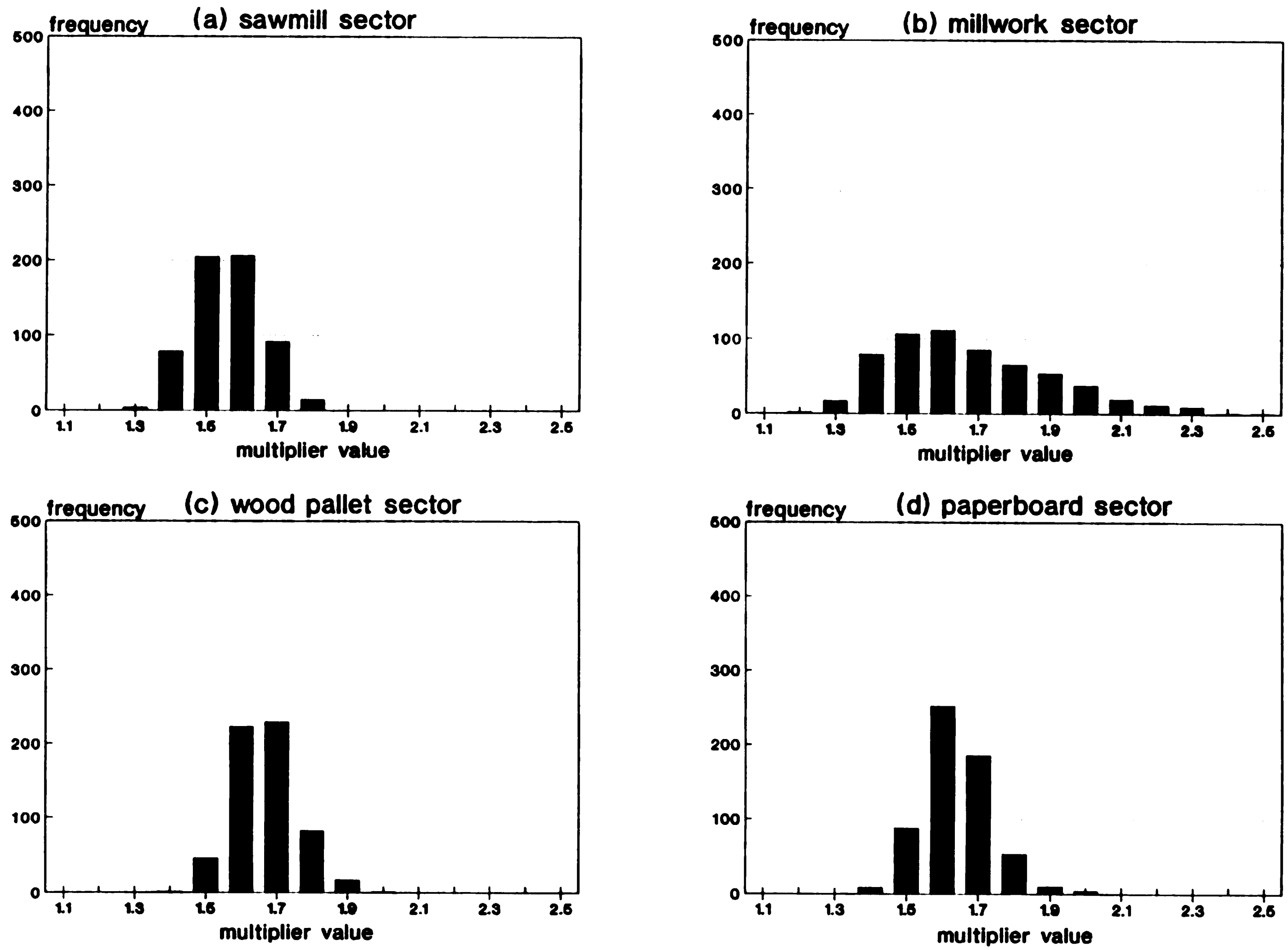

Figure 5. Type I employment multipliers based upon 600 bootstrap replications.

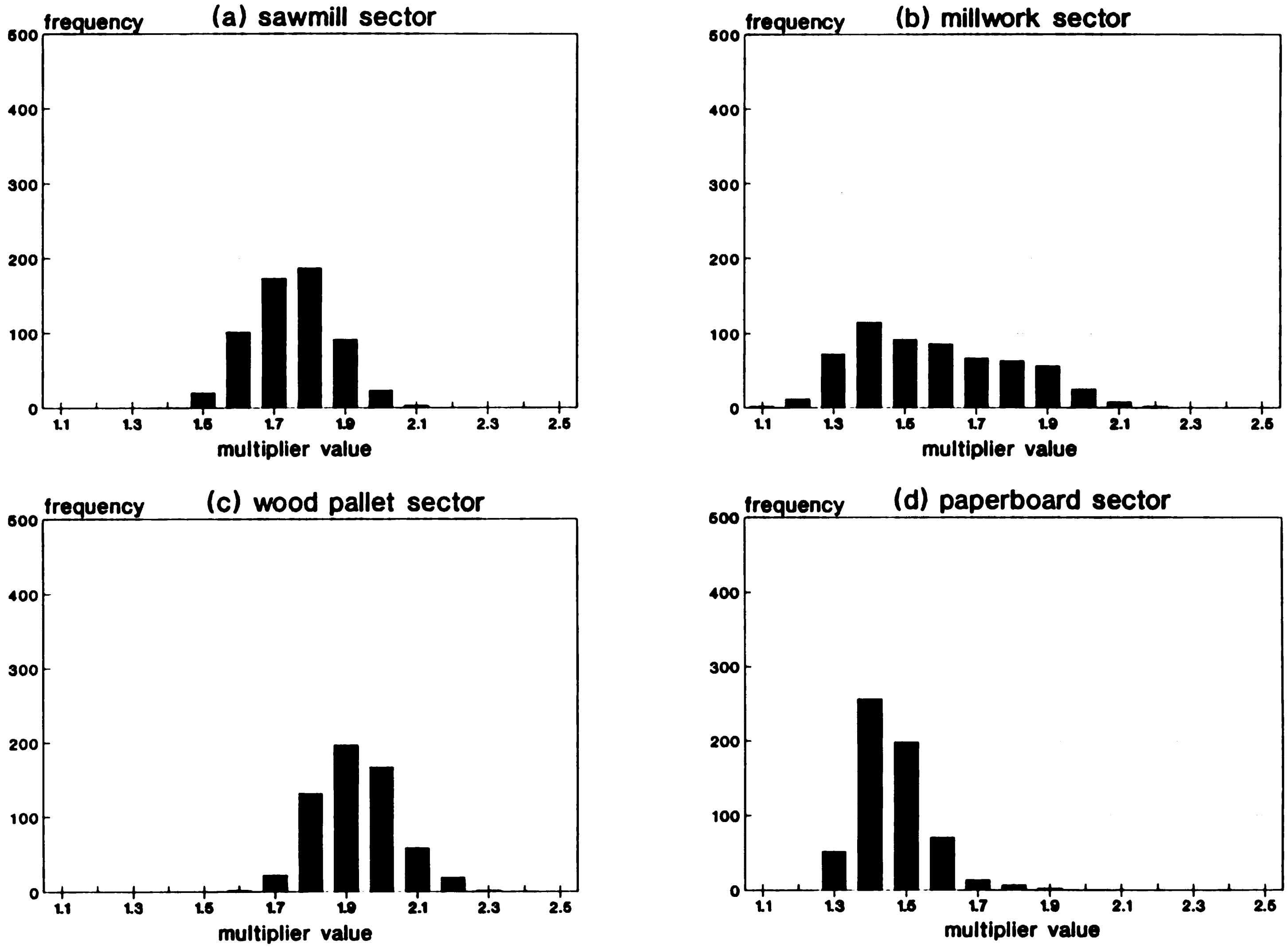

Figure 6. Type I income multiplers based upon 600 bootstrap replications.

important multipliers used. These multipliers use all of the information on the household sector which is quite important in an input-output analysis. The Type I multipliers are presented next in order to present a complete analysis.

Figure 1 contains the histogram for Type II output multipliers. It is very evident that the range of the sawmill sector's multipliers is far less than any of the other sectors. The millwork sector, on the other hand, has its multipliers spread over a very large range due to the fact that the millwork sector is composed of firms engaged in activities representing three separate 4-digit SIC classifications: (1) hardwood dimension and flooring, SIC 2426; (2) millwork, SIC 2431; and structural wood members, not elsewhere classified, SIC 2439. The millwork sector's sample size was also small at only 19 firms sampled, which contributed to the millwork sector's large multiplier range presented here. The wood pallet sector's Type II multipliers varied more than the sawmill sector's multipliers. The paperboard sector's multipliers were a bit more variable than the wood pallet sector but not as much as the millwork sector. Clearly the effort to define the sawmill sector's Type II output multiplier was very successful. Whether the same can be said about the millwork sector depends upon how important a precise estimate of that multiplier is to the ultimate user of the multiplier.

The results for the Type II employment multipliers in Figure 2 are somewhat different. The pattern of the distributions of the sawmill, wood pallet, and paperboard sector are quite similar. The wood pallet sector's distribution is slightly more concentrated than the paperboard sector's distribution. Which, in turn, is slightly more concentrated than the sawmill sector's distribution. The millwork sector's distribution is, again, the most variable.

Figure 3 contains a histogram of the Type II income multiplier. Here the paperboard distribution is the least volatile, followed by the wood pallet sector, then the sawmill sector. Again the millwork sector has the largest range.

A histogram showing the Type I output multiplier is shown in Figure 4. The scale of the histograms showing the Type I multiplier had to be changed from the Type II multiplier because of the extreme concentration of the sawmill sector's Type I multiplier. Three-quarters of all the sawmill's multipliers were concentrated at 1.8. The second least variable sector was the wood pallet sector, followed again by the paperboard sector, and finally the millwork sector.

Figure 5 shows the Type I employment multiplier. The sawmill, wood pallet, and paperboard sectors have very similar distributions. The millwork sector's multipliers are clearly much more variable than the other three sector's multipliers.

Figure 6 contains the Type I income multiplier distribution. The paperboard sector's distribution is more concentrated than the other sectors. The least concentrated sector is the millwork sector, again.

The data are also presented in a tabular format. Table 12 summarizes the data giving the average multiplier for each distribution, the bootstrap estimate of the standard deviation, and values of the multipliers at selected percentiles.

Table 12. Mean Type I and II output, employment, and income multipliers and bootstrap standard deviations for selected sectors.

Type of multiplier	Sector	Mean	Boot-strap s.d.	Percentile 5	10	25	50	75	90	95
Type II output	sawmill	2.763	0.0576	2.667	2.685	2.723	2.762	2.804	2.838	2.863
	millwork	2.315	.2946	1.920	1.992	2.092	2.280	2.510	2.758	2.863
	wood pallet	3.069	.1200	2.881	2.921	2.986	3.068	3.146	3.219	3.270
	paperboard	2.062	.1718	1.809	1.854	1.949	2.042	2.160	2.296	2.356
Type I output	sawmill	1.794	.0442	1.714	1.730	1.770	1.799	1.824	1.845	1.854
	millwork	1.493	.1945	1.240	1.276	1.334	1.473	1.634	1.752	1.825
	wood pallet	1.951	.0704	1.842	1.861	1.900	1.947	1.997	2.044	2.071
	paperboard	1.356	.0839	1.231	1.253	1.296	1.346	1.402	1.471	1.527
Type II employment	sawmill	1.912	.1341	1.708	1.744	1.816	1.911	2.006	2.076	2.117
	millwork	2.103	.2795	1.735	1.773	1.882	2.048	2.274	2.508	2.625
	wood pallet	1.996	.1096	1.824	1.859	1.916	1.991	2.070	2.140	2.189
	paperboard	2.200	.1217	2.003	2.038	2.120	2.191	2.278	2.361	2.400
Type I employment	sawmill	1.559	.1000	1.407	1.432	1.486	1.557	1.629	1.682	1.716
	millwork	1.674	.2351	1.380	1.410	1.492	1.635	1.825	1.992	2.102
	wood pallet	1.666	.0905	1.532	1.560	1.604	1.663	1.720	1.785	1.815
	paperboard	1.640	.0910	1.499	1.525	1.579	1.632	1.699	1.758	1.797
Type II income	sawmill	2.275	.1546	2.031	2.077	2.160	2.274	2.379	2.476	2.524
	millwork	2.057	.3368	1.667	1.717	1.808	2.023	2.278	2.471	2.553
	wood pallet	2.503	.1401	2.280	2.315	2.399	2.497	2.594	2.693	2.749
	paperboard	1.897	.1136	1.740	1.763	1.810	1.878	1.955	2.054	2.120
Type I income	sawmill	1.751	.1190	1.563	1.599	1.663	1.752	1.832	1.906	1.944
	millwork	1.584	.2432	1.283	1.322	1.393	1.558	1.755	1.902	1.966
	wood pallet	1.927	.1078	1.756	1.783	1.847	1.923	1.998	2.073	2.118
	paperboard	1.461	.0873	1.340	1.358	1.394	1.446	1.506	1.582	1.632

The wood pallet sector's average Type II output multiplier is greater than the sawmill sector's multiplier, followed by the millwork sector and trailed by the paperboard sector. Note that the bootstrap estimate of the standard deviation for the wood pallet, paperboard, and millwork sectors are twice, three-times, and five-times as great as the estimate for the sawmill sector.

A similar pattern exists for the Type I output multiplier. The largest average multiplier is the wood pallet sector, followed by the sawmill, millwork and finally paperboard sectors. The bootstrap estimates of the standard deviation follow the same pattern as the Type II multiplier, except the overall estimates are lower. For example the Type I and Type II estimates of standard deviation for the sawmill sector are 0.0442 and 0.0576, respectively. Similar comparisons for the wood pallet, paperboard, and millwork sectors show standard deviations of .0704 and 0.1200, 0.0839 and 0.1718, and 0.1945 and 0.2946, respectively.

The pattern of the Type II employment multiplier is different than the two output multipliers. The largest (average) Type II employment multiplier is found in the paperboard industry followed in descending order by the millwork, wood pallet, and the sawmill sector. Note that the range of multipliers is much narrower than in the output multipliers. The range of the average Type II employment multipliers is from 1.912 in the sawmill sector to 2.200 in the paperboard sector. The range of the average Type II output multipliers is from 2.062 in the paperboard sector to 3.069 in the wood pallet sector. The order of the values of the bootstrap standard deviations was also different from the two output multipliers. The wood pallet sector had the lowest standard deviation at .1096, followed by the paperboard sector (.1217), the sawmill sector (.1341), and finally the millwork sector (.2795).

The Type I employment multipliers were very similar for the four sectors studied. The multipliers ranged from 1.559 in the sawmill sector to 1.674 in the millwork sector, a difference of only 0.115. The bootstrap estimates of the standard deviation followed the same pattern as the Type II employment multipliers in which the wood pallet sector had the lowest value (.0905) and the millwork sector had the highest value (.2351).

The average Type I and Type II income multipliers follow the same order. The wood pallet sector had the highest multiplier, followed by the sawmill sector, the millwork sector, and finally by the paperboard sector. The paperboard sector had the lowest standard deviation, followed in order by the wood pallet, sawmill, and millwork sector, respectively.

The overall average of all bootstrap estimates of standard deviation (i.e., overall all six multipliers) is 0.1459. This value can be divided into two categories, however, the millwork sector average standard deviation is 0.2640, the sawmill, wood pallet, and paperboard sector's stand deviations are 0.1016, 0.1064, and 0.1116, respectively. Clearly, the estimates for the sawmill, wood pallet, and paperboard sectors are superior to the estimates of the millwork sector.

There is a difference in the absolute value of the Type I and II measures of standard deviation. For example, the sawmill, wood pallet, and paperboard sector's estimates of bootstrap standard deviations for the Type I multipliers are almost exactly the same at 0.0877, 0.0896, and 0.0874, respectively. The same measure for the Type II multipliers are 0.1154, 0.1232, and 0.1357, respectively. In contrast, the millwork sector's average bootstrap standard deviations for the Type I and II multipliers are 0.2243 and 0.3036. These comparisons show two things: (1) the sawmill, wood pallet, and paperboard estimates are known with the same precision, and (2) the sawmill, wood pallet, and paperboard estimates are superior to the millwork sector estimates.

Note that the fact that the standard deviation estimates for the Type II multipliers are higher than the Type I multipliers has no significance since the Type II multipliers are larger than the Type I multipliers. The average standard deviation is 7.49 percent of the average Type II multiplier. In contrast, the average standard deviation for all Type I multipliers is 7.39 percent.

Finally, confidence intervals may be constructed from the distribution of multipliers generated. Figure 7 contains the 90 percent Type II output, employment, and income confidence intervals for the four sectors surveyed. Note that the millwork sector's confidence intervals are very wide compared to the other sectors reflecting the uncertainty of the estimates in that sector.

Another factor evident from Figure 7 is that the output values or the four sectors tend to be furthest apart, followed by the income multipliers, and finally the employment multipliers. The narrow differences between the estimates of the employment multipliers show that the analysts cannot distinguish as well between the employment multipliers as they can either the income or output multipliers.

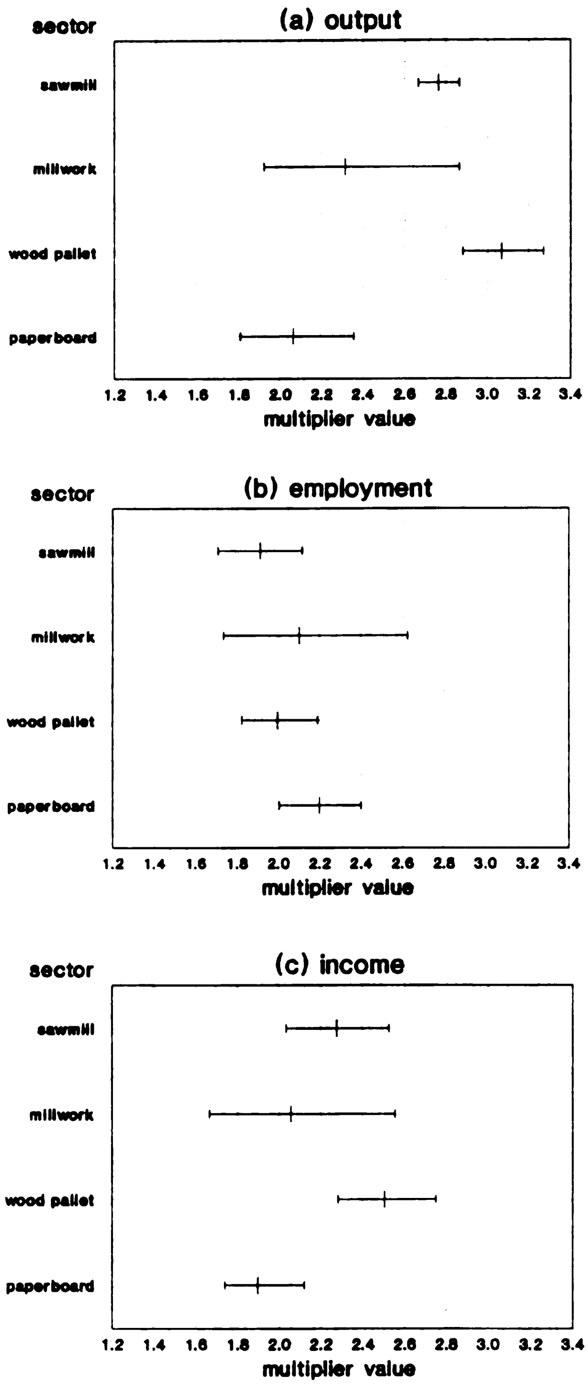

Figure 7. Type II multiplier 90 percent confidence intervals generated using 600 bootstrap replications.

Figure 8 shows the 90 percent confidence intervals for the Type I output, income, and employment multipliers. The message from this comparison is the same as the message from the comparison of the Type II multipliers. The millwork sector's multipliers are not known with very much precision and the employment multipliers are very similar for all four sectors.

Determination of the Effects of Sample Size

The final test conducted in this analysis is a sensitivity analysis of the effects of sample size on the multipliers generated. What is the impact of a reduced sample size on the results? As mentioned in the methods section, three sample sizes were used in the study, full, three-quarters, and one-half.

The full sample used the same sample size in each sector as the original data. The three-quarters sample was designed to measure the effect of "somewhat reduced" effort, the sample size used was three-quarters of the original sample size for each sector. Finally, the one-half sample size measured the effect of "greatly reduced" effort to gather samples. The sample size was one-half of the original sample size. Each trial sample was gathered with replacement from the original sample data base. Details of the procedure are provided above, the sample sizes used for each sector were provided in Table 7.

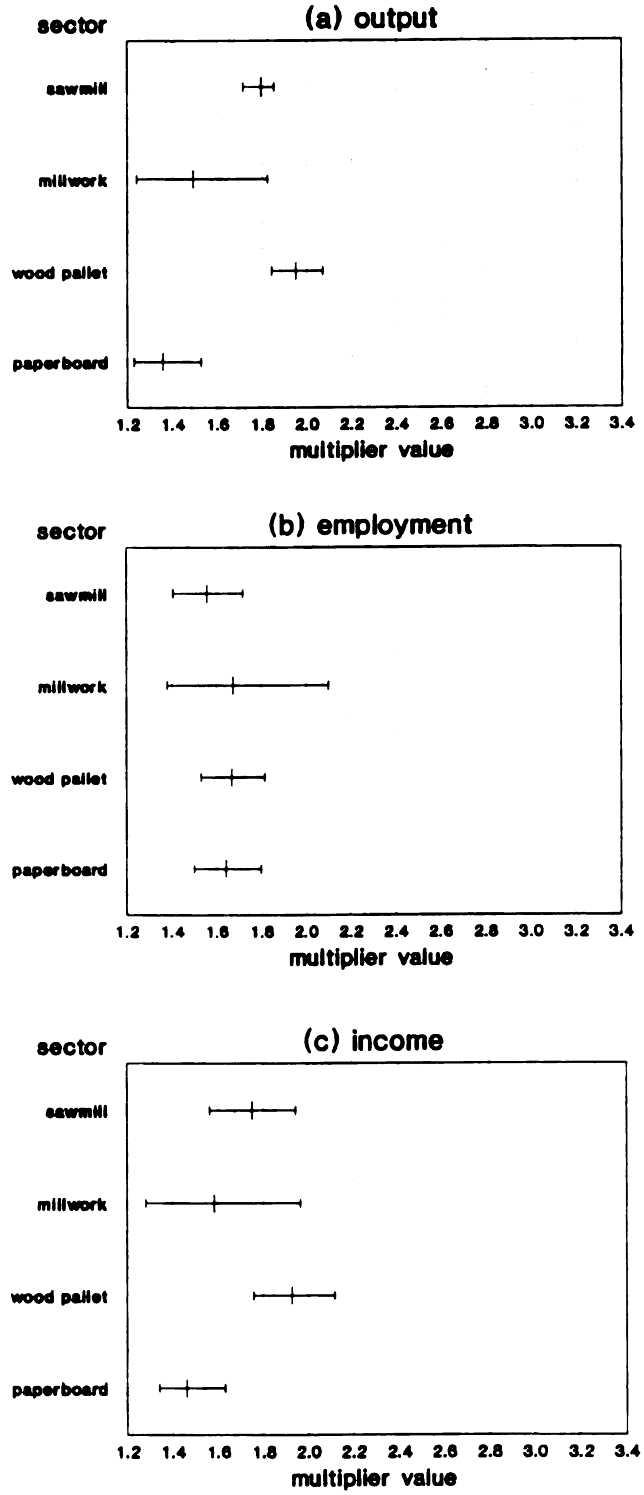

Figure 8. Type I multiplier 90 percent confidence intervals generated using 600 bootstrap replications.

Table 13 shows the mean Type II output multipliers and its associated bootstrap estimate of the standard deviation for each sector and sampling distribution. Both measures were averaged over the 10 trials used in the study. Tables A1 through A4 (in Appendix A) provide the details of the multipliers for each individual trial. Results show that the average multiplier for the sawmill, wood pallet, and paperboard sectors are very similar whether the full, three-quarters, or one-half distributions are examined. For example, the sawmill sector had average multiplier values of 2.761, 2.761, and 2.765 for the full, three-quarters, and one-half sampling distributions, respectively. The data for each individual trial shown in Table A1 reveal that the sawmill sector is not as uniform as the average data suggests. The values of the mean Type II output multiplier of the sawmill sector for the full, three-quarters, and one-half sampling distributions ranged from 2.691 to 2.844, 2.626 to 2.997, and 2.679 to 2.957, respectively.

Table 13. Summary of mean Type II output multiplier values and bootstrap standard deviations for selected sectors using the full, three-quarters, and one-half sample-size distributions.

Sector	Sampling distribution	Mean	Bootstrap standard deviation
sawmills	full	2.761	0.0603
	three-quarters	2.761	.0738
	one-half	2.765	.1036
	average	2.762	.0793
millwork	full	2.328	.2535
	three-quarters	2.423	.2417
	one-half	2.616	.3314
	average	2.456	.2755
wood pallets	full	3.084	.1001
	three-quarters	3.105	.1299
	one-half	3.048	.1527
	average	3.079	.1275
paperboard	full	2.102	.1786
	three-quarters	2.107	.2074
	one-half	2.107	.1862
	average	2.105	.1907

The average bootstrap standard deviations for the sawmill sector are 0.0603, 0.0738, and 0.1036, for the full, three-quarters, and one-half sampling distributions, respectively. Assuming one uses the one-half sampling distribution as a base, increasing the sampling effort to the three-quarters samples decreased, on average, the bootstrap standard deviation by 29 percent. Increasing the sampling effort to the full level, as compared to the one-half sample sizes, decreased the standard deviation by a total of 42 percent.

The millwork sector's mean Type II output multiplier values are not as uniform as the sawmill, wood pallet, or paperboard sectors, but there is a lot more variability in the millwork sector's results in general. The averages are 2.328 for the full sampling distribution, 2.423 for the three-quarters sampling distribution, and 2.616 for the one-half sampling distribution.

The pattern for the wood pallet sector is the same as the sawmill sector, the average multipliers are similar for the three sampling distributions and the bootstrap standard deviations decrease with increasing sample size. Changing the sample size from one-half to three quarters, and from one-half to full, reduced the average bootstrap standard deviations 15 and 34 percent, respectively.

The pattern for the paperboard sector is somewhat different than the pattern for the sawmill sector or wood pallet sector in that the bootstrap standard deviations are similar for the three distributions. The values are 0.1786, 0.2074, and 0.1862 for the full, three-quarters, and one-half sampling distributions, respectively. The variability in the paperboard sector does not decrease with increasing sample size.

Table 14 and 15 summarize the results for the mean Type II employment multipliers and mean Type II income multipliers, respectively. Appendix Tables A5 through A12 in Appendix A contain the detailed results of the analysis. The patterns are the same as the output sectors. The average values for each sampling distribution are very similar for the sawmill, wood pallet and paperboard sectors. The data are more variable for the millwork sector and hence the values of the average multipliers for the three sampling distributions are not as close as the other three sectors studied.

The bootstrap standard deviations for the sawmill, millwork, and wood pallet sectors follow the expected pattern of increasing with decreasing size of the sample. That is not true in the paperboard sector where there is no apparent difference among bootstrap standard deviations for the three levels of sampling intensity for either the employment multiplier or income multiplier.

Table 14 shows that for the sawmill sector's Type II employment multiplier, changing from the one-half to the three-quarters sample size, and changing from the one-half to the full sample size, reduced the bootstrap standard deviations by 31 and 37 percent, respectively.

Similar measures for the millwork sector show 21 and 46 percent reductions in the bootstrap standard deviations.

Table 14. Summary of mean Type II employment multiplier values and bootstrap standard deviations for selected sectors using the full, three-quarters, and one-half sample-size distributions.

Sector	Sampling distribution	Mean	Bootstrap standard deviation
sawmills	full	1.880	0.1235
	three-quarters	1.896	.1337
	one-half	1.871	.1945
	average	1.883	.1506
millwork	full	2.010	.1860
	three-quarters	2.164	.2706
	one-half	2.230	.3437
	average	2.135	.2667
wood pallets	full	2.007	.1044
	three-quarters	2.013	.1562
	one-half	2.054	.1766
	average	2.024	.1457
paperboard	full	2.203	.1346
	three-quarters	2.224	.1412
	one-half	2.161	.1144
	average	2.196	.1301

Table 15. Summary of mean Type II income multiplier values and bootstrap standard deviations for selected sectors using the full, three-quarters, and one-half sample-size distributions.

Sector	Sampling distribution	Mean	Bootstrap standard deviation
sawmills	full	2.215	0.1514
	three-quarters	2.260	.1726
	one-half	2.256	.1874
	average	2.244	.1705
millwork	full	2.008	.1903
	three-quarters	2.085	.2505
	one-half	2.300	.3457
	average	2.131	.2662
wood pallets	full	2.446	.1322
	three-quarters	2.669	.1903
	one-half	2.563	.2262
	average	2.559	.1829
paperboard	full	1.901	.1188
	three-quarters	1.947	.1484
	one-half	1.894	.1381
	average	1.914	.1351

The improvement in the bootstrap standard deviations for the sawmill sector's Type II income multipliers is not as dramatic as the changes for the Type II output or employment multipliers. Table 15 reports, however, that moving from the one-half to three-quarters sample size, reduced the standard deviation by 8 percent. Movement from the one-half to the full sample size reduced the standard deviation by 19 percent. Similar results for the millwork and wood pallet sectors, show improvements of 28 and 45 percent for the millwork sector, and 16 and 42 percent for the wood pallet sector.

Confidence intervals were also generated for each distribution. They are found in Appendix A as Figures A1 through A12. Figure A1, for example, shows 90 percent confidence intervals for the sawmill sector's Type II output multipliers. The results show a tight pattern for the confidence intervals, as expected. And note how much larger the one-half sample confidence intervals are compared to the full sample confidence intervals. Also note Figure A4, that shows the Type II output multiplier for the millwork sector. Note that the confidence intervals are very much wider for the millwork sector than for the sawmill sector.

Use of Study Results

There are two broad categories of users of the results of this analysis, analysts who are contemplating building an input-output modeling system, and decision-makers that want to use these results for the Michigan economy. Analysts may want to use the techniques employed in this analysis in their own studies in order that they may be able to produce confidence intervals for the models that they generate. Modelers studying the forest products sectors could use these results to guide sample collection. Clearly the millwork sector would need more intensive sampling or perhaps disaggregation in future studies. There are no guarantees that the same pattern will hold in other states or in other times, but these results for Michigan should give an indicator of where problems may exist in future studies.

An example may be the best way to show how a decision-maker may use the results of this analysis. Assume that a decision-maker must choose between alternative policies that would yield a change in final demand of $50 million in either the sawmill, millwork, wood pallet, or paperboard sectors. What would the change in total output, income, and employment be in Michigan? What would be the 90 percent confidence intervals about each estimate? In other words, what is our best estimate of the impact, and how much confidence do we have in that prediction?

Table 16 shows the changes in total output caused by an increase in final demand of $50 million. The highest increase occurred in the wood pallet sector, followed by the sawmill sector, millwork, and finally the paperboard sectors. The data also shows that the wood

Table 16. Total output change of the Michigan economy due to a hypothetical increase in final demand for output of sawmill, millwork, wood pallet, and paperboard sectors.

	Sawmill	Millwork	Wood pallet	Paperboard
Impact on output:				
Increase in final demand ($million)	50.0	50.0	50.0	50.0
X average Type II output multiplier	2.763	2.315	3.069	2.062
Total increase in output throughout economy ($million)	138.2	115.8	153.4	103.1
90 percent confidence intervals				
X low multiplier estimate	2.667	1.920	2.881	1.809
X high multiplier estimate	2.863	2.863	3.270	2.356
90 percent confidence interval of total increase in output throughout economy ($million)	133.4-143.2	96.0-143.2	144.0-163.5	90.4-117.8

pallet sector data are known with precision. The range in values are from \$144.0 million to \$163.5 million. The sawmill sector information is also known quite well, the range of the 90 percent confidence intervals was from \$133.4 million to \$143.2 million. The prediction for the millwork sector is more suspect, however, the 90 percent confidence intervals show values ranging from \$96.0 million to \$143.2 million. In other words, total output increase in the millwork sector may be very low or it may be relatively high. The data for the input-output model does not allow us to make any more precise prediction. Finally, data for the paperboard sector show that modest prediction of \$103.1 million is known with more precision than the millwork sector, but less than the sawmill or wood pallet sectors. The 90 percent confidence intervals, \$90.4 million to \$117.8 million, show that the return for the paperboard sector are significantly less than the sawmill or wood pallet sectors.

Table 17 shows results for the total personal income change resulting from a change in final demand of \$50 million in the sawmill, millwork, wood pallet and paperboard sectors. The results are similar to the results for the predicted change in total output. The highest estimated increase in personal income was in the wood pallet sector, followed by the sawmill, millwork, and the paperboard sectors. The 90 percent confidence intervals show that the paperboard data is known with more precision than the other three sectors, followed closely by the sawmill and wood pallet sectors, and trailed again by the millwork sector. The 90 percent confidence intervals are very similar for the sawmill, \$23.3 million to \$29.0 million, and wood pallet sectors, \$24.9 million to \$30.0 million. The range for the millwork sector,

Table 17. Total income change of the Michigan economy due to a hypothetical increase in final demand for output of sawmill, millwork, wood pallet, and paperboard sectors.

	Sawmill	Millwork	Wood pallet	Paperboard
Impact on income:				
Increase in final demand ($million)	50.0	50.0	50.0	50.0
X average personal income direct coefficients ($personal income/$output)	0.22941	.21497	.21830	.20567
Personal income increase in original sector ($million)	11.471	10.749	10.915	10.284
X average Type II personal income multiplier	2.275	2.057	2.503	1.897
Total increase in personal income throughout economy ($million)	26.1	22.1	27.3	19.5
90 percent confidence intervals				
X low multiplier estimate	2.031	1.667	2.280	1.740
X high multiplier estimate	2.524	2.553	2.749	2.120
90 percent confidence interval of total increase in personal income throughout economy ($million)	23.3-29.0	17.9-27.4	24.9-30.0	17.9-21.8

$17.9 million to $27.4 million, is again the highest. The range for the paperboard sector, $18.0 million to $21.8 million, is the narrowest.

Table 18 shows results for total employment change resulting from a change in final demand of $50 million in the sawmill, millwork, wood pallet, and paperboard sectors. The results show the same pattern, as the output, and the income changes. The wood pallet sector shows the largest increase in employment, followed by the sawmill, millwork, and the paperboard sectors. In terms of number of jobs, the data for the paperboard sector indicates a 90 percent confidence interval of 955 jobs to 1,145 jobs, is the most precise. Confidence intervals for the sawmill sector, 1,784 jobs to 2,211 jobs, and the wood pallet sector, 2,376 jobs to 2,851 jobs, are very similar. The least precise data are again from the millwork sector with a range of 1,225 jobs to 1,853 jobs.

If the decision-maker's goal was to improve output, income, or jobs in the Michigan economy, and if all changes in final demand could be obtained with equal effort, the wood pallet sector would be the ideal sector to target. Not only is the wood pallet sector predicted increase in output, income, and employment the highest, the confidence which we can make that prediction is fairly high.

The second best sector to target would be the sawmill sector, since the predicted changes lag the wood pallet sector by only a small amount, and the predicted changes are known with a high degree of precision.

Table 18. Total employment change of the Michigan economy due to a hypothetical increase in final demand for output of sawmill, millwork, wood pallet, and paperboard sectors.

	Sawmill	Millwork	Wood pallet	Paperboard
Impact on employment:				
Increase in final demand ($million)	50.0	50.0	50.0	50.0
X average employment direct coefficients (#jobs/$million output)	20.89	14.12	26.05	9.54
Employment increase in original sector (#jobs)	1044.	706.	1302.	477.
X average Type II employment multiplier	1.912	2.103	1.996	2.220
Total increase in employment throughout economy (#jobs)	1997.	1485.	2600.	1059.
90 percent confidence intervals				
X low multiplier estimate	1.708	1.735	1.824	2.003
X high multiplier estimate	2.117	2.625	2.189	2.400
90 percent confidence interval of total increase in employment throughout economy (#jobs)	1784-2211	1225-1853	2376-2851	955-1145

Results from targeting the millwork sector are uncertain. The level of precision of the estimates are lower than targeting the other three sectors. Efforts to target this sector may be highly rewarding or they may be meager. The 90 percent confidence intervals show results that vary from among the best results to among the worst results. The data, unfortunately, do not support precise estimates of output, income, or employment impacts stemming from a change in final demand in the millwork sector.

The paperboard sector results are lower than the other sectors and the results are known with a fair degree of precision. If the cost and effort required to change the final demand of the four sectors were the same, the paperboard sector would not be a good candidate for targeting. The output, income, and employment effects are lower than the other sectors.

The results of this example assume that the change in final demand is the same for the four sectors studied. Naturally, that assumption may not be realistic. The results could be modified by changing the assumption about sector specific changes in final demand. The new rankings could produce a different order than the order outlined in Tables 16, 17, and 18. For example, assume that a particular set of final demands could produce the same total increase in output ($150 million) for all four sectors, using the average Type II output multipliers. Table 19 shows the results. Although the average predicted impact is the same in the four sectors, the precision in which we know these values are not equivalent. The results for the sawmill sector are very precise, the results for the

Table 19. A hypothetical example where the average total output change for the sawmill, millwork, wood pallet, and paperboard sectors are equal.

	Sawmill	Millwork	Wood pallet	Paperboard
Impact on output:				
Increase in final demand ($million)	54.29	64.79	48.88	72.74
X average Type II output multiplier	2.763	2.315	3.069	2.062
Total increase in output throughout economy ($million)	150.0	150.0	150.0	150.0
90 percent confidence intervals				
X low multiplier estimate	2.667	1.920	2.881	1.809
X high multiplier estimate	2.863	2.863	3.270	2.356
90 percent confidence interval of total increase in output throughout economy ($million)	144.8-155.4	124.4-185.5	140.8-159.8	131.6-171.4

millwork sector are far less precise. The model says that the millwork sector total change may vary from a low of $124.4 million to a high of $185.5, using the 90 percent confidence intervals. The sawmill sector's confidence intervals vary only from a low of $144.8 million to $155.4 million. In the case of the millwork sector, the $150.0 million change in total output is a crude estimate, in the case of the sawmill sector, the $150.0 million change in total output is a relatively precise estimate. The risk-averse decision-maker should implement the policy that increases output in the sawmill sector rather than the millwork sector. The assumption being that the risk-adverse decision-maker would be more uncomfortable with imprecise rather than precise estimates of outcomes. Naturally, many other factors, such as the ability of the area to support additional manufacturing capacity, the development of markets for the products produced, and stability of sales for the different products produced, must also be considered in making a decision concerning which industry to attract to a region.

Hypothesis Evaluation

The hypothesis of this report is that the bootstrap procedure can be effectively and efficiently used to generate confidence intervals and other measures of precision for input-output multipliers.

There were four evaluation criteria: (1) the assumptions of the model used to generate the results must be valid; (2) the results must be useful to analysts; (3) the results must be useful to decision-makers; and (4) the approach must be cost effective.

The analysis by West is currently the best example of an alternative to the bootstrap procedure in the literature. West's technique used two assumptions: (1) the original input data are normally distributed, and (2) the individual direct coefficients are independent. West (1986:page 370) admits that the normality assumption of the original data is one aspect of his study where a "...complete lack of prior information prevails." West also felt that the task of relaxing the independence assumption and developing an appropriate model that recognizes the dependence of the individual direct coefficients, assuming it could be done, is not cost-efficient since "...the cost of increased complexity and data requirements of a more general model ... outweighs the resultant improvement in accuracy (West 1986:page 364)."

The bootstrap procedure provides estimates of the precision of input-output multipliers without any assumptions about the normality of the input coefficients. Since the bootstrap procedure does not require the input data to be normally distributed, nor does it require a model be developed to analyze the data, the bootstrap procedure appears to be superior to the technique used by West when evaluated by the first criterion. Note also that the technique of sampling firms, not individual direct coefficients, means that the bootstrap procedure does not require that individual coefficient data are independent of one another.

The second evaluation criterion is the usefulness to the analyst. Clearly, having information about the precision of input-output multipliers would be an advantage to analysts planning a new input-output study or evaluating a study after its completion. Effort could be directed to sectors where the variability of results were high from sectors where the variability of results were low. The results would be highly useful. Of course, if another appropriate model were to be developed, the results may be just as useful as the bootstrap procedure. An appropriate model is one that does not violate the assumptions underlying input-output analysis.

The third evaluation criterion is the usefulness to decision-makers. Assuming that scarce resources must be used to attract industry to a region, information on the output, employment, and income effects would be important. Since that information is not known with certainty, information on the relative precision of the estimates would be very valuable to such decision-makers. Naturally if such information were available from another appropriate model, the results may be just as useful as the bootstrap procedure.

The fourth evaluation criterion is the cost-effectiveness of the approach. The bootstrap procedure uses very little resources other than what would otherwise be used in a partial survey input-output analysis without estimates of precision. The same survey data that is gathered for the development of the partial survey model is used by the bootstrap procedure. The additional costs are: (1) the cost of development of the SAS procedures, and (2) the computing cost involved. The cost of development of the proper SAS procedures are not excessive since a set of programming steps have to be developed anyway to convert sampled firm data to industry estimates. The only major addition required is the use of a random sampling procedure to generate the bootstrap samples and the effort required to evaluate the results.

The additional computing costs required are not considered significant. The computing time required to generate the multiplier distributions for all Type I and Type II output, employment, and income multipliers was only about 10 hours on an IBM PS/2 Model 80. The sample size, and number of bootstrap replications tests, naturally, took considerably more time to complete but it would not be essential to conduct those analyses in an operational setting. If a model were developed with a large number of sectors, a more powerful computer might be required to conduct the analysis.

If West is correct about the cost and difficulty of developing an appropriate parametric model to analyze the precision of input-output multipliers, the bootstrap procedure has an advantage. It is assumed that the cost of the theoretical work would be far more than the minor expense of implementing the bootstrap procedure since the bootstrap requires no additional theoretical work prior to its use.

In summary, the bootstrap procedure: (1) does not rely upon questionable assumptions about the distributions of the input data; (2) provides results that are useful to analysts; (3) provides results that are useful to decision-makers; and (4) requires costs that are not excessive. Based upon this evidence, the hypothesis of this report is accepted, the bootstrap procedure has been found to be effective and efficiently used to generate measures of precision for multipliers from partial survey input-output models.

CHAPTER 5

SUMMARY AND CONCLUSIONS

The parameters of input-output models are not known with certainty. Analysts compiling the data to generate an input-output model take samples from the firms in the region they are studying and combine these data with the non-survey information they gather to estimate economic interrelationships within the economy. They build input-output models, and generate output, income, and employment multipliers based upon that model. The analysts present these multipliers, generally, without estimates of their reliability due to the lack of theory linking errors or variability of the input data and the input-output model generated. There is also a complete lack of information about the statistical error associated with the secondary data used to build many input-output models.

Early efforts at estimating stochastic input-output models were undertaken by Evans (1954), Quandt (1958, 1959), Gerking (1976a, 1976b), and Ives (1977), among others. Two recent works that are particularly promising were completed by Jackson (1986) and West (1986).

Jackson (1986) argued that multipliers with confidence intervals may be obtained by simulation. He assumes that each coefficient in the A matrix (a_{ij}) may has its own probability density function (pdf) that may be estimated for it. Each direct coefficient is drawn randomly from its particular pdf, once the entire A matrix is complete, the multipliers are generated in the usual fashion. If this tactic is used a large number of times, multipliers and their associated confidence intervals result.

The key problem with Jackson's approach is the fact that the a_{ij} are drawn independently. Data for each sector represent purchases from other firms in the economy, purchases of labor and net imports into the region. The sum of the entire column of coefficients must be equal to 1.00, therefore if a_{11} is the value for one coefficient, all other sectors must sum to $(1-a_{11})$. The data are not independent. Some firms in a region may be labor intensive, others may be capital intensive. It is not correct to mix coefficients of the two types of firms at random. It is proper to find some weighted average production function of all the firms in the region, as done in this analysis.

West (1986) analytically estimated the probability density function of the input-output multipliers assuming that the a_{ij} were normally distributed and independent. He suggested that the costs of increased complexity and the data requirements of a more general model outweigh any resulting improvement in accuracy. The two key problems with this approach is that it is incorrect to assume that the direct coefficients are independent using the argument presented in the previous paragraph, and there is no assurance that the distribution of a_{ij} are normal.

Fortunately, the bootstrap method is a calculation intensive technique that does not rely upon the assumption that the data are normally distributed. The bootstrap procedure is a means of estimating the statistical precision of a measure from a single sample of the data. The procedure, using Monte Carlo simulation, were used to generate all confidence intervals for this analysis.

The primary objective of the report is to test the effectiveness of using the bootstrap technique to generate estimates of the statistical precision of input-output multipliers using a partial-survey input-output model. The statistical precision was illustrated by: (1) providing histograms of the distribution of each multiplier, (2) providing estimates of the average multiplier, the bootstrap estimate of standard deviation, and values at key points of the distribution of multipliers, and (3) providing confidence intervals for each multiplier. Michigan was chosen for a study region since a partial-survey input-output model that highlighted the forestry sectors of the economy was available for analysis.

The technique is calculation intensive, assuming that 2 matrix inversions were required to estimate the Type I and Type II multipliers without estimation of the reliability of the results. The model presented here, with estimates of the reliability of the results, required 1,200 matrix inversions. The test on sample size required 1,200 matrix inversions for each of the 30 trials used making a total of 36,000 matrix inversions required to complete that test. Such large scale computing requirements would have been prohibitively expensive in the past. The availability of large scale computer programs such as SAS for the personal computer environment, and the proliferation of powerful personal computers means that computer-intensive techniques are feasible. The key costs involved are the purchase of the computer, the license fee for the software, the development costs of producing the computer programs, the input of the data, and the processing of the results. The actual execution of the program cost nothing since the programs were executed at night when the personal computer would ordinarily be unoccupied. It makes little difference whether the program executes for 2 hours or for 10 hours in this sort of environment.

Surveyed firm data were used as the unit of sampling in this analysis. Each individual surveyed firm provided an estimate of total output, income, employment, and distribution of purchases. These data were kept intact throughout this analysis. Each bootstrap replication generated an A matrix that represented a weighted average of the coefficients of the sampled firms for each sector. The weights were provided by total output. If both labor-intensive and capital-intensive firms were represented in a sample from a sector, the a_{ij} for the sector would be the weighted average of the coefficients from each firm sampled.

Survey data on individual firms were the primary source of data for the key sectors of the input-output model, non-survey data was used in all other non-key sectors studied. There was no information on the variability of these non-survey sectors, so all variability comes from the surveyed sectors. Data were collected from firms in the economy of Michigan, but of course, not all firms in the state are represented in the data set. All variability represented in these results represents variability in the survey data, it does not represent difference between the sampled and unsampled populations. Those differences remain unknown. Note however, that statistical tests regarding this question are noted in Chappelle et al. (1986).

This analysis used the total output, income, employment, and distribution of purchases information from each sampled firm to estimate the required coefficients. Each firm was drawn into the bootstrap sample using independent random sampling, of course.

Three tests were conducted in this analysis: (1) determination of the number of bootstrap replications required to define the multiplier distributions; (2) the generation and display of the multiplier distributions, estimation of the statistical precision of the multipliers, and development of confidence intervals for each multiplier; and (3) a sensitivity analysis of the effects of sample size on the results.

The first test conducted was to determine the number of bootstrap replications required to define the multiplier distributions. Although 200 replications were found to be adequate for the Type II output multipliers examined, in order to be conservative, 600 replications were used throughout this analysis.

The second test is the heart of this analysis. The information about the multipliers were displayed using three techniques: (1) histograms of the distribution of multipliers, (2) a table summarizing key values, and (3) a figure showing confidence intervals for each multiplier. Multiplier distributions were displayed in Figures 1 through 6. The figures are histograms showing the multipliers for each bootstrap replication. Table 12 lists the average multiplier, the associated measure of statistical precision, the bootstrap standard deviation, and multiplier values at key points in the distribution. Finally, Figures 7 and 8 show 90 percent confidence intervals for each multiplier calculated.

The results vary depending upon both the sector and the multiplier examined. The output multipliers clearly show that the data for the sawmill sector generates very precise estimates. The estimate of bootstrap standard deviation is by far the smallest of all multipliers and sectors studied in this analysis. The next best sector for the output multiplier is the wood pallet sector followed by the paperboard sector.

The results for the employment multipliers were different than the results for the output multipliers. The employment multipliers are very close, with the wood pallet sector providing the best estimate, followed by the paperboard sector, then the sawmill sector.

The results for the income multipliers show that again the reliability of the multipliers were similar, the paperboard sector edged out the wood pallet sector, followed closely by the sawmill sector.

Another pattern emerged from the analysis, the millwork sector had by far the worst multiplier estimates for the four sectors studied. The average bootstrap standard deviation as a percentage of the sawmill sector's output, employment, and income multiplier is 5.26 percent. The corresponding values for the wood pallet, paperboard, and millwork sectors are 4.94, 6.26, and 14.14 percent, respectively. Whether the variability in the millwork sector is acceptable depends upon the needs of the user of the multipliers, but this analysis clearly shows there is a difference between the millwork sector and the other sectors studied.

The third test is a sensitivity test examining the effects of sample size on the results reported. Three levels of sampling intensity were examined, the full, three-quarters, and one-half. The full level used the same sectoral sample sizes as used in the second test. Each trial was started by a selection with replacement from the data base of a sample for the trial. The bootstrap procedure used the trial sample to generate all multipliers. The three-quarters level was used to model a somewhat reduced level of sampling intensity, each trial selected an initial sample using three-quarters of the number of samples as the full sample. The one-half level modeled a greatly reduced level of sampling intensity, a sample size of one-half the full level was used to select an initial sample for each trail.

The results show that sample sizes do matter. The standard deviation for the sawmill's sector Type II output multiplier was reduced by 29 percent when the three-quarters sampling level rather than the one-half sampling level was used. The full sampling level provided a 42 percent decrease in the standard deviation compared to the one-half sampling level. Similar results were displayed in the millwork and wood pallet sector. In fact the millwork sector showed the biggest improvement of all of the sectors studied.

In summary, the bootstrap procedure: (1) does not rely upon questionable assumptions about the distributions of the input data; (2) provides results that are useful to analysts; (3) provides results that are useful to decision-makers; and (4) does not require excessive cost. Therefore, the bootstrap procedure has been found to be effective and efficiently used to generate measures of precision for multipliers from partial survey input-output models.

Future Research

There are two areas of additional research that seem to be very important, they are: (1) additional techniques to generate non-parametric confidence intervals; and (2) an evaluation of the effects of aggregation using the bootstrap technique. Efron (1982) contains other techniques that may be used to generate bootstrap confidence intervals, the most promising one is the jackknife technique. That technique could be applied to a data set with its results compared to the bootstrap technique.

A second area of research is an evaluation of the effects of aggregation using the bootstrap technique. The millwork sector was the only sector examined in this paper that was aggregated from three 4-digit SIC sectors. The sawmill, wood pallet, and paperboard sectors all are single 4-digit SIC sectors. Since the results for the millwork sector were worse than the other sectors, it would be interesting to examine the effects of aggregation has upon the confidence intervals calculated. Could the precision of the multiplier estimates in the millwork sector be improved by disaggregating by 4-digit SIC industry or are the firms in this sector inherently diverse?

LIST OF REFERENCES

LIST OF REFERENCES

Brown, Douglas M.; Giarratani, Frank. 1979. Input-output as a simple econometric model: a comment. The Review of Economics and Statistics. 61(4):621-623.

Chappelle, Daniel E.; Heinen, Suzanne E.; James, Lee M.; Kittleson, Kyle M.; Olson, David D. 1986. Economic impacts of Michigan forest industries: a partial survey-based input-output study. Research Report 472. East Lansing, MI: Michigan State University Agricultural Experiment Station. 17 p.

Diaconis, Persi; Efron, Bradley. 1983. Computer-intensive methods in statistics. Scientific American. 248(5):116-126,128,130.

Efron, B. 1979. Bootstrap methods: another look at the jackknife. The Annals of Statistics. 7(1):1-26.

Efron, Bradley. 1982. The jackknife, the bootstrap and other resampling plans. Philadelphia, PA: Society for Industrial and Applied Mathematics. 92 p.

Evans, W. Duane. 1954. The effect of structural matrix errors on inter-industry relations estimates. Econometrica. 22(4):461-480.

Garhart, Robert, Jr. 1985. The role of error structure in simulations on regional input-output analysis. Journal of Regional Science. 25(3):353-366.

Gerking, S.D. 1976a. Estimation of stochastic input-output models: some statistical problems. Leiden: Martinus Nijhoff. 87 p.

Gerking, Shelby D. 1976b. Input-output as a simple econometric model. The Review of Economics and Statistics. 58(3):274-282.

Gerking, Shelby D. 1979a. Reconciling reconciliation procedures in regional input-output analysis. International Regional Science Review. 4(1):23-36.

Gerking, Shelby D. 1979b. Input-output as a simple econometric model: reply. The Review of Economics and Statistics. 61(4):623-626.

Gerking, Shelby D.; Pleeter, Saul. 1977. Minimum variance sampling in input-output analysis. The Review of Regional Studies. 7(1):59-80.

Hanseman, D.J. 1982. Stochastic input-output analysis: a simulation study. Environment and Planning A. 14(11):1425-1435.

Hanseman, Dennis J.; Gustafson, Elizabeth F. 1981. Stochastic input-output analysis: a comment. The Review of Economics and Statistics. 63(3):468-470.

Hansen, Morris H.; Hurwitz, William N.; Madow, William G. 1953. Sample survey methods and theory: methods and applications. New York: John Wiley & Sons. 638 p. Vol 1.

Heinen, Suzanne E. 1982. Methods and preliminary survey results of the Michigan forest products industry input-output study. Unpublished paper. East Lansing, MI: Michigan State University.

Ives, Edward Ewing. 1977. Input-output models estimated from primary data: sampling considerations and parameter variability. PhD diss. Corvallis, OR: Oregon State University. 193 p.

Jackson, Randall W. 1986. The full-distribution approach to aggregate representation in the input-output modeling framework. Journal of Regional Science. 26(3):515-531.

Jensen, R.C.; Macdonald, S. 1982. Technique and technology in regional input-output. The Annals of Regional Science. 16(2):27-45.

Karaska, Gerald J. 1968. Variation of input-output coefficients for different levels of aggregation. Journal of Regional Science. 8(2):215-227.

Katz, Joseph L.; Burford, Roger L. 1981. The effect of aggregation on the output multipliers in input-output models. The Annals of Regional Science. 15(3):46-54.

McCamley, Francis; Schreiner, Dean; Muncrief, George. 1973. A method for estimating the sampling variances of multipliers derived from a from-to model. The Annals of Regional Science. 7(2):81-89.

Michigan Energy Administration. 1980. A 44 sector input/output model of the Michigan economy. Energy Issue Analysis Series. Lansing, MI: Energy Administration, Michigan Department of Commerce. 20 p.

Miernyk, William H. 1965. The elements of input-output analysis. New York: Random House. 156 p.

Miernyk, William H. 1976. Comments on recent developments in regional input-output analysis. International Regional Science Review. 1(2):47-55.

Miernyk, William H. 1979. Reconciling reconciliation procedures in regional input-output analysis: comment. International Regional Science Review. 4(1):36-38.

Miernyk, William H.; Shellhammer, Kenneth L.; Brown, Douglas M.; Coccari, Ronald L.; Gallagher, Charles J.; Wineman, Wesley H. 1970. Simulating regional economic development. Lexington MA: D.C. Heath & Co. 337 p.

Miller, Ronald E.; Blair, Peter D. 1985. Input-output analysis: foundations and extensions. Englewood Cliffs, NJ: Prentice-Hall, Inc. 464 p.

Park, Se-Hark; Mohtahi, Malek; Kubursi, Atif. 1981. Errors in regional nonsurvey input-output models: analytical and simulation results. Journal of Regional Science. 21(3):321-339.

Quandt, R.E. 1958. Probabilistic errors in the Leontief system. Naval Research Logistics Quarterly. 5(2):155-70.

Quandt, Richard E. 1959. On the solutions of probabilistic Leontief systems. Naval Research Logistics Quarterly. 6(4):295-305.

Richardson, Harry W. 1972. Input-output and regional economics. New York: A Halsted Press Book, John Wiley and Sons. 294 p.

Richardson, Harry W. 1985. Input-output and economic base multipliers: looking backward and forward. Journal of Regional Science. 25(4):607-661.

Stevens, Benjamin H.; Trainer, Glynnis A. 1978. Error generation in regional input-output analysis and its implications for non-survey models. Regional Science Research Institute Discussion Paper 108. Amherst MA: Regional Science Research Institute. 33 p.

West, Guy R. 1986. A stochastic analysis of an input-output model. Econometrica. 54(2):363-374.

APPENDICES

APPENDIX A

Table A1. Mean Type II output multiplier value, bootstrap standard deviation, and multiplier values for the sawmill sector, for 10 bootstrap trials using the full, three-quarters, and one-half samples.

Bootstrap trial	Mean	Bootstrap s.d.	Percentile						
			5	10	25	50	75	90	95
full sample									
1	2.711	0.0470	2.616	2.649	2.680	2.718	2.747	2.772	2.784
2	2.774	.0442	2.707	2.720	2.744	2.771	2.798	2.836	2.859
3	2.740	.0301	2.689	2.701	2.719	2.741	2.761	2.777	2.790
4	2.771	.0505	2.692	2.710	2.733	2.767	2.799	2.840	2.866
5	2.771	.1014	2.603	2.649	2.702	2.774	2.840	2.893	2.929
6	2.810	.0574	2.728	2.742	2.770	2.803	2.849	2.885	2.919
7	2.753	.0858	2.610	2.638	2.694	2.756	2.810	2.854	2.890
8	2.844	.0610	2.758	2.773	2.798	2.837	2.884	2.922	2.953
9	2.741	.0471	2.655	2.678	2.710	2.742	2.772	2.800	2.818
10	2.691	.0790	2.568	2.594	2.641	2.686	2.741	2.790	2.826
avg.	2.761	.0603							
three-quarters sample									
1	2.759	.0915	2.611	2.642	2.699	2.759	2.817	2.874	2.908
2	2.997	.0870	2.845	2.885	2.931	2.995	3.058	3.113	3.116
3	2.626	.0763	2.489	2.528	2.577	2.663	2.679	2.723	2.749
4	2.796	.0683	2.687	2.710	2.748	2.802	2.843	2.876	2.891
5	2.720	.0876	2.582	2.612	2.660	2.717	2.780	2.835	2.873
6	2.798	.0611	2.702	2.722	2.754	2.793	2.836	2.881	2.911
7	2.707	.0883	2.559	2.594	2.647	2.705	2.768	2.823	2.854
8	2.737	.0295	2.685	2.697	2.718	2.737	2.735	2.775	2.790
9	2.799	.0741	2.679	2.705	2.748	2.796	2.845	2.902	2.932
10	2.671	.0745	2.531	2.570	2.625	2.671	2.721	2.762	2.798
avg.	2.761	.0738							
one-half sample									
1	2.712	.1424	2.467	2.529	2.620	2.705	2.813	2.890	2.946
2	2.744	.1117	2.554	2.599	2.668	2.743	2.818	2.894	2.924
3	2.697	.0744	2.558	2.584	2.647	2.698	2.742	2.805	2.840
4	2.749	.1234	2.548	2.592	2.663	2.750	2.834	2.904	2.953
5	2.679	.0997	2.536	2.554	2.610	2.669	2.736	2.805	2.861
6	2.726	.0969	2.565	2.604	2.661	2.725	2.798	2.846	2.875
7	2.957	.1003	2.809	2.839	2.890	2.945	3.023	3.083	3.117
8	2.774	.0942	2.632	2.658	2.707	2.769	2.829	2.898	2.959
9	2.883	.0754	2.780	2.794	2.826	2.874	2.933	2.989	3.014
10	2.735	.1174	2.538	2.582	2.648	2.743	2.818	2.880	2.927
avg.	2.765	.1036							
overall avg.	2.762	.0793							

Table A2. Mean Type II output multiplier value, bootstrap standard deviation, and multiplier values for the millwork sector, for 10 bootstrap trials using the full, three-quarters, and one-half samples.

Bootstrap trial	Mean	Bootstrap s.d.	Percentile 5	10	25	50	75	90	95
full sample									
1	2.747	0.0959	2.589	2.627	2.681	2.743	2.807	2.871	2.916
2	2.461	.2654	2.060	2.136	2.201	2.271	2.471	2.641	2.788
3	2.478	.2545	2.064	2.163	2.305	2.493	2.663	2.776	2.820
4	2.267	.4351	1.818	1.862	1.956	2.113	2.713	2.813	2.874
5	2.202	.3111	1.838	1.885	1.970	2.116	2.391	2.702	2.807
6	2.032	.2091	1.747	1.787	1.861	1.972	2.151	2.344	2.456
7	2.192	.2573	1.834	1.880	1.972	2.120	2.331	2.731	2.835
8	2.431	.2124	2.100	2.164	2.283	2.433	2.564	2.692	2.777
9	2.530	.3416	2.083	2.158	2.305	2.470	2.781	2.989	3.060
10	1.941	.1528	1.738	1.754	1.804	1.887	2.012	2.156	2.368
avg.	2.328	0.2535							
three-quarters sample									
1	2.256	.4438	1.792	1.833	1.919	2.075	2.670	2.827	2.916
2	2.803	.1561	2.417	2.504	2.721	2.840	2.929	2.995	3.026
3	2.423	.2738	2.001	2.074	2.231	2.430	2.624	2.757	2.810
4	2.290	.2927	1.897	1.953	2.079	2.248	2.467	2.696	2.794
5	2.279	.3960	1.845	1.895	1.990	2.150	2.638	2.775	2.841
6	1.983	.1600	1.753	1.791	1.849	1.933	2.066	2.221	2.359
7	2.882	.1485	2.561	2.652	2.787	2.911	2.991	3.059	3.123
8	1.971	.1683	1.750	1.776	1.838	1.924	2.055	2.203	2.336
9	2.472	.2944	2.054	2.115	2.268	2.453	2.698	2.829	2.903
10	2.869	.0829	2.732	2.766	2.811	2.967	2.926	2.976	3.003
avg.	2.423	.2417							
one-half sample									
1	2.179	.3468	1.762	1.815	1.897	2.048	2.300	2.921	3.223
2	2.809	.1270	2.595	2.634	2.727	2.828	2.900	2.949	2.976
3	2.822	.1191	2.627	2.680	2.738	2.812	2.893	2.989	3.049
4	2.490	.3768	1.941	2.020	2.223	2.500	2.775	2.939	3.007
5	2.536	.5463	1.766	1.820	1.944	2.179	2.918	3.006	3.041
6	2.901	.1682	2.582	2.667	2.797	2.923	3.025	3.100	3.152
7	2.922	.1551	2.617	2.686	2.837	2.958	3.035	3.083	3.110
8	2.413	.5712	1.756	1.807	1.953	2.214	2.988	3.056	3.103
9	2.564	.3573	1.988	2.099	2.316	2.578	2.828	2.990	3.058
10	2.527	.5460	1.965	2.022	2.130	2.365	3.109	3.219	3.285
avg.	2.616	.3314							
overall avg.	2.456	.2755							

Table A3. Mean Type II output multiplier value, bootstrap standard deviation, and multiplier values for the pallet sector, for 10 bootstrap trials using the full, three-quarters, and one-half samples.

Bootstrap trial	Mean	Bootstrap s.d.	Percentile 5	10	25	50	75	90	95
			full sample						
1	2.997	0.0758	2.878	2.904	2.942	2.995	3.045	3.099	3.131
2	3.196	.0959	3.037	3.073	3.131	3.190	3.259	3.324	3.355
3	3.145	.1009	2.978	3.022	3.081	3.141	3.215	3.268	3.299
4	3.067	.0969	2.900	2.941	3.005	3.069	3.132	3.188	3.215
5	3.029	.1053	2.868	2.898	2.955	3.026	3.101	3.176	3.206
6	3.039	.1032	2.865	2.905	2.968	3.039	3.113	3.178	3.212
7	3.103	.1696	2.857	2.897	2.985	3.094	3.213	3.323	3.396
8	2.960	.0799	2.831	2.860	2.907	2.956	3.009	3.069	3.106
9	3.059	.0934	2.919	2.950	2.997	3.046	3.115	3.178	3.227
10	3.245	.0800	3.111	3.139	3.191	3.251	3.305	3.343	3.373
avg.	3.084	0.1001							
			three-quarters sample						
1	3.085	.1135	2.907	2.940	3.011	3.084	3.153	3.219	3.256
2	3.325	.1136	3.114	3.184	3.241	3.325	3.401	3.471	3.506
3	3.230	.1228	3.018	3.069	3.154	3.236	3.313	3.384	3.427
4	3.189	.1132	2.993	3.040	3.114	3.192	3.272	3.325	3.356
5	2.888	.1730	2.655	2.683	2.751	2.852	2.974	3.147	3.278
6	2.809	.1266	2.638	2.660	2.719	2.782	2.875	2.997	3.070
7	3.045	.1386	2.819	2.870	2.943	3.047	3.143	3.222	3.255
8	3.187	.0942	3.022	3.065	3.126	3.189	3.246	3.311	3.358
9	3.312	.2193	3.003	3.050	3.160	3.306	3.450	3.581	3.664
10	2.983	.0836	2.853	2.879	2.922	2.979	3.045	3.091	3.137
avg.	3.105	.1299							
			one-half sample						
1	3.206	.2081	2.939	2.981	3.059	3.164	3.326	3.182	3.587
2	3.209	.1826	2.965	3.002	3.071	3.194	3.326	3.458	3.522
3	3.053	.1481	2.812	2.861	2.952	3.047	3.149	3.255	3.315
4	2.939	.1366	2.736	2.775	2.841	2.916	3.018	3.132	3.220
5	2.775	.1378	2.573	2.598	2.679	2.765	2.858	2.960	3.004
6	3.090	.1803	2.790	2.868	2.976	3.085	3.218	3.324	3.349
7	2.937	.1761	2.681	2.721	2.810	2.909	3.029	3.175	3.272
8	2.851	.1485	2.631	2.665	2.742	2.839	2.939	3.047	3.122
9	3.034	.1143	2.853	2.891	2.948	3.036	3.111	3.183	3.230
10	3.386	.0942	3.217	3.259	3.323	3.397	3.457	3.503	3.530
avg.	3.048	.1527							
overall avg.	3.079	.1275							

Table A4. Mean Type II output multiplier value, bootstrap standard deviation, and multiplier values for the paperboard sector, for 10 bootstrap trials using the full, three-quarters, and one-half samples.

Bootstrap trial	Mean	Bootstrap s.d.	Percentile						
			5	10	25	50	75	90	95
full sample									
1	2.360	0.2841	1.986	2.046	2.144	2.332	2.558	2.740	2.810
2	2.354	.2119	2.062	2.114	2.208	2.342	2.472	2.629	2.724
3	2.240	.2377	1.880	1.955	2.086	2.230	2.419	2.535	2.599
4	1.636	.0911	1.499	1.522	1.567	1.629	1.689	1.755	1.806
5	1.950	.1166	1.778	1.807	1.866	1.939	2.018	2.103	2.173
6	2.097	.1441	1.906	1.930	1.982	2.072	2.188	2.296	2.376
7	2.289	.2794	1.895	1.962	2.094	2.268	2.497	2.638	2.704
8	2.342	.2311	1.978	2.059	2.185	2.332	2.498	2.629	2.673
9	1.959	.1203	1.756	1.803	1.873	1.947	2.029	2.141	2.196
10	1.797	.0697	1.693	1.710	1.748	1.791	1.837	1.888	1.917
avg.	2.102	0.1786							
three-quarters sample									
1	1.995	.1962	1.715	1.758	1.851	1.979	2.099	2.262	2.345
2	2.134	.2279	1.776	1.859	1.974	2.113	2.298	2.441	2.498
3	2.147	.1943	1.908	1.942	2.012	2.096	2.240	2.436	2.553
4	2.187	.2891	1.810	1.867	1.960	2.113	2.339	2.704	2.851
5	2.053	.1844	1.796	1.844	1.919	2.014	2.145	2.308	2.463
6	2.185	.2506	1.821	1.876	2.023	2.175	2.360	2.510	2.563
7	1.911	.1142	1.747	1.774	1.824	1.889	1.974	2.082	2.170
8	1.859	.1254	1.676	1.706	1.761	1.839	1.928	2.035	2.109
9	2.187	.1499	1.967	2.006	2.070	2.173	2.279	2.401	2.447
10	2.414	.3426	2.006	2.064	2.167	2.356	2.687	2.861	2.921
avg.	2.107	.2074							
one-half sample									
1	1.784	.1067	1.607	1.622	1.684	1.757	1.846	1.959	2.042
2	2.194	.1863	1.925	1.969	2.053	2.151	2.303	2.463	2.594
3	2.187	.1990	1.903	1.953	2.051	2.177	2.315	2.420	2.504
4	2.062	.1288	1.868	1.902	1.967	2.050	2.133	2.233	2.319
5	1.956	.2012	1.651	1.693	1.783	1.926	2.074	2.209	2.465
6	2.015	.1687	1.795	1.814	1.864	1.946	2.093	2.265	2.538
7	2.690	.2062	2.461	2.485	2.536	2.622	2.802	2.996	3.109
8	2.204	.4029	1.810	1.863	1.910	2.033	2.549	2.750	2.828
9	2.018	.1550	1.818	1.840	1.885	1.976	2.102	2.258	2.330
10	1.958	.1071	1.792	1.823	1.875	1.938	2.021	2.120	2.189
avg.	2.107	.1862							
overall avg.	2.105	.1907							

Table A5. Mean Type II employment multiplier value, bootstrap standard deviation, and multiplier values for the sawmill sector, for 10 bootstrap trials using the full, three-quarters, and one-half samples.

Bootstrap trial	Mean	Bootstrap s.d.	Percentile 5	10	25	50	75	90	95
full sample									
1	1.973	.1554	1.721	1.767	1.872	1.978	2.079	2.164	2.212
2	1.856	.1234	1.684	1.712	1.765	1.823	1.933	2.013	2.069
3	1.942	.1255	1.736	1.782	1.857	1.941	2.029	2.106	2.140
4	2.096	.1380	1.848	1.905	2.007	2.104	2.194	2.258	2.307
5	1.756	.1265	1.575	1.609	1.662	1.744	1.832	1.927	1.988
6	1.875	.1342	1.663	1.712	1.781	1.869	1.953	2.056	2.099
7	1.701	.0830	1.578	1.595	1.641	1.698	1.756	1.816	1.850
8	2.072	.1521	1.823	1.878	1.972	2.074	2.173	2.270	2.320
9	1.858	.1128	1.674	1.720	1.774	1.849	1.938	2.016	2.069
10	1.668	.0845	1.534	1.559	1.606	1.660	1.723	1.777	1.832
avg.	1.880	0.1235							
three-quarters sample									
1	1.702	.0868	1.567	1.593	1.642	1.697	1.759	1.812	1.852
2	1.715	.1048	1.550	1.583	1.638	1.708	1.784	1.856	1.901
3	1.916	.1595	1.658	1.712	1.804	1.914	2.026	2.123	2.176
4	1.907	.0961	1.752	1.782	1.837	1.905	1.969	2.033	2.073
5	1.885	.1356	1.700	1.724	1.789	1.874	1.967	2.055	2.107
6	1.747	.0968	1.586	1.624	1.678	1.746	1.807	1.871	1.920
7	1.941	.1697	1.676	1.725	1.821	1.949	2.054	2.153	2.193
8	2.117	.1384	1.832	1.899	2.019	2.125	2.230	2.318	2.351
9	2.146	.1817	1.827	1.907	2.032	2.151	2.280	2.373	2.425
10	1.890	.1676	1.634	1.675	1.772	1.885	1.993	2.112	2.177
avg.	1.896	.1337							
one-half sample									
1	1.798	.1415	1.579	1.628	1.693	1.786	1.893	1.981	2.035
2	1.836	.1414	1.629	1.667	1.741	1.822	1.925	2.018	2.073
3	2.008	.1750	1.711	1.778	1.890	2.012	2.133	2.234	2.290
4	1.674	.1080	1.518	1.541	1.595	1.661	1.739	1.814	1.876
5	1.766	.5359	1.581	1.616	1.683	1.763	1.845	1.906	1.961
6	1.844	.1481	1.624	1.659	1.735	1.829	1.932	2.049	2.116
7	1.704	.1352	1.501	1.542	1.608	1.684	1.790	1.887	1.934
8	2.200	.1826	1.888	1.956	2.095	2.209	2.325	2.410	2.446
9	1.941	.2295	1.604	1.663	1.784	1.929	2.089	2.235	2.296
10	1.944	.1482	1.725	1.759	1.834	1.936	2.039	2.136	2.186
avg.	1.871	.1945							
overall avg.	1.883	.1506							

Table A6. Mean Type II employment multiplier value, bootstrap standard deviation, and multiplier values for the millwork sector, for 10 bootstrap trials using the full, three-quarters, and one-half samples.

Boot-strap trial	Mean	Boot-strap s.d.	Percentile 5	10	25	50	75	90	95
full sample									
1	2.045	0.0854	1.903	1.932	1.987	2.051	2.107	2.151	2.173
2	2.522	.3458	1.975	2.081	2.297	2.512	2.765	2.947	3.076
3	2.043	.2745	1.657	1.716	1.827	2.024	2.198	2.424	2.558
4	1.753	.1276	1.588	1.607	1.660	1.731	1.833	1.931	1.993
5	1.881	.1464	1.679	1.713	1.773	1.868	1.986	2.075	2.111
6	1.856	.2075	1.592	1.621	1.680	1.809	1.964	2.116	2.303
7	1.981	.2464	1.654	1.694	1.787	1.929	2.128	2.346	2.452
8	2.279	.2485	1.954	1.985	2.062	2.273	2.441	2.594	2.696
9	2.032	.0820	1.891	1.928	1.979	2.039	2.089	2.127	2.160
10	1.703	.0958	1.578	1.593	1.627	1.684	1.747	1.839	1.902
avg.	2.010	0.1860							
three-quarters sample									
1	1.774	.1311	1.611	1.630	1.680	1.741	1.884	1.937	1.965
2	2.791	.5436	1.781	1.832	2.406	2.833	3.170	3.508	3.724
3	2.269	.3656	1.779	1.829	1.972	2.273	2.474	2.746	2.892
4	2.109	.2732	1.721	1.789	1.916	2.065	2.277	2.473	2.607
5	1.824	.1387	1.651	1.676	1.726	1.800	1.927	2.000	2.028
6	1.898	.1504	1.676	1.708	1.786	1.875	1.999	2.084	2.216
7	2.623	.5757	1.919	1.958	2.103	2.703	3.019	3.268	3.422
8	1.819	.1205	1.632	1.660	1.719	1.792	1.901	2.027	2.098
9	2.045	.1035	1.872	1.911	1.981	2.053	2.123	2.166	2.196
10	2.480	.3037	2.100	2.135	2.217	2.486	2.645	2.926	3.061
avg.	2.164	.2706							
one-half sample									
1	1.917	.2359	1.648	1.679	1.755	1.864	2.129	2.203	2.235
2	2.365	.5138	1.635	1.751	1.960	2.384	2.672	3.003	3.137
3	2.022	.1179	1.835	1.871	1.943	2.020	2.101	2.165	2.199
4	2.457	.5147	1.843	1.881	2.013	2.472	2.792	3.092	3.239
5	2.176	.0971	1.594	1.625	1.749	2.049	2.495	2.977	3.252
6	1.939	.1094	1.747	1.795	1.869	1.947	2.015	2.067	2.088
7	2.823	.6944	1.961	2.032	2.164	2.912	3.287	3.613	3.726
8	2.140	.5424	1.581	1.607	1.752	2.019	2.456	2.919	3.143
9	2.523	.5007	1.862	1.958	2.089	2.543	2.795	3.112	3.258
10	1.940	.1107	1.758	1.792	1.863	1.938	2.018	2.087	2.126
avg.	2.230	.3437							
overall avg.	2.135	.2667							

Table A7. Mean Type II employment multiplier value, bootstrap standard deviation, and multiplier values for the pallet sector, for 10 bootstrap trials using the full, three-quarters, and one-half samples.

Boot-strap trial	Mean	Boot-strap s.d.	Percentile 5	10	25	50	75	90	95
full sample									
1	2.004	0.1016	1.845	1.870	1.935	1.997	2.065	2.138	2.180
2	2.054	.1154	1.881	1.912	1.975	2.046	2.123	2.205	2.254
3	2.045	.0888	1.905	1.934	1.982	2.037	2.099	2.164	2.208
4	1.943	.0621	1.841	1.866	1.901	1.938	1.982	2.026	2.047
5	1.919	.0991	1.769	1.795	1.849	1.912	1.977	2.049	2.086
6	1.909	.0995	1.750	1.784	1.837	1.903	1.972	2.042	2.092
7	2.061	.1575	1.827	1.881	1.949	2.041	2.151	2.274	2.376
8	2.083	.0698	1.963	1.990	2.037	2.079	2.129	2.180	2.209
9	1.995	.1282	1.808	1.839	1.906	1.986	2.074	2.165	2.216
10	2.053	.1218	1.859	1.909	1.964	2.044	2.128	2.219	2.275
avg.	2.007	0.1044							
three-quarters sample									
1	1.928	.1063	1.759	1.799	1.854	1.920	1.996	2.064	2.112
2	2.280	.1511	2.038	2.097	2.169	2.272	2.368	2.479	2.529
3	2.001	.1630	1.766	1.812	1.882	1.985	2.098	2.210	1.918
4	1.781	.0847	1.634	1.670	1.721	1.785	1.837	1.885	2.458
5	2.062	.1816	1.813	1.839	1.912	2.025	2.158	2.347	2.098
6	1.853	.1091	1.707	1.727	1.765	1.827	1.909	2.013	2.334
7	2.062	.1450	1.840	1.877	1.958	2.047	2.154	2.259	2.334
8	2.027	.1558	1.834	1.870	1.946	2.025	2.105	2.188	2.222
9	2.394	.3454	1.918	1.987	2.154	2.374	2.614	2.838	2.957
10	1.740	.1199	1.566	1.596	1.655	1.731	1.804	1.905	1.954
avg.	2.013	.1562							
one-half sample									
1	2.546	.3420	2.078	2.142	2.287	2.480	2.747	3.227	3.179
2	2.112	.3161	1.672	1.727	1.863	2.078	2.323	2.542	2.687
3	2.066	.2027	1.745	1.792	1.916	2.057	2.202	2.330	2.443
4	2.008	.1639	1.777	1.813	1.880	1.977	2.100	2.248	2.333
5	1.846	.1007	1.714	1.730	1.776	1.828	1.901	1.985	2.025
6	2.049	.1304	1.843	1.884	1.952	2.038	2.133	2.229	2.291
7	2.033	.1394	1.830	1.864	1.936	2.009	2.114	2.231	2.301
8	1.776	.0957	1.622	1.656	1.706	1.770	1.832	1.906	1.969
9	1.990	.1463	1.765	1.811	1.886	1.978	2.078	2.179	2.230
10	2.115	.1290	1.900	1.954	2.030	2.108	2.195	2.292	2.340
avg.	2.054	.1766							
overall avg.	2.024	.1457							

Table A8. Mean Type II employment multiplier value, bootstrap standard deviation, and multiplier values for the paperboard sector, for 10 bootstrap trials using the full, three-quarters, and one-half samples.

Bootstrap trial	Mean	Bootstrap s.d.	Percentile						
			5	10	25	50	75	90	95
			full sample						
1	2.257	0.1718	2.017	2.054	2.130	2.228	2.364	2.490	2.587
2	2.340	.1670	2.108	2.150	2.219	2.307	2.436	2.567	2.671
3	2.178	.1439	1.958	1.997	2.076	2.167	2.267	2.372	2.441
4	1.824	.0914	1.684	1.707	1.760	1.820	1.884	1.934	1.974
5	2.223	.1153	2.031	2.072	2.139	2.210	2.295	2.390	2.448
6	2.321	.1283	2.127	2.162	2.219	2.295	2.394	2.511	2.603
7	2.360	.2164	2.051	2.099	2.219	2.349	2.495	2.624	2.709
8	2.291	.1612	2.036	2.091	2.174	2.270	2.389	2.511	2.592
9	2.135	.0793	2.000	2.028	2.075	2.132	2.187	2.245	2.285
10	2.104	.0717	1.985	2.012	2.057	2.104	2.151	2.195	2.219
avg.	2.203	0.1346							
			three-quarters sample						
1	2.022	.1171	1.852	1.882	1.934	2.006	2.093	2.195	2.249
2	2.114	.1337	1.890	1.942	2.011	2.109	2.204	2.309	2.365
3	2.063	.0640	1.942	1.977	2.022	2.062	2.102	2.144	2.172
4	2.280	.1929	1.986	2.032	2.143	2.260	2.381	2.555	2.678
5	2.331	.1576	2.086	2.124	2.200	2.301	2.415	2.539	2.688
6	2.175	.1632	1.913	1.960	2.051	2.165	2.287	2.397	2.466
7	2.217	.1254	2.032	2.061	2.119	2.220	2.284	2.385	2.477
8	2.222	.1299	1.988	2.035	2.105	2.197	2.317	2.424	2.518
9	2.331	.1440	2.134	2.166	2.218	2.299	2.416	2.530	2.634
10	2.484	.1839	2.218	2.265	2.347	2.458	2.607	2.727	2.803
avg.	2.224	.1412							
			one-half sample						
1	2.108	.1480	1.855	1.916	2.005	2.098	2.214	2.307	2.369
2	2.230	.0983	2.095	2.113	2.145	2.201	2.282	2.374	2.470
3	2.205	.0923	2.088	2.101	2.127	2.180	2.247	2.341	2.424
4	2.181	.0932	2.074	2.084	2.110	2.154	2.231	2.320	2.368
5	1.993	.0788	1.819	1.858	1.927	1.991	2.068	2.126	2.166
6	2.021	.1109	2.029	2.060	2.111	2.177	2.268	2.364	2.481
7	2.377	.2626	2.081	2.117	2.178	2.301	2.491	2.793	2.907
8	2.119	.1024	1.947	1.980	2.035	2.101	2.170	2.284	2.404
9	2.194	.0927	2.066	2.078	2.121	2.173	2.245	2.328	2.370
10	2.184	.0649	2.100	2.110	2.130	2.177	2.218	2.263	2.311
avg.	2.161	.1144							
overall avg.	2.196	.1301							

Table A9. Mean Type II income multiplier value, bootstrap standard deviation, and multiplier values for the sawmill sector, for 10 bootstrap trials using the full, three-quarters, and one-half samples.

Bootstrap trial	Mean	Bootstrap s.d.	Percentile 5	10	25	50	75	90	95
			full sample						
1	2.338	0.1615	2.058	2.126	2.228	2.340	2.457	2.546	2.594
2	2.153	.1485	1.946	1.978	2.043	2.142	2.250	2.349	2.407
3	2.214	.1274	2.009	2.048	2.126	2.212	2.306	2.386	2.416
4	2.446	.1666	2.159	2.223	2.331	2.453	2.560	2.664	2.720
5	2.118	.1899	1.835	1.883	1.982	2.098	2.235	2.373	2.460
6	2.192	.1604	1.943	1.988	2.077	2.178	2.295	2.420	2.485
7	2.094	.1429	1.872	1.907	1.993	2.093	2.178	2.275	2.352
8	2.483	.1793	2.161	2.246	2.356	2.469	2.612	2.723	2.786
9	2.154	.1320	1.952	1.987	2.060	2.145	2.237	2.331	2.388
10	1.960	.1056	1.793	1.834	1.883	1.951	2.025	2.095	2.156
avg.	2.215	0.1514							
			three-quarters sample						
1	2.016	.1095	1.841	1.878	1.944	2.006	2.084	2.159	2.191
2	2.125	.1900	1.822	1.887	1.986	2.111	2.245	2.379	2.453
3	2.256	.1997	1.934	2.007	2.132	2.257	2.387	2.491	2.556
4	2.246	.1404	2.039	2.072	2.144	2.226	2.337	2.436	2.517
5	2.245	.1873	1.971	2.009	2.123	2.233	2.360	2.495	2.553
6	2.085	.1295	1.890	1.921	1.990	2.071	2.153	2.271	2.327
7	2.309	.1935	1.994	2.053	2.181	2.318	2.435	2.542	2.598
8	2.442	.1936	2.124	2.191	2.327	2.444	2.574	2.668	2.705
9	2.636	.1936	2.290	2.367	2.519	2.648	2.764	2.881	2.936
10	2.243	.1962	1.933	1.995	2.108	2.239	2.375	2.501	2.561
avg.	2.260	.1726							
			one-half sample						
1	2.151	.1969	1.878	1.914	2.003	2.127	2.264	2.414	2.521
2	2.165	.1865	1.883	1.935	2.035	2.139	2.269	2.435	2.542
3	2.350	.2076	2.000	2.068	2.220	2.359	2.491	2.596	2.667
4	2.059	.1458	1.836	1.878	1.954	2.041	2.149	2.252	2.308
5	2.053	.1140	1.853	1.900	1.975	2.050	2.137	2.199	2.252
6	2.255	.1941	1.967	2.025	2.114	2.237	2.373	2.510	2.591
7	2.082	.1930	1.801	1.848	1.952	2.062	2.198	2.334	2.429
8	2.577	.1980	2.233	2.318	2.447	2.591	2.715	2.799	2.863
9	2.486	.2351	2.076	2.165	2.332	2.493	2.663	2.784	2.841
10	2.378	.2105	2.063	2.125	2.223	2.368	2.526	2.642	2.717
avg.	2.256	.1874							
overall avg.	2.244	.1705							

Table A10. Mean Type II income multiplier value, bootstrap standard deviation, and multiplier values for the millwork sector, for 10 bootstrap trials using the full, three-quarters, and one-half samples.

Boot-strap trial	Mean	Boot-strap s.d.	Percentile 5	10	25	50	75	90	95
			full sample						
1	2.030	0.0986	1.884	1.911	1.963	2.023	2.092	2.164	2.195
2	2.419	.1951	2.078	2.158	2.296	2.425	2.556	2.660	2.710
3	2.205	.1981	1.928	1.974	2.055	2.211	2.316	2.442	2.525
4	1.805	.2692	1.510	1.550	1.615	1.721	2.049	2.150	2.201
5	1.880	.2549	1.574	1.609	1.694	1.821	2.066	2.216	2.290
6	1.734	.1970	1.477	1.507	1.574	1.681	1.845	2.016	2.143
7	1.835	.2295	1.540	1.576	1.659	1.778	1.956	2.162	2.346
8	2.284	.1805	1.997	2.064	2.148	2.281	2.402	2.515	2.579
9	2.208	.1425	1.981	2.024	2.112	2.204	2.302	2.382	2.448
10	1.678	.1376	1.471	1.499	1.548	1.626	1.735	1.969	2.171
avg.	2.008	0.1903							
			three-quarters sample						
1	1.836	.3041	1.509	1.531	1.606	1.731	2.048	2.260	2.366
2	2.510	.3094	1.842	1.888	2.320	2.542	2.741	2.909	3.016
3	2.148	.2450	1.798	1.837	1.974	2.156	2.294	2.447	2.519
4	2.028	.2901	1.622	1.706	1.808	1.992	2.220	2.408	2.519
5	1.872	.2839	1.554	1.585	1.667	1.785	2.085	2.254	2.347
6	1.827	.1899	1.554	1.581	1.672	1.771	1.939	2.172	2.343
7	2.286	.3485	1.857	1.878	1.965	2.331	2.527	2.680	2.747
8	1.820	.1972	1.537	1.574	1.653	1.749	1.927	2.248	2.396
9	2.217	.1425	1.994	2.041	2.121	2.211	2.309	2.393	2.455
10	2.304	.1940	2.021	2.064	2.162	2.303	2.426	2.551	2.622
avg.	2.085	.2505							
			one-half sample						
1	1.973	.4410	1.523	1.557	1.671	1.830	2.303	2.586	2.674
2	2.333	.3076	1.884	1.959	2.069	2.387	2.531	2.681	2.758
3	2.161	.1310	1.961	1.989	2.057	2.148	2.238	2.333	2.420
4	2.453	.3032	2.001	2.064	2.230	2.479	2.665	2.809	2.884
5	2.044	.5463	1.490	1.538	1.639	1.896	2.408	2.784	2.878
6	2.534	.2565	2.090	2.211	2.361	2.534	2.715	2.855	2.938
7	2.455	.3920	1.934	1.990	2.084	2.520	2.731	2.878	2.959
8	2.170	.6499	1.491	1.543	1.681	1.942	2.790	2.951	3.013
9	2.553	.2587	2.087	2.189	2.404	2.596	2.738	2.842	2.905
10	2.349	.1713	2.064	2.110	2.190	2.301	2.429	2.541	2.627
avg.	2.300	.3457							
overall avg.	2.131	.2622							

Table A11. Mean Type II income multiplier value, bootstrap standard deviation, and multiplier values for the pallet sector, for 10 bootstrap trials using the full, three-quarters, and one-half samples.

Bootstrap trial	Mean	Bootstrap s.d.	Percentile						
			5	10	25	50	75	90	95
full sample									
1	2.447	0.1684	2.210	2.246	2.326	2.435	2.539	2.674	2.756
2	2.419	.1429	2.212	2.251	2.317	2.407	2.505	2.613	2.678
3	2.490	.1305	2.290	2.334	2.400	2.482	2.578	2.657	2.709
4	2.393	.1055	2.227	2.263	2.317	2.388	2.465	2.526	2.563
5	2.260	.1078	2.093	2.124	2.179	2.254	2.326	2.404	2.450
6	2.428	.1350	2.213	2.252	2.339	2.421	2.514	2.602	2.674
7	2.744	.1272	2.541	2.579	2.652	2.744	2.821	2.911	2.958
8	2.144	.0878	2.022	2.034	2.075	2.131	2.200	2.259	2.319
9	2.508	.1938	2.236	2.280	2.363	2.495	2.625	2.749	2.852
10	2.632	.1230	2.440	2.483	2.542	2.630	2.708	2.789	2.843
avg.	2.446	0.1322							
three-quarters sample									
1	2.519	.1839	2.255	2.299	2.389	2.502	2.616	2.777	2.866
2	2.913	.2388	2.558	2.626	2.744	2.892	3.052	3.234	3.332
3	2.798	.2193	2.480	2.590	2.636	2.762	2.940	3.085	3.182
4	2.391	.1238	2.194	2.225	2.309	2.388	2.470	2.546	2.595
5	3.024	.1666	2.745	2.815	2.913	3.012	3.145	3.245	3.318
6	2.683	.1841	2.367	2.437	2.543	2.681	2.808	2.911	2.988
7	2.713	.2048	2.379	2.457	2.568	2.698	2.832	2.978	3.082
8	2.579	.1550	2.329	2.389	2.469	2.570	2.680	2.783	2.843
9	2.921	.3011	2.457	2.544	2.718	2.917	3.116	3.311	3.417
10	2.149	.1258	1.962	1.997	2.053	2.134	2.231	3.313	2.385
avg.	2.669	.1903							
one-half sample									
1	2.784	.4209	2.268	2.353	2.484	2.691	3.083	3.337	3.463
2	2.556	.2916	2.123	2.197	2.354	2.535	2.748	2.945	3.046
3	2.419	.2474	2.060	2.109	2.247	2.404	2.580	2.718	2.867
4	2.548	.2532	2.184	2.247	2.357	2.507	2.685	2.899	3.059
5	2.722	.1842	2.430	2.489	2.595	2.714	2.846	2.952	3.016
6	2.597	.1282	2.372	2.438	2.513	2.594	2.684	2.756	2.813
7	2.541	.2270	2.218	2.282	2.387	2.522	2.685	2.831	2.909
8	2.583	.2003	2.261	2.322	2.446	2.571	2.709	2.851	2.939
9	2.349	.1722	2.113	2.146	2.225	2.329	2.459	2.574	2.651
10	2.533	.1367	2.315	2.361	2.441	2.519	2.621	2.709	2.771
avg.	2.563	.2262							
overall avg.	2.559	.1829							

Table A12. Mean Type II income multiplier value, bootstrap standard deviation, and multiplier values for the paperboard sector, for 10 bootstrap trials using the full, three-quarters, and one-half samples.

Boot-strap trial	Mean	Boot-strap s.d.	Percentile 5	10	25	50	75	90	95
full sample									
1	2.066	0.1886	1.805	1.856	1.916	2.026	2.183	2.322	2.481
2	2.041	.1708	1.807	1.842	1.905	2.006	2.138	2.296	2.403
3	1.850	.1051	1.699	1.726	1.775	1.845	1.912	1.993	2.031
4	1.589	.0591	1.497	1.514	1.547	1.586	1.629	1.658	1.692
5	1.898	.1036	1.735	1.766	1.816	1.877	1.957	2.047	2.126
6	2.000	.1469	1.807	1.838	1.889	1.966	2.075	2.214	2.669
7	2.008	.1599	1.770	1.814	1.881	1.996	2.120	2.220	2.286
8	1.989	.1443	1.776	1.809	1.882	1.964	2.080	2.180	2.265
9	1.807	.0630	1.709	1.735	1.761	1.799	1.845	1.892	1.922
10	1.758	.0467	1.679	1.694	1.727	1.753	1.786	1.821	1.848
avg.	1.901	0.1188							
three-quarters sample									
1	1.760	.0947	1.622	1.649	1.687	1.740	1.819	1.893	1.946
2	1.829	.1127	1.665	1.690	1.749	1.814	1.889	1.995	2.064
3	1.858	.1040	1.718	1.740	1.779	1.832	1.906	2.016	2.078
4	2.202	.2729	1.826	1.882	1.981	2.142	2.337	2.618	2.819
5	1.980	.1429	1.772	1.804	1.859	1.944	2.051	2.186	2.330
6	1.962	.1682	1.733	1.768	1.829	1.932	2.067	2.187	2.279
7	1.856	.1081	1.703	1.725	1.767	1.827	1.910	2.011	2.098
8	1.847	.1217	1.678	1.707	1.754	1.822	1.916	2.015	2.080
9	1.918	.1271	1.751	1.770	1.811	1.886	1.972	2.105	2.221
10	2.257	.2320	1.940	1.996	2.083	2.215	2.402	2.600	2.688
avg.	1.947	.1484							
one-half sample									
1	1.753	.0844	1.611	1.638	1.689	1.744	1.803	1.877	1.924
2	1.845	.0964	1.709	1.724	1.757	1.814	1.890	1.981	2.100
3	1.898	.0903	1.771	1.787	1.819	1.862	1.938	2.043	2.119
4	1.795	.0796	1.698	1.705	1.728	1.770	1.837	1.921	1.990
5	1.792	.0996	1.642	1.669	1.708	1.775	1.835	1.917	1.994
6	2.014	.1885	1.763	1.794	1.842	1.938	2.097	2.306	2.598
7	2.238	.3942	1.830	1.872	1.952	2.097	2.455	2.825	3.046
8	1.982	.2177	1.750	1.774	1.821	1.901	2.068	2.352	2.455
9	1.819	.0590	1.735	1.747	1.770	1.803	1.850	1.907	1.944
10	1.800	.0717	1.698	1.712	1.741	1.782	1.842	1.909	1.964
avg.	1.894	.1381							
overall avg.	1.914	.1351							

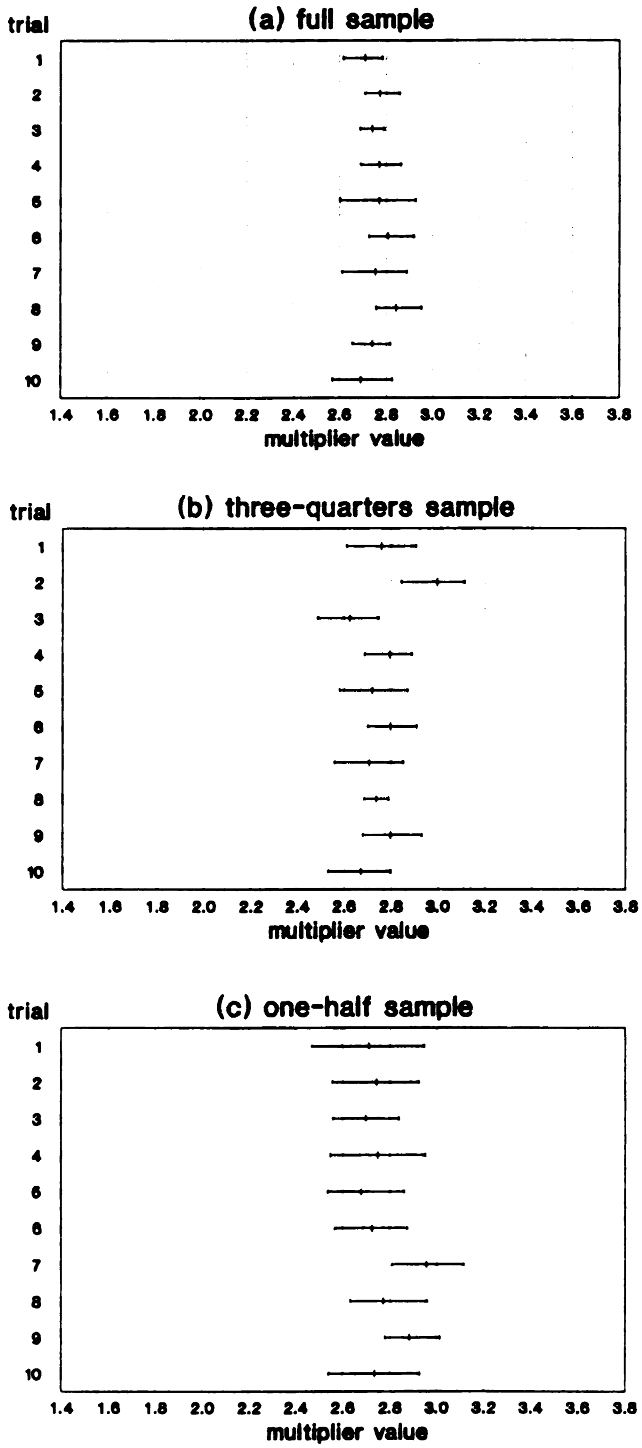

Figure A1. Type II output multiplier 90 percent confidence intervals for the sawmill sector generated using 600 bootstrap replications, and 10 trials.

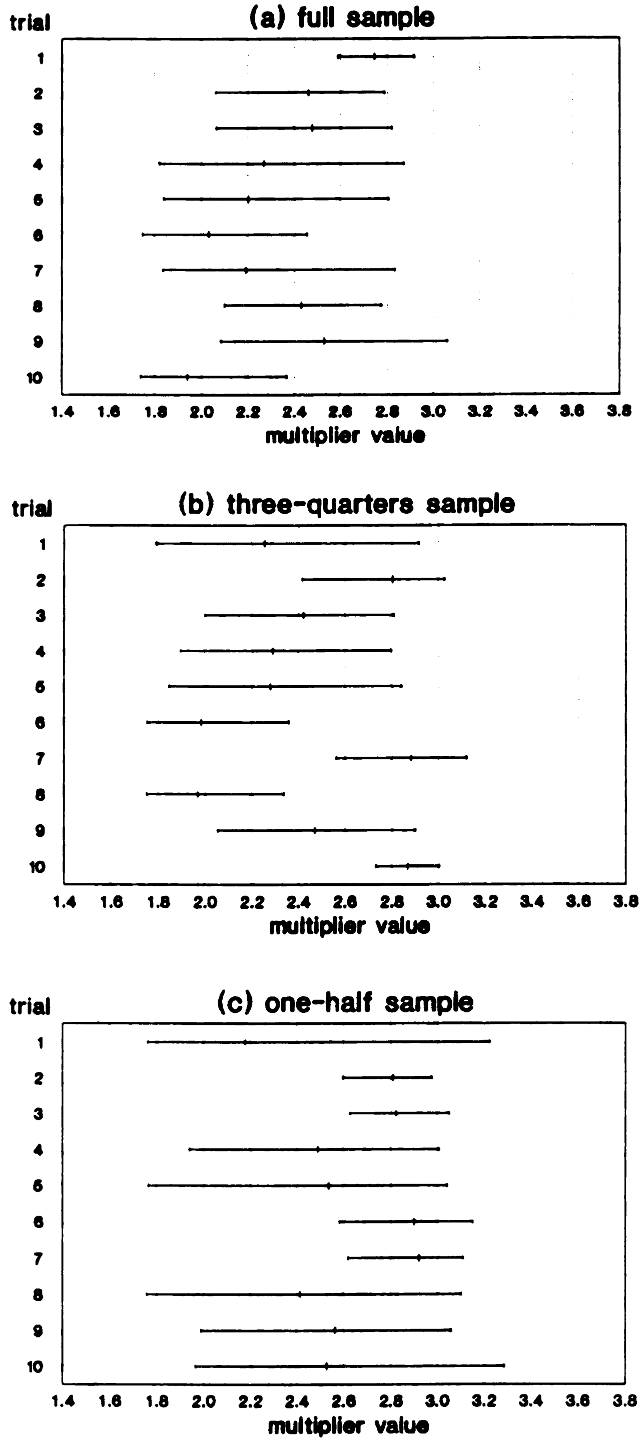

Figure A2. Type II output multiplier 90 percent confidence intervals for the millwork sector generated using 600 bootstrap replications, and 10 trials.

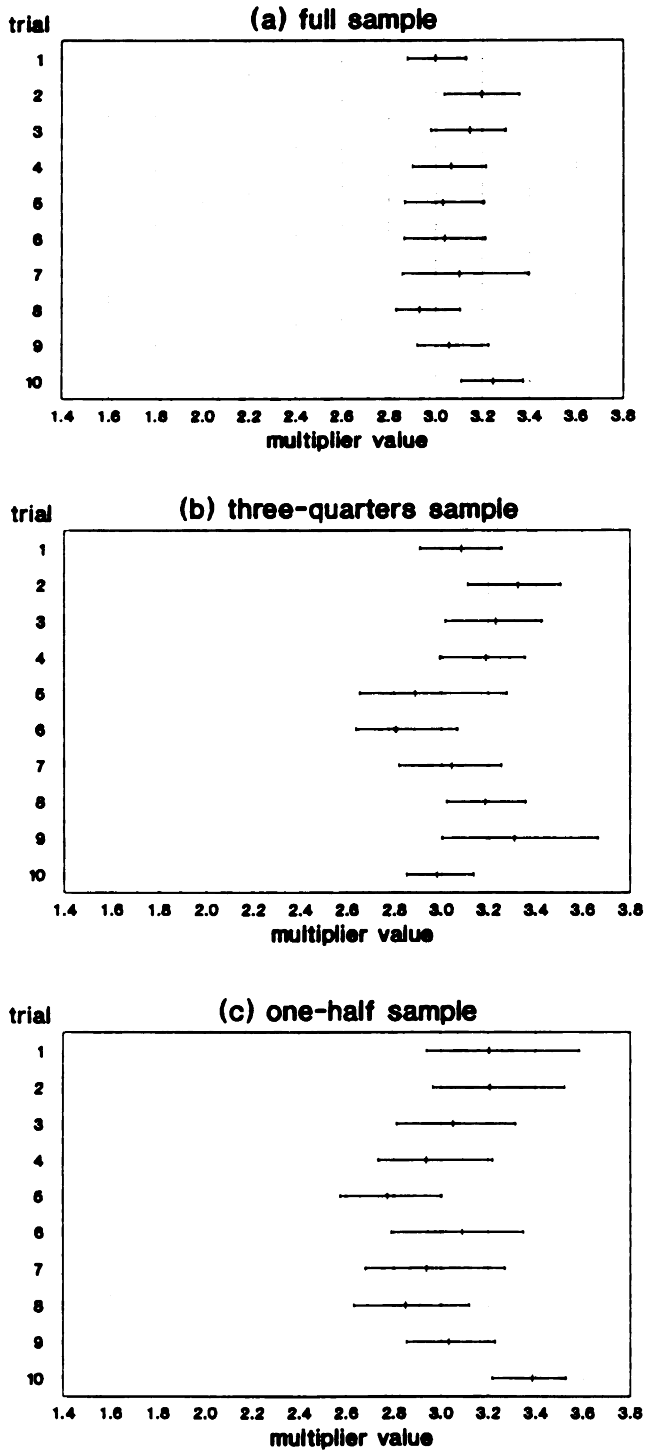

Figure A3. Type II output multiplier 90 percent confidence intervals for the wood pallet sector generated using 600 bootstrap replications, and 10 trials.

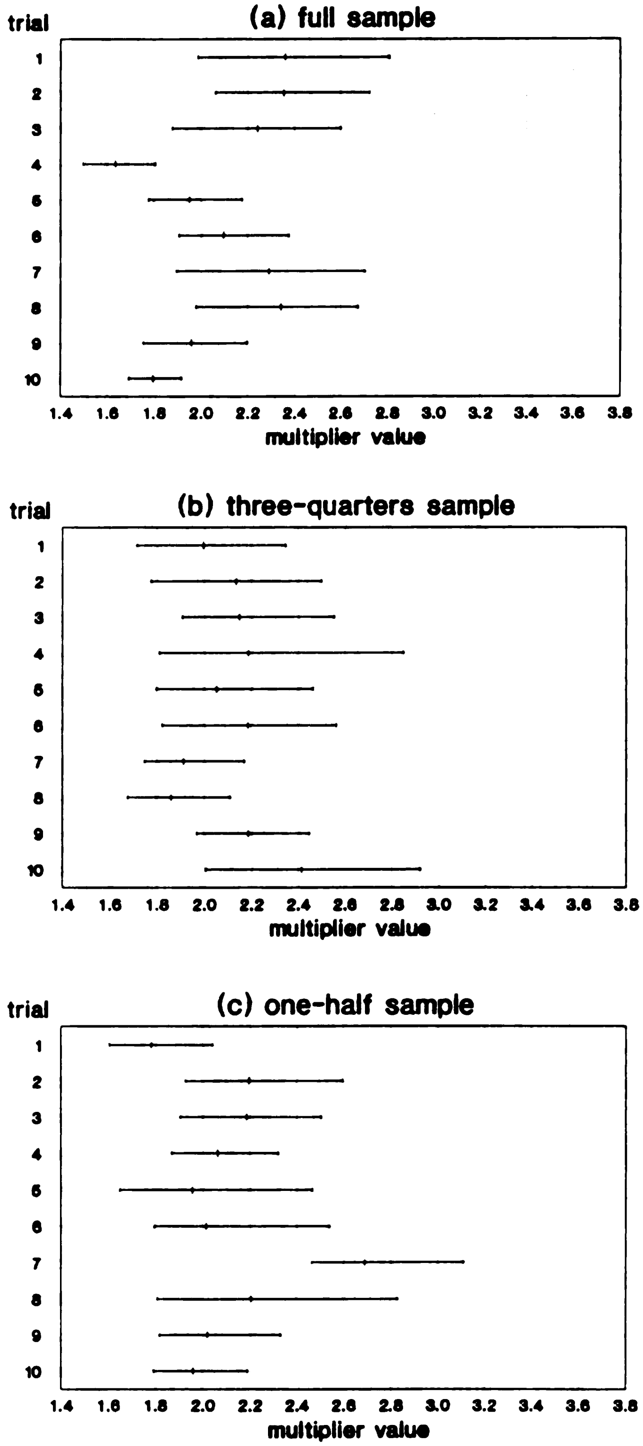

Figure A4. Type II output multiplier 90 percent confidence intervals for the paperboard sector generated using 600 bootstrap replications, and 10 trials.

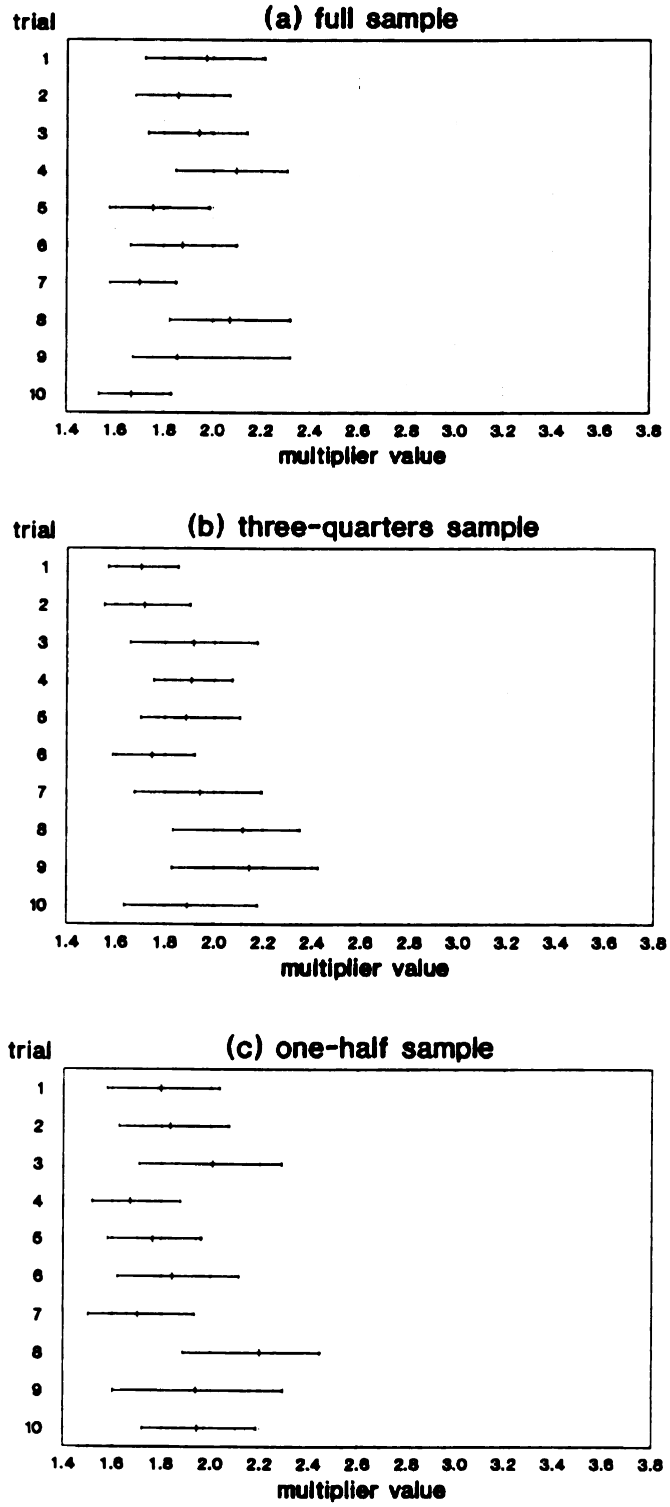

Figure A5. Type II employment multiplier 90 percent confidence intervals for the sawmill sector generated using 600 bootstrap replications, and 10 trials.

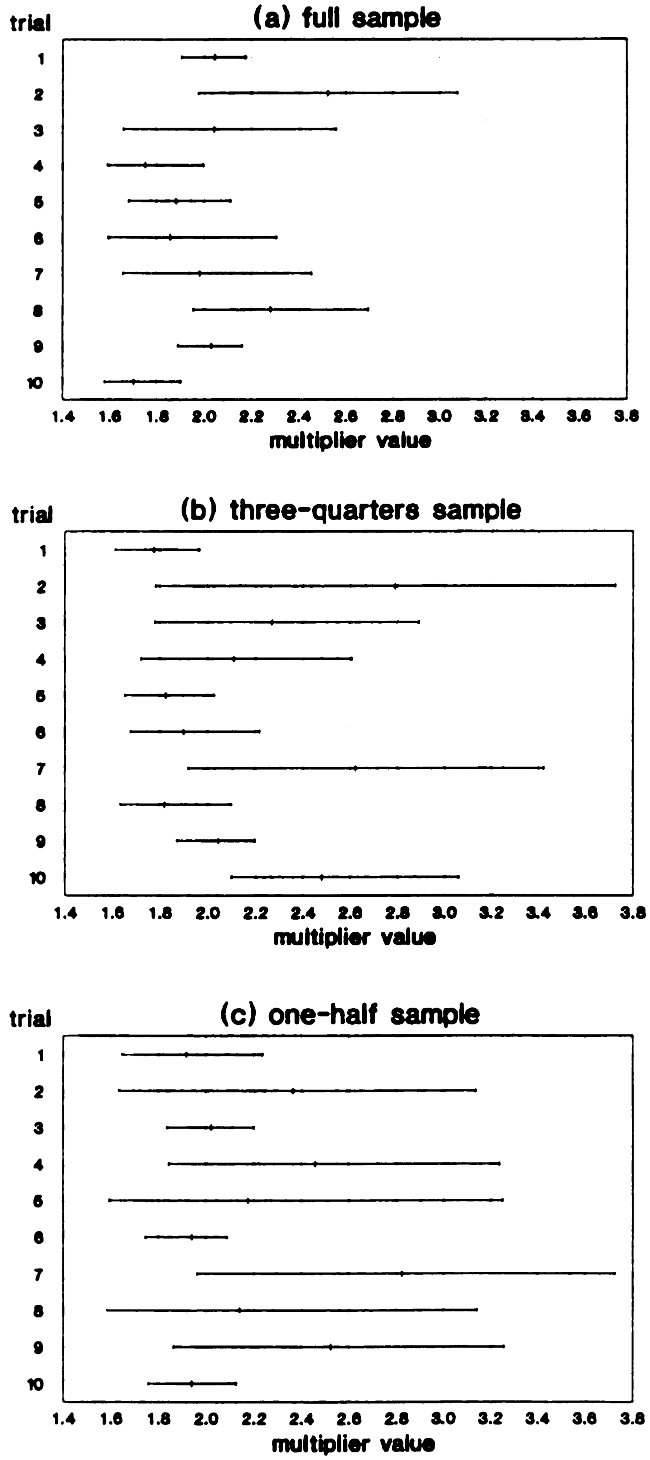

Figure A6. Type II employment multiplier 90 percent confidence intervals for the millwork sector generated using 600 bootstrap replications, and 10 trials.

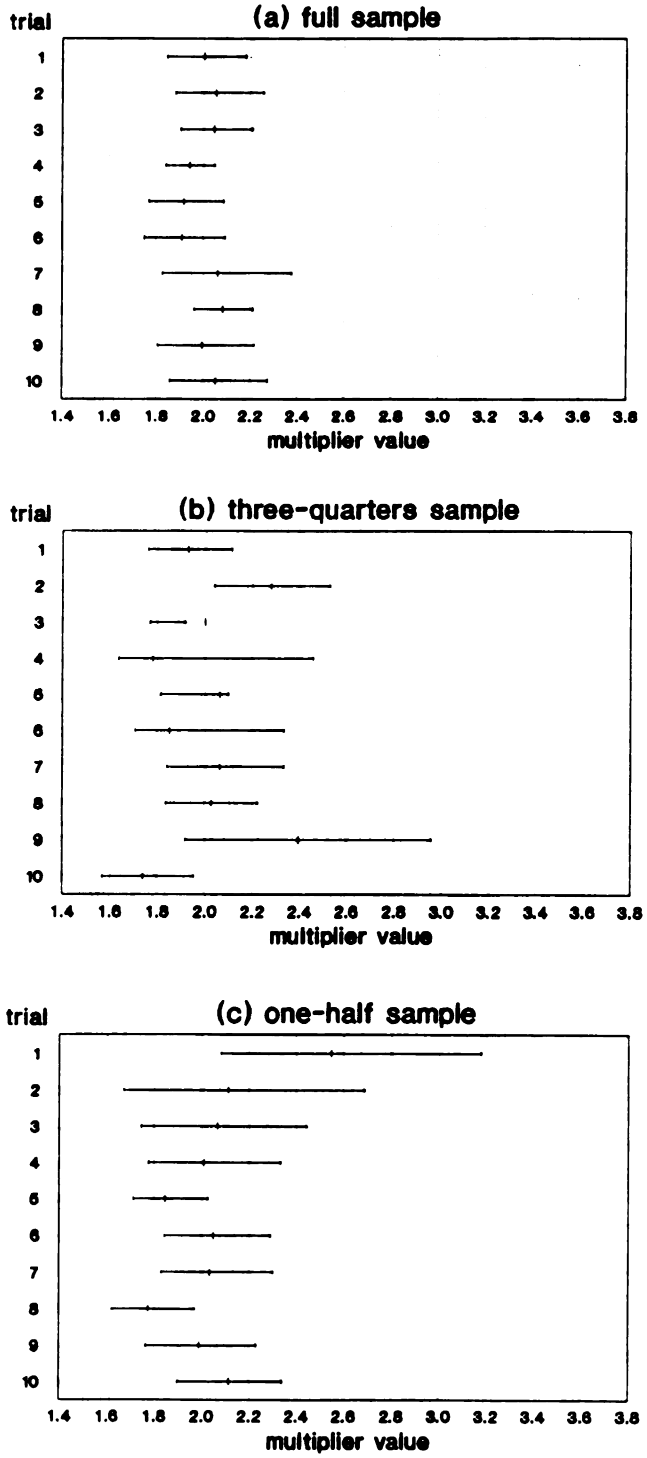

Figure A7. Type II employment multiplier 90 percent confidence intervals for the wood pallet sector generated using 600 bootstrap replications, and 10 trials.

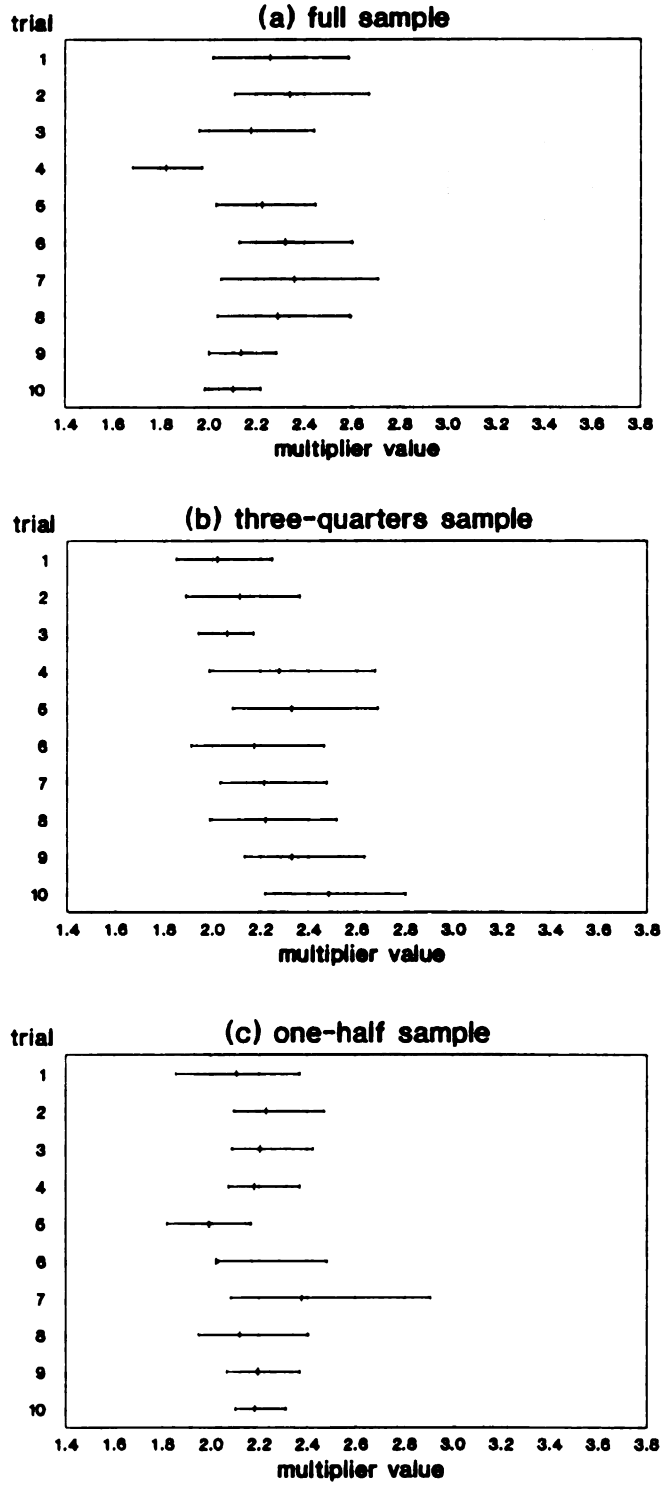

Figure A8. Type II employment multiplier 90 percent confidence intervals for the paperboard sector generated using 600 bootstrap replications, and 10 trials.

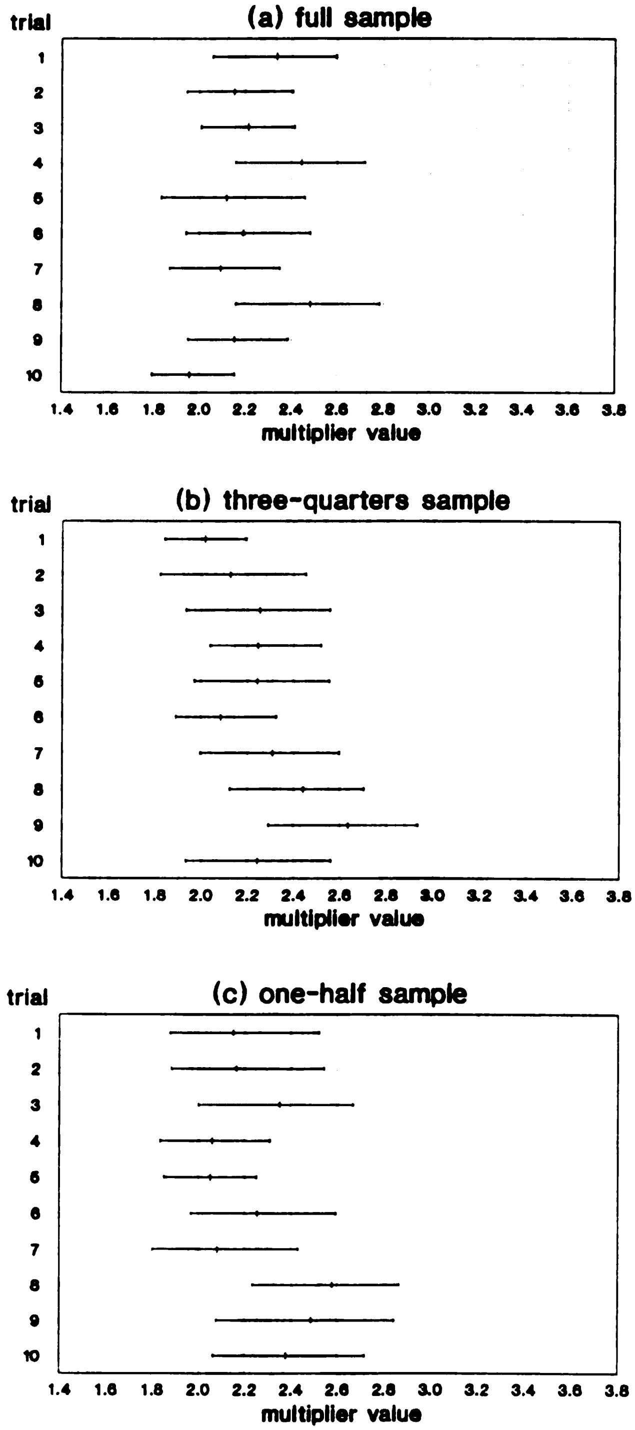

Figure A9. Type II income multiplier 90 percent confidence intervals for the sawmill sector generated using 600 bootstrap replications, and 10 trials.

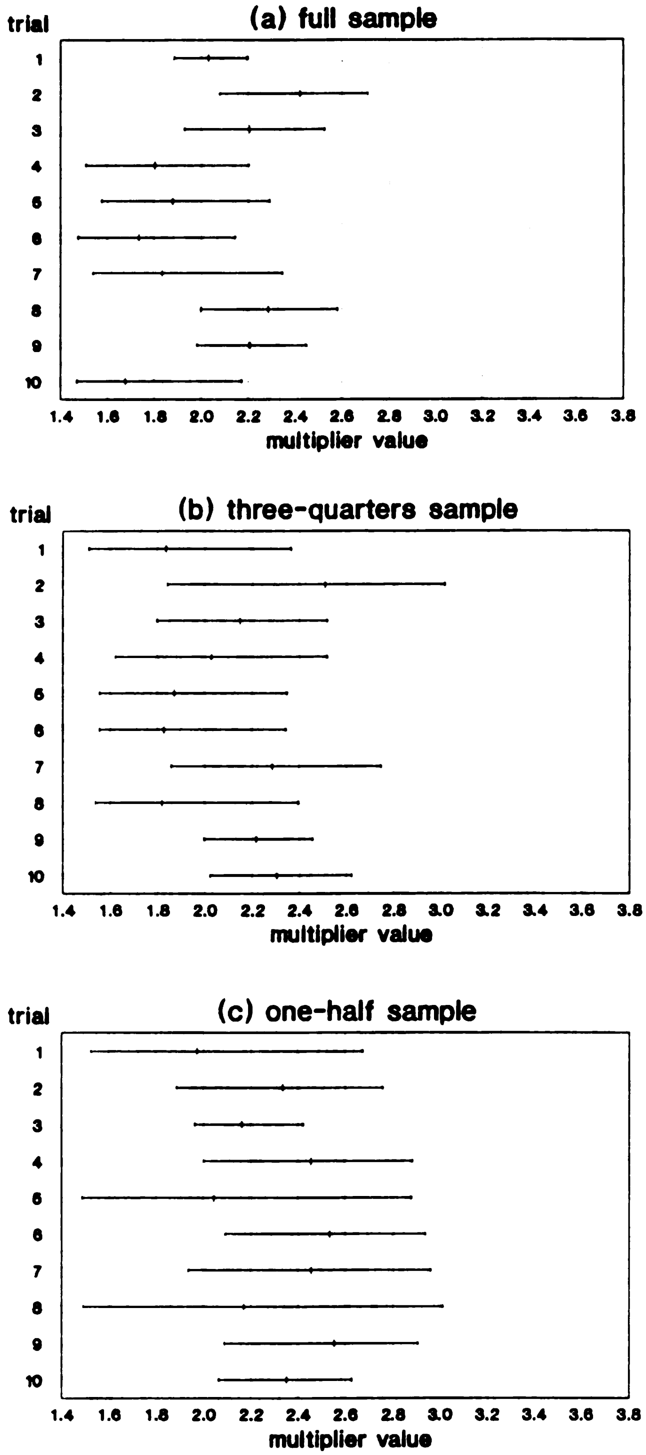

Figure A10. Type II income multiplier 90 percent confidence intervals for the millwork sector generated using 500 bootstrap replications, and 10 trials.

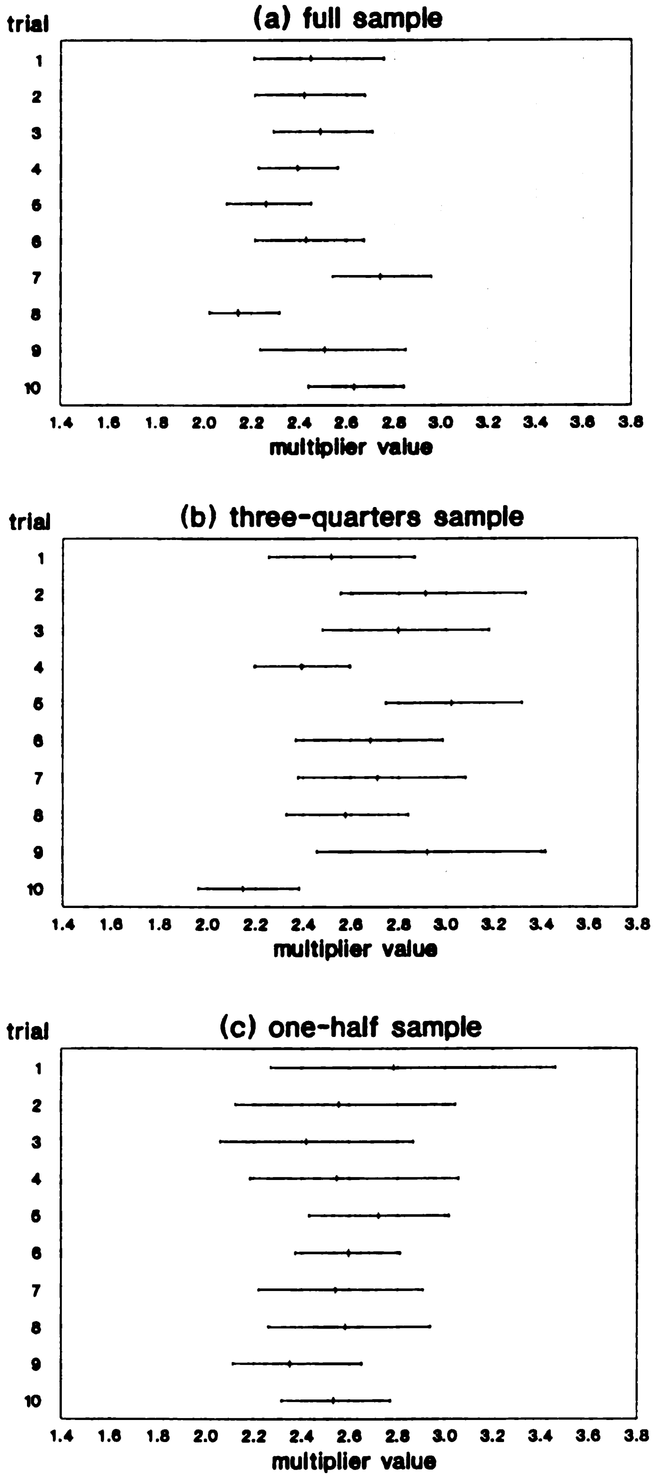

Figure A11. Type II income multiplier 90 percent confidence intervals for the wood pallet sector generated using 600 bootstrap replications, and 10 trials.

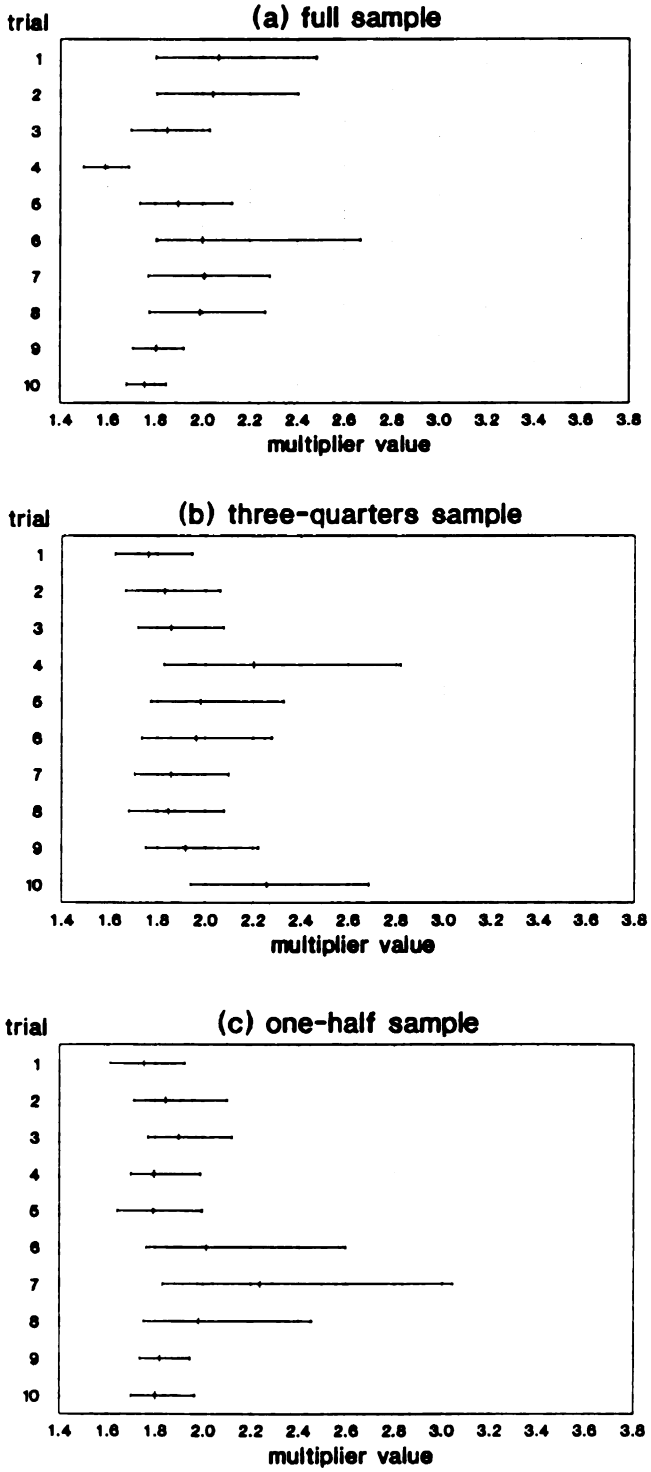

Figure A12. Type II income multiplier 90 percent confidence intervals for the paperboard sector generated using 600 bootstrap replications, and 10 trials.

APPENDIX B

FOREST INDUSTRIES STUDY

MICHIGAN STATE UNIVERSITY

DEPARTMENT OF FORESTRY

NATURAL RESOURCES BLDG.

EAST LANSING, MI 48824

Forest Industries Questionnaire

Your answers to the first five questions will permit us to classify your establishment by size and industry, and to identify your net inventory changes. By "establishment" we mean a place of business or specific location of an economic activity which includes the physical structure, equipment and employees (e.g., a manufacturing plant).

Q-1 What were your total sales in 1980 from your Michigan operations of all products?

TOTAL SALES $ ___________

Q-2 Please list your major products or services and what percentage each is of total sales (dollars).

	PRODUCTS	PERCENTAGE
1.	__________	__________
2.	__________	__________
3.	__________	__________
4.	__________	__________

Q-3 What are your current capacities (please list products in same order as in Q-2 and include units of output per day or month)?

Capacity is the maximum amount that the establishment can produce with the given quantity and quality of inputs and the given technology. Note that if an additional shift is possible with the current equipment but that additional production would not be possible without additional labor, that this would constitute an expansion of capacity and should be included in Q-7.

	PRODUCT	CAPACITY
1.	__________	__________
2.	__________	__________
3.	__________	__________
4.	__________	__________

Q-4 What was your total average monthly employment and payroll during 1980?

Please estimate employment in terms of "full time equivalents." This means that all part time employees should be converted to the number employed if they had been working full time.

1. Number of Employees ______________

2. Payroll $______________

Q-5 What was the value of your closing inventory of finished goods in 1979 and 1980?

1979 Value of Inventory $______________

1980 Value of Inventory $______________

The next three questions will help us forecast planned changes in Michigan's industrial capacity.

Q-6 Do you plan to expand (or reduce) your firms's capacity in Michigan during the next 2 years? (Circle appropriate response.)

1.	Expanding	If no change
2.	Remain about the same	is planned,
3.	Reducing	skip from
		here to Q-9.

Q-7 Please describe the type of expansion (reduction) and estimated cost planned by your firm in Michigan (types of capital equipment, improvements, shutdown, etc.) in the next 2 years.

Q-8 What will be the capacity of your firm (in units of output per day or month) after the planned expansion (or reduction) is completed?

	PRODUCT	CAPACITY
1.	___________	___________
2.	___________	___________
3.	___________	___________
4.	___________	___________

The next two questions organize your annual sales and purchases by industrial group. When weighted by the size of your establishment and adjusted for net inventory change, your answers along with others will show the structure of sales and purchases for all industries in Michigan. There is an example with each question to help make it clear.

Q-9 What were your sales in 1980 to the user groups shown below?

Please write your answers as percentages of total sales (revenues) from your Michigan operations and exclude sales of surplus equipment resulting from capacity reductions. The word user is emphasized because we must know the groups that actually use your products for additional production or final consumption. If most of your sales are to wholesalers or retailers who pass your product on to others, please write the percentage of sales by the user group of final destination (either final users such as state and local government, federal government or households or intermediate groups 01 to 39). Also, please put and "X" next to sales that "passed through" a wholesaler or retailer. If you do not know the final user destinations, please place the information in the wholesalers or retailers rows as appropriate.

Example: Sales Group	Percentage of total sales	Percent to Michigan users
01 Agricultural products and services	6%	95%
.. ....		
16 Integrated pulp and paper or paperboard mills	5%	5%
.. ....		
19 Converted paper and paperboard products	11%	%
.. ....		
36 Wholesale (except forest products) and retail trade	8%	25%
37 Finance, insurance and real estate	1%	100%
.. ....		
42 Personal consumption	12%	60%
43 Capital formation	6%	100%
.. ....		
45 Exports	7%	%
Total	100%	

The hypothetical establishment sold 6% of its total sales to agricultural products and services. Of this, 95 percent went to Michigan establishments. They also sold 11 percent of their total sales to converted paper and paperboard products, all of which were outside of the state. Twelve percent of their total sales went to personal consumption (or households), 60% of which were located in Michigan. The rest of this example can be explained in a similar fashion. The figures in the "percentage of Total Sales" column should sum to 100 percent.

	Group	Percentage of Total Sales	Percent to Michigan Users
01	Agricultural products and services	_____%	_____%
02	Mining (including metals, minerals, crude petroleum, natural gas)	_____%	_____%
03	Construction	_____%	_____%
04	Food and kindred products	_____%	_____%
05	Textiles and apparel	_____%	_____%
06	National forests (stumpage sellers)	_____%	_____%
07	State forests (stumpage sellers)	_____%	_____%
08	Other stumpage sellers	_____%	_____%
09	Logging contractors	_____%	_____%
10	Sawmills and planing mills	_____%	_____%
11	Millwork, flooring, structural members	_____%	_____%
12	Wood furniture and fixtures	_____%	_____%
13	Veneer and plywood	_____%	_____%

14	Wood pallets, boxes and skids	_____%	_____%
15	Other lumber and wood products	_____%	_____%
16	Integrated pulp and paper or	_____%	_____%
17	Paper mills, (not integrated with pulp manufacture) except building paper mills)	_____%	_____%
18	Paperboard containers and boxes	_____%	_____%
19	Converted paper and paperboard	_____%	_____%
20	Building paper and building board mills	_____%	_____%
21	Other paper products	_____%	_____%
22	Wholesale trade, forest products only	_____%	_____%
23	Printing, publishing and allied industries	_____%	_____%
24	Chemicals and allied products (includes plastics and synthetic materials, drugs, industrial organic chemicals, agricultural chemicals)	_____%	_____%
25	Petroleum refining	_____%	_____%
26	Rubber and leather products	_____%	_____%
27	Stone, clay, glass & concrete products	_____%	_____%
28	Primary metal industries	_____%	_____%
29	Fabricated metal products, except machinery and transportation equipment	_____%	_____%
30	Machinery	_____%	_____%
31	Transportation equipment	_____%	_____%
32	Misc. manufacturing	_____%	_____%

33	Transportation and communication	____%	____%
34	Electrical and gas utilities	____%	____%
35	Water and sanitary service	____%	____%
36	Wholesale (except forest products) & retail trade	____%	____%
37	Finance, insurance and real estate	____%	____%
38	Other services (including major group 08)	____%	____%

Final Users

39	Federal government (except stumpage sellers)	____%	____%
40	State government (except stumpage sellers)	____%	____%
41	Local government (except stumpage sellers)	____%	____%
42	Personal consumption	____%	____%
43	Capital formation	____%	____%
44	Change in inventory	____%	____%
45	Exports	____%	____%

Q-10 What were your purchases in 1980 from the industry groups shown below? Please write your answers as percentage of total sales from your Michigan operations (your answer to Q-1). Total sales is used as the base of calculating percentages in this question because it is generally easier to estimate in a consistent fashion among business firms. In your answer, please exclude purchases of capital equipment (these were requested earlier).

If most of your purchases and expenditures are from wholesalers or retailers who bought the products from others, please write the percentages of total sales under the industry group that actually made the product. Please put an "X" next to purchases that "passed through" a wholesaler or retailer. Purchases from a wholesaler or retailer which cannot be traced to an industry of origin should be placed under Group 36, wholesale (except forest products) and retail trade.

Taxes are to be recorded as purchased from appropriate governments. Wages and salaries are to be recorded as purchases from households. As before, it is important to identify the portion of your purchases from industry groups in Michigan. If you do not provide a specific estimate, we will assume all your purchases from that group are imported into Michigan. Because capital expenditures are excluded, the percentages need not add to 100. For further description of industry groups please refer to the enclosed appendix. The example below may help clarify our directions.

Example: PURCHASES	Purchases as a percentage of total sales	Percent from Michigan industries
.. ....		
10 Sawmills and planing mills	4%	98%
.. ....		
16 Integrated pulp and paper or paperboard mills	X24%	75%
.. ....		
24 Chemicals and allied products (includes plastics and synthetic materials, drugs, industrial organic chemicals, agricultural chemicals)	5%	40%
33 Transportation and communication	2%	85%
.. ....		
37 Finance, insurance and real estate	1%	%
Payment Sector		
.. ....		
40 State government (except stumpage sellers)	1%	100%
41 Local government (except stumpage sellers)	1%	100%
42 Households (labor)	17%	85%
.. ....		
44 Other payments (e.g. rent)	21%	50%
TOTAL	100%	

In this example, the establishment spends four percent of its total purchases on products from the sawmills and planing mills industry, 98% of which is from Michigan producers. This establishment spent 24 percent of its payments on products from integrated pulp and paper or paperboard mills and 75 percent of these purchases were from establishments in Michigan. The "X" indicates these products are mainly bought from a wholesaler. One percent of the establishment's purchases were from the finance, insurance and real estate sector and all of these purchases were from establishments outside of Michigan. The figures for the other processing sectors can be similarly interpreted.

Among the payments sectors, one percent of the establishment's purchases are paid to state government (except stumpage sellers) as income taxes, workmen's compensation insurance, unemployment insurance, etc., of which all are to Michigan. Seventeen percent of the purchases is paid to Households as wages, salaries and dividends, 85% of which went to Michigan households.

	Purchases as a percentage of total sales	Percent from Michigan industries
01 Agricultural products and services	_____%	_____%
02 Mining (including metals, minerals, crude petroleum, nat. gas)	_____%	_____%
03 Construction	_____%	_____%
04 Food and kindred products	_____%	_____%
05 Textiles and apparel	_____%	_____%
06 National forests (stumpage sellers)	_____%	_____%
07 State forests (stumpage sellers)	_____%	_____%
08 Other stumpage sellers	_____%	_____%
09 Logging contractors	_____%	_____%
10 Sawmills and planing mills	_____%	_____%
11 Millwork, flooring, structural members	_____%	_____%
12 Wood furniture and fixtures	_____%	_____%
13 Veneer and plywood	_____%	_____%
14 Wood pallets and skids	_____%	_____%
15 Other lumber and wood products	_____%	_____%
16 Integrated pulp and paper or paperboard mills	_____%	_____%
17 Paper mills (not integrated with pulp manufacture), except building paper mills	_____%	_____%
18 Paperboard containers and boxes	_____%	_____%

19 Converted paper and paperboard products	_____%	_____%
20 Building paper and building board mills	_____%	_____%
21 Other paper products	_____%	_____%
22 Wholesale trade, forest products only	_____%	_____%
23 Printing, publishing and allied industries	_____%	_____%
24 Chemicals and allied products (includes plastics and synthetic materials, drugs, industrial organic chemicals, agricultural chemicals)	_____%	_____%
25 Petroleum refining	_____%	_____%
26 Rubber and leather products	_____%	_____%
27 Stone, clay, glass & concrete products	_____%	_____%
28 Primary metal industries	_____%	_____%
29 Fabricated metal products, except machinery & transportation equipment	_____%	_____%
30 Machinery	_____%	_____%
31 Transportation equipment	_____%	_____%
32 Misc. manufacturing	_____%	_____%
33 Transportation and communication	_____%	_____%
34 Electrical and gas utilities	_____%	_____%
35 Water and sanitary service	_____%	_____%
36 Wholesale (except forest products) & retail trade	_____%	_____%
37 Finance, insurance and real estate	_____%	_____%

38 Other services (including major group 08)	_____%	_____%

<u>Payments Sector</u>

39 Federal government, including taxes (except stumpage sellers)	_____%	_____%
40 State government, including taxes (except stumpage sellers)	_____%	_____%
41 Local government, including taxes (except stumpage sellers)	_____%	_____%
42 Households (labor costs, including fringe benefits)	_____%	_____%
43 Other payments (e.g. rent and profit)	_____%	_____%

TOTAL _____%

Q-11 If roundwood or wholetree chips are one of your inputs, please indicate below the percentages of your wood raw material inputs by geographic source:

Upper Peninsula _____%

Northern Lower Peninsula _____%

Southern Lower Peninsula _____%

What percentage of your total wood inputs came from your own lands _____%

Total volume of roundwood or wholetree chips used as input in 1980 was _____ (Circle one of following measurement scales: cords, tons, MBd.Ft. or cunnits.)

Total volume of residual chips used as an input in 1980 was _____ tons.

Q-12 Main transportation mode used to move raw materials is_________.
For final products the main transportation mode is_____________.
(Examples of transportation media are trucks, railroad, and barge).

Please print the name and address of your establishment.

____

____

____

Please print the name and telephone number of the person completing this questionnaire.

Name: ____

Phone #: ____

Do you wish to receive a copy of the results (Please circle your answer.)

1. Yes

2. No

Is there anything else you would like to tell us about your establishment, firm or industry? If so, please use the following space.