Symposium on Auditing Research

By an Audit Group at the University of Illinois
at Urbana-Champaign

Department of Accountancy
University of Illinois at Urbana-Champaign
<table>
<thead>
<tr>
<th>Name</th>
<th>Affiliation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Robert K. Ashton</td>
<td>The University of Texas at Austin</td>
</tr>
<tr>
<td>Russell M. Baetfield</td>
<td>The University of Arizona</td>
</tr>
<tr>
<td>Michael J. Barrett</td>
<td>University of Missouri—Kansas City</td>
</tr>
<tr>
<td>William C. Boynton</td>
<td>University of Virginia</td>
</tr>
<tr>
<td>Douglas R. Carmichael</td>
<td>American Institute of Certified Public Accountants</td>
</tr>
<tr>
<td>Neil C. Churchill</td>
<td>Harvard University</td>
</tr>
<tr>
<td>Jesse V. Dillard</td>
<td>The Ohio State University</td>
</tr>
<tr>
<td>Roland E. Dukes</td>
<td>Cornell University</td>
</tr>
<tr>
<td>William L. Felix, Jr.</td>
<td>University of Washington</td>
</tr>
<tr>
<td>Thomas E. Gibbons</td>
<td>Potter &amp; Company, CPAs</td>
</tr>
<tr>
<td>James Godfrey</td>
<td>University of Michigan</td>
</tr>
<tr>
<td>James L. Goodfellow</td>
<td>Touche Ross &amp; Co.</td>
</tr>
<tr>
<td>Jerry Gottlieb</td>
<td>Laventhal &amp; Horwath</td>
</tr>
<tr>
<td>Richard A. Grimmund</td>
<td>The University of Iowa</td>
</tr>
<tr>
<td>Roy J. Groves</td>
<td>Ernst &amp; Ernst</td>
</tr>
<tr>
<td>Robert E. Hamilton</td>
<td>University of Minnesota</td>
</tr>
<tr>
<td>James A. Heintz</td>
<td>Indiana University</td>
</tr>
<tr>
<td>Neal B. Hitzig</td>
<td>Arthur Young &amp; Company</td>
</tr>
<tr>
<td>Gary L. Holstrum</td>
<td>University of Florida</td>
</tr>
<tr>
<td>Daniel L. Jensen</td>
<td>Purdue University</td>
</tr>
<tr>
<td>Douglas A. Johnson</td>
<td>Arizona State University</td>
</tr>
<tr>
<td>Edward J. Joyce</td>
<td>University of Chicago</td>
</tr>
<tr>
<td>Robert S. Kaplan</td>
<td>Carnegie-Mellon University</td>
</tr>
<tr>
<td>William R. Kinney, Jr.</td>
<td>The University of Iowa</td>
</tr>
<tr>
<td>Richard B. Lea</td>
<td>University of Massachusetts</td>
</tr>
<tr>
<td>W. Morley Lemon</td>
<td>McMaster University</td>
</tr>
<tr>
<td>Donald A. Leslie</td>
<td>Clarkson, Gordon &amp; Company</td>
</tr>
<tr>
<td>Robert Libby</td>
<td>University of Chicago</td>
</tr>
<tr>
<td>John O. Milton</td>
<td>Cooper &amp; Lybrand</td>
</tr>
<tr>
<td>Theodore J. Mock</td>
<td>University of Southern California</td>
</tr>
<tr>
<td>Shane Moriarity</td>
<td>University of Oklahoma</td>
</tr>
<tr>
<td>Philip M. J. Reckers</td>
<td>University of Maryland</td>
</tr>
<tr>
<td>J. Hal Renaud</td>
<td>Arizona State University</td>
</tr>
<tr>
<td>Larry E. Rittenberg</td>
<td>University of Wisconsin-Madison</td>
</tr>
<tr>
<td>Donald M. Roberts</td>
<td>University of Illinois at Urbana-Champaign</td>
</tr>
<tr>
<td>Jack C. Robertson</td>
<td>The University of Texas</td>
</tr>
<tr>
<td>Robert S. Roussey</td>
<td>Arthur Andersen &amp; Co.</td>
</tr>
<tr>
<td>Gerald L. Salamon</td>
<td>The University of Iowa</td>
</tr>
<tr>
<td>Joseph G. San Miguel</td>
<td>Harvard University</td>
</tr>
<tr>
<td>Richard G. Schroeder</td>
<td>Texas A &amp; M University</td>
</tr>
<tr>
<td>William R. Scott</td>
<td>Queens University</td>
</tr>
<tr>
<td>John K. Shank</td>
<td>The Ohio State University</td>
</tr>
<tr>
<td>Matthew D. Shedd</td>
<td>Ernst &amp; Ernst</td>
</tr>
<tr>
<td>Howard F. Steffler</td>
<td>The University of Kansas</td>
</tr>
<tr>
<td>Kenneth W. Stringer</td>
<td>Deloitte, Haskins &amp; Sells</td>
</tr>
<tr>
<td>Robert J. Swierenga</td>
<td>Cornell University</td>
</tr>
<tr>
<td>Michael T. Vanceck</td>
<td>North Texas State University</td>
</tr>
<tr>
<td>Chris Wahlig</td>
<td>McGladrey, Hansen, Dunn &amp; Company</td>
</tr>
<tr>
<td>Wanda A. Wallace</td>
<td>The University of Rochester</td>
</tr>
<tr>
<td>Bart H. Ward</td>
<td>University of Oklahoma</td>
</tr>
<tr>
<td>Carl S. Warren</td>
<td>University of Georgia</td>
</tr>
<tr>
<td>William F. Wright</td>
<td>Pec, Marvin, Mitchell &amp; Co.</td>
</tr>
</tbody>
</table>

Co-Directors:  
Philip E. Fess  
Frederick L. Neumann  
Soong H. Park  
Joseph J. Schultz, Jr.  
Richard E. Ziegler

The proceedings of this and past Symposia are available from:  
The Center for International Education  
and Research in Accounting  
322 Commerce West  
University of Illinois at Urbana-Champaign  
Urbana, Illinois 61801

The Department is grateful for the financial support provided by the Price Waterhouse & Co. Foundation.
Symposium on Auditing Research

By an Audit Group at the University of Illinois
at Urbana-Champaign

Department of Accountancy
University of Illinois at Urbana-Champaign
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preface</td>
<td></td>
<td>v</td>
</tr>
<tr>
<td>Foreword</td>
<td></td>
<td>vi</td>
</tr>
<tr>
<td>Session 1</td>
<td>DEVELOPING A FINANCIAL PLANNING MODEL FOR AN ANALYTICAL REVIEW: A FEASIBILITY STUDY, Robert S. Kaplan</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>DISCUSSANT'S RESPONSE, Neal B. Hitzig</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>DISCUSSANT'S RESPONSE, Richard B. Lea</td>
<td>38</td>
</tr>
<tr>
<td>Session 2</td>
<td>THE EFFECT OF MEASUREMENT ERROR ON REGRESSION RESULTS IN ANALYTICAL REVIEW, William R. Kinney, Jr. and Gerald L. Salamon</td>
<td>49</td>
</tr>
<tr>
<td></td>
<td>DISCUSSANT'S RESPONSE, Robert K. Elliott</td>
<td>65</td>
</tr>
<tr>
<td></td>
<td>DISCUSSANT'S RESPONSE, Wanda Wallace</td>
<td>70</td>
</tr>
<tr>
<td>Session 3</td>
<td>AUDIT TESTS FOR INTERNAL CONTROL RELIANCE, William L. Felix, Jr. and James L. Goodfellow</td>
<td>85</td>
</tr>
<tr>
<td></td>
<td>DISCUSSANT'S RESPONSE, Abraham D. Akresh</td>
<td>106</td>
</tr>
<tr>
<td></td>
<td>DISCUSSANT'S RESPONSE, William A. Scott</td>
<td>113</td>
</tr>
<tr>
<td>Session 4</td>
<td>&quot;SUBJECT TO&quot; AUDIT OPINIONS: A PRELIMINARY INVESTIGATION OF STATEMENT USERS' AND STATEMENT ISSUERS' PERCEPTIONS, John K. Shank and Jesse F. Dillard</td>
<td>121</td>
</tr>
<tr>
<td></td>
<td>DISCUSSANT'S RESPONSE, Douglas R. Carmichael</td>
<td>150</td>
</tr>
<tr>
<td></td>
<td>DISCUSSANT'S RESPONSE, Robert Libby</td>
<td>189</td>
</tr>
<tr>
<td>Session 5</td>
<td>THOUGHTS ON PRACTICAL AUDITING RESEARCH, Ray J. Groves</td>
<td>197</td>
</tr>
<tr>
<td>Session 6</td>
<td>EVALUATING THE COMPETENCE OF INTERNAL AUDIT DEPARTMENTS, Thomas E. Gibbs and Richard G. Schroeder</td>
<td>207</td>
</tr>
<tr>
<td></td>
<td>DISCUSSANT'S RESPONSE, John G. Milton</td>
<td>225</td>
</tr>
<tr>
<td></td>
<td>DISCUSSANT'S RESPONSE, Edward J. Joyce</td>
<td>232</td>
</tr>
<tr>
<td>Session 7</td>
<td>TO BLOW THE WHISTLE OR NOT: AN EMPLOYEE'S DILEMMA WITH INTERNAL CONTROL IMPLICATIONS, Douglas A. Johnson</td>
<td>239</td>
</tr>
<tr>
<td></td>
<td>DISCUSSANT'S RESPONSE, Robert S. Roussey</td>
<td>257</td>
</tr>
<tr>
<td></td>
<td>DISCUSSANT'S RESPONSE, Robert H. Ashton</td>
<td>265</td>
</tr>
</tbody>
</table>
The papers contained in this volume represent the proceedings of the Third Symposium on Auditing Research held at the University of Illinois at Urbana-Champaign on October 26 and 27, 1978. The purpose of this symposium was to provide a forum for the discussion of research on current issues in auditing among practitioners and academicians.

We are especially grateful to the persons whose works compose this volume. Six research papers, each accompanied by two discussants' comments (one from practice and one from academia) and Ray J. Groves' "Thoughts on Practical Auditing Research" make up this volume. The papers were selected to provide for as many different timely issues as practicable. We hope that the interchange of ideas and the avenues developed in these authors' works will continue to result in further stimulating quality research into the many issues confronting the auditing profession.

Finally, we are appreciative of the financial support from the Price Waterhouse & Co. Foundation, which made this symposium possible.

An Audit Group at the University of Illinois at Urbana-Champaign

JOSEPH J. SCHULTZ, JR.
FREDERICK L. NEUMANN
PHILIP E. FESS
SOONG H. PARK
RICHARD E. ZIEGLER
Foreword

Recent social, political, and economic developments have fostered the rise of the accountability function in society. It seems to grow at an ever increasing rate. In response to this rise in accountability, the role of the auditor in society has assumed greater importance and the profession associated with it has, of necessity, had to assume greater responsibilities. Thus, a continuing need exists for the improvement of the auditing function to assure the coordination and systematic operation of social and economic life.

The Third Symposium on Auditing Research represents the Department of Accountancy's effort to gather together scholars and practitioners to examine both opportunities and obligations of the auditing profession. In particular, the symposium seeks to direct attention to the many complex and involved issues associated with the development of the discipline underlying the auditing profession. The Department hopes, through this symposium, to provide an opportunity for scholarly discussion and exchanges of views of interest at the leading edge of auditing thought.

The passage of the Foreign Corrupt Practice Act suggests fundamental thinking is needed to provide a broad based auditing discipline if society is to cope with its many problems. And from this emerges the second main objective of the symposium: to encourage academic auditing scholars to think boldly and to act decisively to the end that the scope and role of auditing is fitted well with the needs of society.

We wish to acknowledge with great appreciation the financial support provided by Price Waterhouse & Co. Foundation which made this symposium possible.

Norton M. Bedford, Head
Department of Accountancy

V. K. Zimmerman, Dean
College of Commerce and
Business Administration
Session

ONE
Developing a Financial Planning Model for an Analytical Review: A Feasibility Study

ROBERT S. KAPLAN

The role of the analytic review in an audit has received considerable attention in the last few years. While always a nominal part of any audit, the use of a formal statistical procedure for the analytic review was not discussed until very recently (Stringer, 1975). Given the acceptance by a large public accounting firm of regression analysis for analytic review and motivated by the financial support offered for research in this area by the Peat, Marwick, Mitchell Research Opportunities in Auditing (1976), research has been initiated on the use of formal statistical models for the analytic review (Kinney, 1977, 1978; Albrecht and McKeown, 1977; Lev, 1978; and Dukes and Swieringa, 1978).

The objective of the analytic review is to test the reasonableness of the reported financial statements in light of the past history of the firm and contemporary conditions in the economy and in the firm's industry. Traditionally this has been accomplished by ratio and trend analysis. Regression analysis provides a more general basis for performing an analytic review. Unlike ratio analysis, regression techniques allow for a fixed as well as a variable component describing the relationship between two variables and permit the use of multiple

This research was supported by a grant from the Peat, Marwick, Mitchell Foundation, through its Research Opportunities in Auditing program. Computational assistance was provided by Bela Gang Dharan and M. Sundaresan. Katherine Schipper gave helpful advice on model estimation.
explanatory variables. Also formal statistical tests can be applied to the estimated coefficients and to the predictions, in the test period, of the variable of interest. Kinney (1977) has demonstrated how the statistical conclusions from a regression-based analytic review can be used to reduce the sample size of a subsequent test of details.

In order to use regression analysis, the auditor must decide on the key relationship or relationships to be investigated. This is no different from what he has traditionally done when selecting the ratios or trends to be computed. Nevertheless, the auditor is still focusing on one relationship at a time and relying on his intuition (or human information capabilities) for aggregating across these diverse tests in order to arrive at a conclusion as to the overall reasonableness of the firm's financial statements. A natural extension, therefore, for an analytic review is to attempt to construct a financial model of the firm which can project complete pro forma income statement and balance sheets based on historical relationships and contemporary economic conditions.

Financial planning models provide a comprehensive and integrated set of relationships among the various accounts in a set of financial statements. Starting from a key input variable, such as sales, the expected values of many other accounts can be estimated. These other accounts include cost of goods sold, inventory, accounts receivable, accounts payable, overhead, selling and administrative expenses, and wages and salaries. From these accounts, an estimate of funds and profit from operations can be obtained and integrated with investment and financial transactions to obtain pro forma financial statements. These pro forma statements can then be compared with the statements prepared by management and significant deviations highlighted. If the pro forma statements generally agree with those supplied by management, the auditor can have increased assurance about the overall reasonableness of the financial statements. Significant deviations between pro forma accounts and management-prepared statements provide the basis for further investigation to determine the source of the discrepancy. Financial planning models could therefore provide an excellent basis for the analytic review in the annual audit. In addition, they could be used for reviewing the interim statements of a company and for assessing the realism of any forecasted statements a company may choose to prepare.

In this paper we explore the feasibility and judgments required to build a financial planning model that will produce pro forma statements of sufficient accuracy that they can be used by auditors for their analytic reviews and perhaps to reduce reliance on other audit tests.
We will attempt to use relatively straightforward statistical procedures to serve as prototypes of what could actually be developed in the field without requiring the active help of a Ph.D. in econometrics.

Financial planning models began to be developed about fifteen years ago. Early references include Gershetski (1969), Schrieber (1970), and Boulden and Buffa (1970). In a survey at the beginning of this decade Gershetski (1970) indicated that less than 100 companies were using such models, but their popularity has expanded considerably in the last few years. A more recent survey by Naylor and Gattis (1976) estimated that nearly 2000 firms in the United States, Canada, and Europe are developing, experimenting with, or using a corporate planning model. Such models are typically used by top management, and while most using organizations have sales in excess of $100 Million, at least 7 percent of firms had sales below $50 Million. These models are used for cash flow analysis, financial forecasting, balance sheet projections, financial analysis, preparing pro forma financial statements and profit planning (Naylor and Gattis, 1976; see also Naylor and Schauland, 1976).

Financial planning models have usually been constructed for planning horizons of five to ten years. They have been developed mainly for internal planning purposes although bankers are beginning to use financial planning models to project refunding capabilities when analyzing a corporate loan request. But there has yet to be any reported description of using financial planning models in the audit process.

To be useful in the analytic review phase of an audit, financial planning models must have different characteristics than the traditional models used by firms for their internal planning function. We have already mentioned the desirability of keeping the estimation of these models relatively simple and inexpensive so that they have a chance of being widely used in practice. The time frame for an audit-based financial planning model will be much shorter than models built for internal planning. Auditors typically concentrate on a one year time period, and it seems sensible to develop a model with this time frame in mind. In addition, at the time of the audit, all, or at least most, of the year of interest will have passed so that the auditor can use the values of many discretionary variables (such as new financing or investment) that must otherwise be forecasted in traditional applications of financial planning models. A further advantage is knowing the monthly sales figures rather than having to estimate this key variable. Offsetting these significant advantages, the accuracy required from a financial planning to be useful for an audit is much
higher than that obtained from traditional planning models with their time horizons of five to ten years. Because of the stringent materiality standards in an audit, important account balances must probably be estimated with errors of less than five percent. In some cases, one percent accuracy may be required.

These characteristics of financial planning models for analytic review parallel the situation for statistical sampling in auditing. When sampling, the auditor has much more information than is typically available in survey sampling. This information can be used as the basis for extensive stratification and for auxiliary information estimation (see Kaplan, 1973) but the accuracy requirements from the statistical sampling procedure are much more severe than one encounters in nonauditing sampling applications. Therefore, special statistical procedures (see Anderson-Teitlebaum, 1973) have been developed for audit sampling to exploit the greater information available to auditors and meet the stringent precision dictated by the materiality standard.

These considerations guided the modeling approach used in this study. There is no question that the most important variable for understanding the activity of a firm during a period is its sales figure. From this figure, we can estimate cost of sales, overhead costs and various current asset and liability accounts. Since net income is a residual figure obtained after subtracting expenses from revenues, small errors in estimating either expenses or revenues will lead to extremely large percentage errors in net income. Therefore, in this first effort at building a financial model, we will treat sales as an exogenous variable and not try to estimate it within the model. A further investigation could attempt to estimate the monthly sales figure using time series, cross-sectional, or structural methods. We will take the sales figure as given and estimate expenses and account balances using the reported sales number. This procedure gives us the best opportunity to produce an accurate net income figure.

A second judgment is not to go beyond predicting net income before taxes. The reported tax expense of a company is influenced by many complex factors which are unlikely to be stationary over time. Considerations such as the investment tax credit, accelerated depreciation, foreign tax credit, and the effect of the inventory costing method makes the relationship between net income before taxes and reported tax expense a complicated function. Estimating the tax expense of a company as a function of operations and investment policy could justify a separate study.
Third, the structure of the long-term assets and liabilities of a firm are controllable by discretionary management decisions. While much research has been performed on the investment and financing decisions of firms, the output of this research is not (and may never be) sufficiently accurate to predict the annual or monthly changes in the fixed asset and long-term liability accounts. Since it should be relatively easy to audit these accounts by traditional methods, we have not attempted to model changes in these accounts in this study.

DATA
Monthly balance sheets and income statements were obtained from a large diversified industrial company. Initially, we estimated the model on thirty-six months of data from January 1974 to December 1976. Data from January to December 1977 are used as a test period. One advantage of developing these models is that new data are continually being generated. (One disadvantage is that companies occasionally make major acquisitions or divestitures so that historical data are no longer representative of future relationships.)

The income statement for this company with 1976 percentages of sales are the following:

<table>
<thead>
<tr>
<th>Item</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Net sales</td>
<td>100</td>
</tr>
<tr>
<td>Cost of sales</td>
<td>(85)</td>
</tr>
<tr>
<td>Depreciation</td>
<td>(2)</td>
</tr>
<tr>
<td>Gross margin</td>
<td>13</td>
</tr>
<tr>
<td>Selling expense</td>
<td>(4)</td>
</tr>
<tr>
<td>Administrative expense</td>
<td>(4)</td>
</tr>
<tr>
<td>Research expense</td>
<td>(1)</td>
</tr>
<tr>
<td>Earnings from operations</td>
<td>5</td>
</tr>
<tr>
<td>Miscellaneous expense</td>
<td>(2)</td>
</tr>
<tr>
<td>Interest expense</td>
<td>(2)</td>
</tr>
<tr>
<td>Net income before taxes</td>
<td>1</td>
</tr>
</tbody>
</table>

The major current asset categories with December 31, 1976, percentages of total current assets, are as follows:

<table>
<thead>
<tr>
<th>Item</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cash</td>
<td>(3)</td>
</tr>
<tr>
<td>Accounts receivable</td>
<td>(34)</td>
</tr>
<tr>
<td>Inventories</td>
<td>(82)</td>
</tr>
<tr>
<td>Less: LIFO Reserve</td>
<td>25</td>
</tr>
<tr>
<td>Net inventories</td>
<td>(57)</td>
</tr>
<tr>
<td>Prepaid expenses</td>
<td>(7)</td>
</tr>
<tr>
<td>Total current assets</td>
<td>(100)</td>
</tr>
</tbody>
</table>
The only current liability that will be estimated is the accounts payable account.

ESTIMATING THE MODEL

Cost of Goods Sold
Clearly, obtaining an accurate forecast of the cost of goods sold account is vital if we hope to obtain a reasonably accurate forecast of net income before taxes. We estimated the model:

$$\text{CGS}_t = b_0 + b_1 \cdot \text{Sales}_t,$$

and for various subperiods obtained the following coefficient estimates.*

<table>
<thead>
<tr>
<th>Time Period</th>
<th>Number of Observations</th>
<th>$b_0$</th>
<th>$b_1$</th>
<th>$R^2$</th>
<th>CGS/Sales</th>
</tr>
</thead>
<tbody>
<tr>
<td>1974-76</td>
<td>36</td>
<td>4949</td>
<td>.771</td>
<td>.88</td>
<td>.833</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.36)</td>
<td>(16.1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1975-76</td>
<td>24</td>
<td>-701</td>
<td>.849</td>
<td>.96</td>
<td>.839</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.27)</td>
<td>(23.9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1976</td>
<td>12</td>
<td>17756</td>
<td>.620</td>
<td>.73</td>
<td>.848</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2.07)</td>
<td>(5.58)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*(t-statistics appear in parentheses below the estimated coefficients)*

The explanatory power of the simple regression is very high, as one might expect, but the estimates of the constant and the slope coefficient are highly dependent on the choice of a time period. The ratio of cost of goods sold to sales (suppressing the constant term in the regression) is much more stable.

Initial tests on a holdout sample of the first six months in 1977 showed that the average ratio approach gave the more accurate results than any of the regression models with a constant term. This result was confirmed when data from 1977 became available for estimation:

<table>
<thead>
<tr>
<th>Time Period</th>
<th>Number of Observations</th>
<th>$b_0$</th>
<th>$b_1$</th>
<th>$R^2$</th>
<th>CGS/Sales</th>
</tr>
</thead>
<tbody>
<tr>
<td>1975-77</td>
<td>56</td>
<td>479</td>
<td>.835</td>
<td>.96</td>
<td>.841</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.23)</td>
<td>(29.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1976-77</td>
<td>24</td>
<td>10820</td>
<td>.708</td>
<td>.89</td>
<td>.846</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2.65)</td>
<td>(13.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1977</td>
<td>12</td>
<td>8605</td>
<td>.738</td>
<td>.94</td>
<td>.845</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.86)</td>
<td>(12.8)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
As before, the coefficients are strongly affected by the time period but
the ratio of Cost of Goods Sold to Sales stays remarkably constant. We
decided that a reasonable model would be to assume that Cost of
Goods Sold varied strictly proportionately with sales and use the
1974-76 ratio:

\[ \text{CGS} = 0.84 \times \text{Sales}. \]

**Depreciation**

Depreciation should be a function of the fixed asset account. For this
company, there was a single fixed asset account for buildings, ma-
chinery, and equipment (BM&E). We estimated the model:

\[ \text{Depreciation} = c_0 + c_1 \times \text{Gross BM&E}. \]

The constant was insignificant in each regression as would be ex-
pected given the nature of the account. The \( R^2 \) of the regressions was
around 0.60. A refinement of estimating depreciation each month as a
function of the average BM&E balance (by averaging the beginning
and ending monthly amount) did not improve the results at all. As
with the cost of goods sold account, we will assume a strictly propor-
tional relationship based on the average ratio of depreciation to
BM&E:

\[ \text{Depreciation} = 0.00325 \times \text{Gross BM&E}. \]

A monthly depreciation rate of .00325 implies an average depreciable
life of 25 years. If the company were to start shifting its fixed asset mix
to shorter lived assets, the above historic relationship would under-
estimate depreciation expense. Also an \( R^2 \) of 0.60 indicates a strong
relationship between depreciation and Gross BM&E but that a con-
siderable degree of unexplained monthly fluctuations occurs in de-
preciation expense.

**Selling, Research, and Administrative Expenses**

Separate regressions were run for each of these general and adminis-
trative expense categories. The independent variables were sales and a
linear time trend. The linear time trend variable explains increases in
the fixed cost component over time. The variable coefficient of sales
was significant in the regressions for selling expense (0.9 percent of
sales) and for research expense (0.3 percent of sales) but not for
administrative expense. The fixed cost component and linear time
trend coefficient were significant for all three regressions. For con-
venience, we will display the estimated relationship just for the
aggregate of these three accounts:
SR & A Expense \(= 3475 + 67.2 \times T + .0122 \times \text{Sales}; \overline{R^2} = .68.\) (5.44) (8.66) (1.55)

The time trend variable \(T\) equals 1 for January 1974, 2 for February 1974, \ldots, and 36 for December 1976. For relatively short periods of time this specification is probably adequate, but one should not use this relationship to forecast indefinitely into the future. The linear time trend implies a constantly decreasing percentage increase in fixed costs which seems unrealistic unless the company can continually contain general inflationary increases in its costs. We tried a theoretically superior specification using indices of inflation instead of the time trend variable, but the results did not improve.

Upon examining the residuals, we observed that the residuals for December 1974 and December 1975 were twice as large as any other residual. These unusually large expenses occurred in all three components of SR&A expenses. While unusually large residuals should be investigated whenever they occur (indeed, that is a principal purpose of an analytic review), the occurrence of these residuals in December suggested that special accruals were being made in these fiscal end-of-year months. If true, the observations from these two months would not be comparable with the other months in the estimation period. A regression using the remaining thirty-four months yielded:

\[
\text{SR & A Expense} = 2917 + 68.9 \times T + .018 \times \text{Sales}; \overline{R^2} = .86. \\
(7.21) (14.1) (3.64)
\]

A regression for 23 months in 1975-76, excluding December 1975, yielded:

\[
\text{SR & A Expense} = 2697 + 65.8 \times T + .022 \times \text{Sales}; \overline{R^2} = .74. \\
(4.97) (5.85) (2.77)
\]

These regressions have higher explanatory power, and the variable coefficient of sales is larger, implying a stronger proportional relationship than indicated by the first set of regressions. (1977 data gave an even higher coefficient to the sales variable.) Both of the above regressions are close to each other, and there will be little difference in the predictive ability between them. We will display results using the coefficients from the more recent data:

\[
\text{SR & A Expense} = 2700 + 66 \times T + .022 \times \text{Sales}. 
\]

The procedure of excluding months in the estimation period with large residuals should be part of the judgment used when constructing models for analytic reviews. The model is being estimated on unaudited data. We are attempting to construct relationships based
on normal operations with the expectation that unusual deviations in a month will be subject to special investigation. Particularly when unusual results occur in the closing month of the year, we would not want our estimated relationship to be affected by one or two outliers.

**Interest Expense**

The interest expense category requires a different type of model. A simple regression of interest expense versus interest bearing debt would ignore a changing mix between short-term and long-term debt, the difference in interest rate between debt of different maturities, and changing interest rate on short-term debt. This category requires an explicit recognition of external factors such as the prime interest rate, and careful definition of the different categories of debt. For example, included in the long-term debt balance sheet account was a significant amount of bank debt with an interest rate keyed to the prime lending rate. Included in the bank account in current liabilities is the currently maturing portion of long-term debt. Further investigation revealed that this debt is the long-term bank debt with less than one year to maturity. Consequently, its interest rate would be keyed to the prime rate as with other short-term debt. In general, however, if some portion of the long-term bonds were maturing within a year, the appropriate interest rate for this debt, classified as current liability, would be the historical interest rate on the bond (most likely the coupon rate) and not the prime rate.

The company had three long-term bond issues. The monthly interest expense from each of these bonds was computed simply as

\[
\text{Book Value of Bond} \times \text{Coupon Rate}/12.
\]

The long-term interest expense (LTIE) was obtained by summing the monthly interest expense from each bond.

The short-term interest expense (STIE) was computed by summing together the notes payable, short- and long-term, and the current portion of long-term debt and multiplying this sum by the average prime interest rate for the month. The prime rate is determined from the *Federal Reserve Bulletin*. In the second half of 1976, the company obtained an additional $30 million term loan at an interest rate .5 percent above prime. This higher interest expense was included in the STIE computation.

The final category of debt was long-term lease obligations. The bulk of these arose from leasing property that was acquired by municipally financed industrial development loans. Consequently, the
interest rate on these leases was below the company's normal borrowing rate. An average rate of 6 percent was assumed for these obligations.

The monthly interest expense is therefore estimated as:

$$\text{Interest Expense} = \text{LTIE} + \text{STIE} + (.06/12) \times \text{Lease Obligations}.$$  

Notice that unlike the previous accounts, we are not using any statistical analysis to estimate interest expense. It seems much more appropriate to estimate this account directly from a simple model rather than from a historical regression analysis.

There are a number of factors that serve to limit the accuracy of this simple model. First, there could be considerable fluctuation in short-term debt within a month that will never show up on the end of month balance sheet. Such fluctuations will affect monthly interest expense but will be difficult to trace. Second, the company has many semi-autonomous divisions, and the loans to the different divisions could be made at various interest rates rather than the single prime rate we assumed. Third, the prime rate is a nationwide average. It may differ across banks and may fluctuate within the month. Given that the interest expense is only 1 to 2 percent of sales, however, it did not seem cost effective to attempt to build any more elaborate models for this account. Large deviations between predicted and actual expense could be investigated during the audit.

During the "estimation period" of 1974-76, monthly errors of as high as 10 percent did occur. On an annual basis, cumulative predicted interest overstated actual interest expense by 5.4 percent in 1974, understated actual interest expense by 2.4 percent in 1975, and understated interest expense by 5.8 percent in 1976. Interest expense was consistently underestimated during the last seven months of 1976, suggesting a shift in the relationship which deserves further scrutiny.

This completed our estimation of income statement accounts. Miscellaneous income and expense was not forecasted due to the erratic nature of this account. The next phase was to estimate current asset and liability accounts.

**Cash**

A variety of relationships was explored to explain the variation in the cash account. Among the independent variables that were tried were sales, cost indices, and a trend variable. Ultimately, the model that performed as well as any assumed that cash followed a random
walk (strictly speaking, a martingale process) so that we predicted cash in a month by its level in the preceding month:

\[ \text{CASH}_t = \text{CASH}_{t-1}. \]

**Accounts Receivable**

We performed regressions with either accounts receivable (AR) or collections (C) as the dependent variable. A better fit was obtained when we explained collections as a function of past sales and inferred accounts receivable from the accounting identity. The following model provided a good fit:

\[
C_t = -808 + .2 \times S_t + .49 \times S_{t-1} + .31 \times S_{t-2}; R^2 = 0.84.
\]

\[
(0.56) \quad (1.14) \quad (2.39)
\]

We required the coefficients on the sales variables to sum to one so that all sales would eventually be collected (bad debt expense was negligible). The treasury bill rate was tried as an additional independent variable to explore whether collections might be delayed when short-term interest rates were high. The variable proved to be insignificant.

To obtain accounts receivable in month \( t \) (AR\(_t\)) we used the identity:

\[ \text{AR}_t = \text{AR}_{t-1} + S_t - C_t. \]

and, ignoring the insignificant constant term, we have:

\[ \text{AR}_t = \text{AR}_{t-1} + .8 \times S_t - .49 \times S_{t-1} - .31 \times S_{t-2}. \]

**Inventory**

We started by modeling the flow account, Purchases (PUR). This variable was obtained from the balance sheet inventory account (I\(_t\)) using the accounting identity:

\[ \text{PUR}_t = I_t - I_{t-1} + CGS_t. \]

The simple linear regression of purchases on sales gave a reasonable fit:

\[
\text{PUR}_t = 4322 + .778 \times S_t; R^2 = .55.
\]

\[
(0.48) \quad (6.25)
\]

Lagged sales variables were also tried as independent variables as a surrogate for forecasting expected sales in the next period, but these variables did not improve the explanatory power of the regression. The relatively low \( R^2 \) is caused by two large residuals at the end of 1974 (the onset of the recession) when sales dropped unexpectedly. By ignoring the constant term, so that purchases are assumed propor-
tional to sales we obtained:

\[ \text{PUR}_t = 0.83 \times S_t. \]

Using the accounting identity

\[ I_t = I_{t-1} + \text{PUR}_t - \text{CGS} \]

and recalling that

\[ \text{CGS} \equiv 0.84 S_t, \]

we are left with a simple random walk model as the best predictor for inventory

\[ I_t = I_{t-1}. \]

This effect was confirmed by a simple regression of inventory on sales in which the sales variable was insignificant and the residuals were highly autocorrelated.

Modeling the inventory relationship is challenging since there are two disparate reasons why inventories might vary with current sales. First, if the company anticipates increases in sales in the near future, it could start increasing production now to smooth out the anticipated demand on production. The second reason is quite different. The company could be experiencing an unexpected drop in sales (such as happened in late 1974). Until this sales slump is identified as something other than a transitory fluctuation, production keeps putting out finished good inventory at the previous rates and inventories will expand relative to sales. Thus, there are two quite different behavioral models leading to inventory changes, and it will be difficult to capture both effects with a simple model. This inventory equation problem apparently plagues complex macro-models of the economy also. For our purposes, it makes sense to use a relatively simple model and leave it to the auditor's judgment to understand which circumstance is occurring when a deviation from the model is signaled.

**Prepaid Expense**

A simple time trend

\[ \text{Prepaid Expense}_t = 11250 + 278 \times t; \quad R^2 = .76 \]

\( (20.2) \quad (10.6) \)

gave the best fit for this account. There was no correlation between prepaid expense and sales.

**Accounts Payable**

Modeling the accounts payable account (AP) incorporated a number of ideas and relationships used in estimating other accounts. We start
with the accounting identity relating changes in accounts payable to purchases (\(\text{PUR}_t\)) and payments (\(\text{PAY}_t\)):

\[
\text{AP}_t = \text{AP}_{t-1} + \text{PUR}_t - \text{PAY}_t.
\]

We model payments in month \(t\) as a function of accounts payable at the start of the month \(t\) and purchases during month \(t\). This led to a model of payments as:

\[
\text{PAY}_t = 12300 + .662 \times \text{AP}_{t-1} + .324 \times \text{PUR}_t; R^2 = .81.
\]

(2.57) (6.82) (5.29)

As we would have expected from our prior specification, replacing the independent variable, purchases, with sales lowered the explanatory power of the model. Substituting back into the accounting identity yielded the accounts payable equation:

\[
\text{AP}_t = -12300 + .398 \times \text{AP}_{t-1} + .676 \times \text{PUR}_t.
\]

When predicting accounts payable in a month, one can use either actual purchases for the month or the predicted purchases from the relationship

\[
\text{PUR}_t = .83 \times \text{S}_t.
\]

The estimated equation for accounts payable does give a better fit than assuming a random walk for this account. The regression

\[
\text{AP}_t = b_0 + b_1 \text{AP}_{t-1}
\]

had a \(R^2\) of .64. Using the predicted accounts payable as the independent variable, with purchases estimated from sales each month, yielded a regression with an \(R^2\) of .68. Using actual purchases rather than predicted purchases gave a much better fit, with the \(R^2\) rising to .86. Thus, the effort of building an explicit model for accounts payable based on purchases and payments enabled us to obtain a good prediction equation for this account.

This completes the estimation of the financial model. We now examine the predictions in a test year. By comparing the predictions from the model with actual results in 1977, we can simulate the output an auditor might receive after having constructed a model based on 1974-76 data.

**Predictions from the Model**

The estimated equations for the income statement and balance sheet were applied to a new set of data from calendar year 1977. This exercise replicates how the model, estimated on historic 1974-76 data, could be used during the 1977 audit. As in the estimation phase of the project, we assume that the monthly sales figure is known as well as the monthly gross plant and equipment, and the composition of
short- and long-term debt. Exhibits la-ld present a comparison of the actual with forecasted income statements and balance sheet accounts. The income statement on a quarterly basis is also presented in each table, with each predicted category shown as a percentage of sales. The annual income statement comparison appears in exhibit ld. At the company's request, the data are disguised to maintain confidentiality. Thus the sales figures are normalized so that 1977 sales equals 10000. The expense data are scaled to maintain their actual percentages to sales. Balance sheet accounts are scaled so that the reported January balance in each account is 10000.

It is clear, from examining the income statements in exhibits la-ld, that while there are occasionally large deviations for some accounts on a monthly basis, the quarterly and annual forecasts come close to the actual statement. The net income before taxes for the entire year is forecasted to within about 10 percent of actual net income. Thus using a model developed on historical data and projecting for a new calendar period using only data on monthly sales, gross investment, and interest bearing liabilities, we are able to forecast annual income and expense data quite accurately. Of course, the accuracy at the annual level, to within 10 percent of the true figure, is an average of much larger errors on a quarterly or monthly basis which tend to cancel each other out when added together. For example, net income before tax is overestimated by 80 percent in quarter 4 and by 47 percent in quarter 2, is underestimated by 37 percent in quarter 1, and is predicted to within 1 percent in quarter 3. These relatively large percentage errors are not unexpected given the relatively low percentage (3 to 4 percent) of pretax profit to sales in this company. The errors as a percentage of sales or as a percentage of total assets are much smaller. The cancellation of errors suggests that the monthly errors are due to random fluctuations rather than any consistent bias in the model.

Whether the accuracy of the projected income statements is sufficient evidence for the auditor to rely on or perhaps reduce other tests or procedures during the quarterly review or the annual audit remains to be answered by auditors. I am interested in the remarks of my discussants on this issue. I doubt whether one can regularly expect to obtain accuracy much better than what is reported for the income and expense accounts in exhibits la - ld. Given that monthly closings may not be done with the care of quarterly or annual closings, it is surprising how small the monthly errors are for individual accounts.
## Exhibit 1a. Actual (Act.) and Forecasted (F/C) Results—Selected Income Statement and Balance Sheet Items: First Quarter 1977

<table>
<thead>
<tr>
<th></th>
<th>January</th>
<th>February</th>
<th>March</th>
<th>1st Quarter</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Act.</td>
<td>F/C</td>
<td>Act.</td>
<td>F/C</td>
</tr>
<tr>
<td>Sales</td>
<td>870</td>
<td>870</td>
<td>807</td>
<td>807</td>
</tr>
<tr>
<td>Cost of sales</td>
<td>736</td>
<td>731</td>
<td>651</td>
<td>678</td>
</tr>
<tr>
<td>Depreciation</td>
<td>18</td>
<td>18</td>
<td>17</td>
<td>18</td>
</tr>
<tr>
<td>Gross profit</td>
<td>114</td>
<td>121</td>
<td>139</td>
<td>111</td>
</tr>
<tr>
<td>Comm'L., adm. and res. expense</td>
<td>84</td>
<td>80</td>
<td>75</td>
<td>78</td>
</tr>
<tr>
<td>Operating profit</td>
<td>30</td>
<td>41</td>
<td>66</td>
<td>33</td>
</tr>
<tr>
<td>Interest expense</td>
<td>10</td>
<td>13</td>
<td>14</td>
<td>17</td>
</tr>
<tr>
<td>Pretax profit</td>
<td>20</td>
<td>28</td>
<td>51</td>
<td>16</td>
</tr>
<tr>
<td>Cash</td>
<td>1000</td>
<td>985</td>
<td>659</td>
<td>1000</td>
</tr>
<tr>
<td>Accounts receivable</td>
<td>1060</td>
<td>790</td>
<td>1017</td>
<td>989</td>
</tr>
<tr>
<td>Prepaid expenses</td>
<td>1000</td>
<td>945</td>
<td>869</td>
<td>957</td>
</tr>
<tr>
<td>Inventory, LIFO</td>
<td>1000</td>
<td>985</td>
<td>1031</td>
<td>1000</td>
</tr>
<tr>
<td>Accounts payable</td>
<td>1000</td>
<td>918</td>
<td>866</td>
<td>960</td>
</tr>
</tbody>
</table>
### Exhibit 1b. Actual (Act.) and Forecasted (F/C) Results—Selected Income Statement and Balance Sheet Items: Second Quarter 1977

<table>
<thead>
<tr>
<th></th>
<th>April</th>
<th></th>
<th>May</th>
<th></th>
<th>June</th>
<th></th>
<th>2nd Quarter</th>
<th></th>
<th>(% of Sales)</th>
<th></th>
<th>(% of Sales)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sales</td>
<td>884</td>
<td>884</td>
<td>752</td>
<td>752</td>
<td>685</td>
<td>685</td>
<td>2321</td>
<td>(100)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>745</td>
<td>743</td>
<td>641</td>
<td>631</td>
<td>603</td>
<td>575</td>
<td>1989</td>
<td>(85.7)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depreciation</td>
<td>20</td>
<td>18</td>
<td>18</td>
<td>18</td>
<td>18</td>
<td>18</td>
<td>56</td>
<td>(2.4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>53</td>
<td>(2.3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gross profit</td>
<td>119</td>
<td>124</td>
<td>93</td>
<td>103</td>
<td>64</td>
<td>92</td>
<td>276</td>
<td>(11.9)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comm'1., adm. and</td>
<td>81</td>
<td>78</td>
<td>61</td>
<td>75</td>
<td>66</td>
<td>73</td>
<td>268</td>
<td>(9.0)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>res. expense</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>226</td>
<td>(9.7)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Operating profit</td>
<td>38</td>
<td>45</td>
<td>32</td>
<td>28</td>
<td>(2)</td>
<td>19</td>
<td>68</td>
<td>(2.9)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interest expense</td>
<td>17</td>
<td>15</td>
<td>18</td>
<td>15</td>
<td>13</td>
<td>14</td>
<td>42</td>
<td>(1.8)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pretax profit</td>
<td>21</td>
<td>30</td>
<td>19</td>
<td>13</td>
<td>(15)</td>
<td>5</td>
<td>26</td>
<td>(1.1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cash</td>
<td>693</td>
<td>784</td>
<td>1022</td>
<td>693</td>
<td>1065</td>
<td>1022</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accounts receivable</td>
<td>1115</td>
<td>934</td>
<td>1052</td>
<td>1062</td>
<td>822</td>
<td>1141</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prepaid expenses</td>
<td>825</td>
<td>981</td>
<td>866</td>
<td>993</td>
<td>702</td>
<td>1005</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inventory, LIFO</td>
<td>1095</td>
<td>1044</td>
<td>1087</td>
<td>1095</td>
<td>1109</td>
<td>1087</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accounts payable</td>
<td>903</td>
<td>932</td>
<td>879</td>
<td>816</td>
<td>921</td>
<td>1029</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>July</td>
<td>August</td>
<td>September</td>
<td>3rd Quarter</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------</td>
<td>---------------</td>
<td>---------------</td>
<td>----------------</td>
<td>----------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sales</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sales</td>
<td>967</td>
<td>967</td>
<td>789</td>
<td>789</td>
<td>827</td>
<td>827</td>
<td>2582 (100)</td>
<td>2582</td>
<td>(100)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cost of sales</td>
<td>814</td>
<td>812</td>
<td>670</td>
<td>663</td>
<td>690</td>
<td>694</td>
<td>2174 (84.2)</td>
<td>2169</td>
<td>(84)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depreciation</td>
<td>19</td>
<td>17</td>
<td>18</td>
<td>17</td>
<td>18</td>
<td>17</td>
<td>55 (2.1)</td>
<td>52</td>
<td>(2.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gross profit</td>
<td>134</td>
<td>137</td>
<td>100</td>
<td>109</td>
<td>118</td>
<td>115</td>
<td>358 (13.7)</td>
<td>361</td>
<td>(14.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comm', adm. and res.</td>
<td>79</td>
<td>78</td>
<td>71</td>
<td>74</td>
<td>71</td>
<td>74</td>
<td>221 (8.6)</td>
<td>226</td>
<td>(8.8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Operating profit</td>
<td>55</td>
<td>59</td>
<td>30</td>
<td>35</td>
<td>47</td>
<td>40</td>
<td>182 (5.1)</td>
<td>195</td>
<td>(5.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interest expense</td>
<td>16</td>
<td>13</td>
<td>13</td>
<td>15</td>
<td>13</td>
<td>15</td>
<td>41 (1.6)</td>
<td>43</td>
<td>(1.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pretax profit</td>
<td>89</td>
<td>46</td>
<td>17</td>
<td>20</td>
<td>34</td>
<td>25</td>
<td>90 (3.5)</td>
<td>91</td>
<td>(3.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cash</td>
<td>881</td>
<td>1065</td>
<td>1126</td>
<td>881</td>
<td>1153</td>
<td>1126</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accounts receivable</td>
<td>873</td>
<td>706</td>
<td>945</td>
<td>850</td>
<td>787</td>
<td>1031</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prepaid expenses</td>
<td>778</td>
<td>1018</td>
<td>788</td>
<td>1090</td>
<td>744</td>
<td>1042</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inventory, LIFO</td>
<td>1110</td>
<td>1109</td>
<td>1149</td>
<td>1110</td>
<td>1120</td>
<td>1149</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accounts payable</td>
<td>893</td>
<td>766</td>
<td>1094</td>
<td>881</td>
<td>923</td>
<td>919</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>October</td>
<td>November</td>
<td>December</td>
<td>4th Quarter</td>
<td>Annual</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------------------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
<td>-------------------</td>
<td>-----------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sales</td>
<td>981</td>
<td>981</td>
<td>794</td>
<td>794</td>
<td>820</td>
<td>820</td>
<td>2595</td>
<td>(100)</td>
<td>2595</td>
<td>(100)</td>
<td>10000</td>
</tr>
<tr>
<td>Cost of sales</td>
<td>807</td>
<td>824</td>
<td>697</td>
<td>657</td>
<td>718</td>
<td>689</td>
<td>2217</td>
<td>(85.4)</td>
<td>2180</td>
<td>(84)</td>
<td>8453</td>
</tr>
<tr>
<td>Depreciation</td>
<td>18</td>
<td>17</td>
<td>18</td>
<td>17</td>
<td>19</td>
<td>17</td>
<td>55</td>
<td>(2.1)</td>
<td>51</td>
<td>(2.0)</td>
<td>218</td>
</tr>
<tr>
<td>Gross profit</td>
<td>156</td>
<td>146</td>
<td>79</td>
<td>110</td>
<td>88</td>
<td>114</td>
<td>323</td>
<td>(12.5)</td>
<td>364</td>
<td>(14.0)</td>
<td>1328</td>
</tr>
<tr>
<td>Comm', adm. and res. exp</td>
<td>79</td>
<td>76</td>
<td>76</td>
<td>72</td>
<td>71</td>
<td>72</td>
<td>226</td>
<td>(8.7)</td>
<td>220</td>
<td>(8.5)</td>
<td>893</td>
</tr>
<tr>
<td>Operating profit</td>
<td>77</td>
<td>63</td>
<td>8</td>
<td>38</td>
<td>17</td>
<td>43</td>
<td>97</td>
<td>(3.7)</td>
<td>144</td>
<td>(5.5)</td>
<td>435</td>
</tr>
<tr>
<td>Interest expense</td>
<td>15</td>
<td>14</td>
<td>13</td>
<td>14</td>
<td>13</td>
<td>14</td>
<td>42</td>
<td>(1.6)</td>
<td>48</td>
<td>(1.6)</td>
<td>164</td>
</tr>
<tr>
<td>Pretax profit</td>
<td>62</td>
<td>49</td>
<td>(10)</td>
<td>24</td>
<td>4</td>
<td>29</td>
<td>56</td>
<td>(2.1)</td>
<td>102</td>
<td>(3.9)</td>
<td>271</td>
</tr>
<tr>
<td>Cash</td>
<td>692</td>
<td>1153</td>
<td>1346</td>
<td>692</td>
<td>864</td>
<td>1346</td>
<td>692</td>
<td>1153</td>
<td>864</td>
<td>1346</td>
<td></td>
</tr>
<tr>
<td>Accounts receivable</td>
<td>1013</td>
<td>783</td>
<td>1122</td>
<td>991</td>
<td>966</td>
<td>1152</td>
<td>1013</td>
<td>783</td>
<td>966</td>
<td>1152</td>
<td></td>
</tr>
<tr>
<td>Prepaid expenses</td>
<td>777</td>
<td>1034</td>
<td>898</td>
<td>1066</td>
<td>771</td>
<td>1079</td>
<td>777</td>
<td>1034</td>
<td>771</td>
<td>1079</td>
<td></td>
</tr>
<tr>
<td>Inventory, LIFO</td>
<td>1048</td>
<td>1120</td>
<td>1066</td>
<td>1048</td>
<td>1091</td>
<td>1066</td>
<td>1048</td>
<td>1120</td>
<td>1091</td>
<td>1066</td>
<td></td>
</tr>
<tr>
<td>Accounts payable</td>
<td>777</td>
<td>697</td>
<td>953</td>
<td>896</td>
<td>947</td>
<td>976</td>
<td>777</td>
<td>697</td>
<td>947</td>
<td>976</td>
<td></td>
</tr>
</tbody>
</table>
Probably we are fortunate in examining a company with excellent procedures for accumulating and reporting data on a monthly basis.

Since all the equations apart from interest expense were obtained from statistical analysis of historical data, it is possible to compute the standard error of forecast for the monthly expense account projections. This was not done somewhat due to time limitations but also because almost all the projections, particularly in the critical cost of sales account, were reasonably close to actual values and therefore almost surely within a 2σ confidence interval. Research needs to be performed to integrate the information from the standard error of forecast for individual accounts into, first, an overall conclusion on the reasonableness of the income statement and, second, on the degree that other audit tests can be reduced based on the output from the financial model. Kinney [1977] has provided an excellent start in this direction by showing how a regression analysis on an individual account can be used to reduce subsequent tests of details on that same account. Considerably more opportunities for research remain to be done in this area.

The forecasted monthly current asset and accounts payable balances are also shown in exhibits 1a-1d. Unlike the generally excellent fit that was obtained for income and expense accounts, there are large deviations between the forecasted and actual balance sheet accounts. Of course, for the cash and inventory accounts, where we are forecasting the balance in month t to be the same as the balance in the previous month (t − 1), we are admitting our inability to explain monthly changes in these accounts. For the cash account, the average monthly error can be easily computed by subtracting the cash balance in December 1976 (985) from the level in December 1977 (864) and dividing by the number of intervening months (12). Thus the average error is −10. The standard deviation of the monthly errors (computed from the monthly deviation from the actual mean error of −10) is 384, or about 35 percent of the average 1977 cash level. Thus, there is considerable unexplained fluctuation about the average cash level. This unexplained fluctuation is consistent with variations experienced in the estimation period (1974-76).

For the (LIFO) inventory account, the mean error is 8.5 and the standard deviation of the monthly errors is 33. This standard error is about 8 percent of the 1977 average inventory balance of 1079. The standard deviation of the monthly inventory level (from its annual mean value) is 43, so that using the autoregressive model (I_t = I_{t-1})
does explain about 40 percent of the monthly fluctuations. This performance is better than nothing but not sensational.

For the accounts receivable category, we have a more formal model relating payments to current and lagged sales. Unfortunately, the predictions from this model during 1977 were not very good, as an examination of exhibits 1a - 1d reveals. The average error is 12.4, and the standard deviation of monthly errors is 190. A simple auto-regressive model \( AR_t = AR_{t-1} \) works better in this time period yielding predictions with a higher mean error of 21 but a standard deviation of 148.

For prepaid expenses we extrapolate an increased trend identified in the 1974-76 time period. Unfortunately this trend reverses with a 131 drop in February 1977 and a continued decline of an additional 98 during the remainder of 1977. Therefore, the model predictions for this account are poor. In re-examining the 1974-76 estimation period, the linear time trend gave only a slightly better fit \( (R^2 = .75) \) to the data than the auto-regressive model \( (R^2 = .74) \). Given the 1977 results, the auto-regressive model again appears to provide a better fit to the data. Assuming that prepaid expenses in month \( t \) equals prepaid expenses in month \( t-1 \) gives an average error of -16 (reflecting the decline in this account balance) and a standard deviation of error of 72 (or slightly less than 10 percent of average balance).

Thus, for each of the four current asset accounts (cash, receivables, prepaid expenses, and inventory), we are unable to explain monthly changes in account balances. Our best and most robust model is to predict the account balance in a month as the actual balance in the previous month. Additional work will be performed to see if we can overturn this sad conclusion, particularly for the accounts receivable category where we should be able to model payments as a function of sales, but we must probably conclude that for this company, current asset balances fluctuate randomly from month to month. A similar model is currently being developed for another company. This will give us a chance to see whether we will have better luck in being able to predict monthly balances in these current asset accounts.

The accounts payable predictions are similarly disappointing. The mean monthly error in this account is 23. The standard deviation of the twelve monthly errors is 117 which is more than 12 percent of the average monthly balance of this account. This estimation was done under the most favorable circumstances using actual purchases in each month as computed by:

\[
PUR_t = f_t - f_{t-1} + CGS_t.
\]
When we attempted to estimate purchases, from sales, by:

\[ \text{PUR}_t = 0.23 \times S_t \]

the mean monthly error jumped to 57 and the standard deviation of error to 131.

Using the auto-regressive model \( (AP_t = AP_{t-1}) \) gave a mean error of 5.8 and a standard deviation of error of 124. Thus, the estimated model has less fluctuation about its mean error than the auto-regressive model but the estimated model has a much larger mean error. Overall, the structural model is slightly superior with a root mean square error of 114 while the auto-regressive model has a root mean square error of 119.*

The pattern that a structural model, estimated using regression analysis on historical data, has lower predictive ability than a simple auto-regressive model highlights one of the limitations of regression analysis in an analytic review or indeed for any forecasting exercise. It is common for the explanatory power of a regression model to be much lower in a holdout (or future) period than would be expected based on the explanatory power \( (\bar{R}^2) \) in the estimation period. There could be a number of reasons for this phenomenon, including model mis-specification. But the \( \bar{R}^2 \) in a prediction period is bound to be less than that obtained from the estimation sample. It is all too easy to lose sight of the very important piece of information used in the regression analysis on historical data that is simply not available when we use a model to predict future values of a dependent variable. The \( \bar{R}^2 \) of a regression indicates the percentage of variation, as measured from the mean value of the dependent variable, that is explained by variation in the independent variables. Working with historical data, we assume that if there is no explanatory power from the regression, we cannot explain past values of the dependent variable better than just using the historical mean value. When attempting to predict future values of the dependent variable, however, we obviously do not even know the mean of these future observations. Thus, after the fact, we might have done better using the mean value of the dependent variable (had we known it) than to have used our historically based regression model. If the mean of future values of the dependent variable differs significantly from the mean of the predicted values, the regression model is bound to predict much less well than we would have expected or hoped for. The auto-regressive model

---

*The root mean square (RMS) error of residuals is

\[ \text{RMS} = \left[ \frac{\sum e_i^2}{12} \right]^{1/2}, \]
does a much better job of at least getting the mean of the predicted values close to the mean of the actual values. Its mean error is the difference between beginning and ending values of the variable divided by the number of observations in the prediction period (12 in this case). Unfortunately, the simple auto-regressive model gives us no insight into the causes of monthly differences in account balances. One should not generalize too readily from small samples and even limited model building on the small sample. We will need to examine the problem of estimating and predicting current asset and liability accounts on many more companies before we concede our inability to construct structural models with good predictive ability on these accounts.

DIVISIONAL DATA
All the analysis described so far has used the consolidated financial statements of the company. The company is multi-divisional and prepares its monthly financial statements for eleven separate divisions. After intra-company eliminations, these divisions are consolidated into the corporation statements that provided the basis for the analysis described in the previous sections. In principle, there should be gains from estimating separate models for each division and aggregating these into overall predictions for the company. In practice, this procedure may not be that beneficial. Estimating separate models for each division could lead to increased sampling error on the estimated coefficients for each division so that the overall predictions could have higher variability. A particular problem, for this company, is the amount of intra-company sales (fluctuating from 4 to 10 percent of consolidated sales) which requires eliminations from sales and associated expense accounts before the consolidated figures can be obtained. Even larger eliminations are required for the balance sheet accounts.

The benefits from estimating separate divisional models will be greatest when there is significant variation among the different divisions in the estimated relationships such as cost of sales to sales and depreciation to gross plant and equipment. In such cases, as the proportion of a division's sales or fixed assets to total company sales or assets changes over time, the consolidated model will appear unstable. In fact, however, the relationships could be stable at the divisional level. The shift in coefficients noted at the consolidated level would be due to the changed weightings of the divisional coefficients because of the shift in relative importance of the divisions
to the company. A second benefit from estimating divisional models occurs if the company makes a major acquisition or divestiture. If only the consolidated model had been estimated, a major restructuring of a firm’s assets and equities would make the entire model invalid. With divisional models available, it would still be feasible to maintain the models for those divisions that were still intact and develop new models for the newly acquired or reorganized divisions.

Because of these potential benefits, we estimated separate expense models at the divisional level. This would enable us to obtain predictions of the consolidated pretax income statement by aggregating across the separate divisional models. The first problem was to choose a reasonable level of aggregation. One division accounted for 57 percent of consolidated sales in 1976 and clearly represented a major entity. There were four other divisions in roughly the same line of business, none having more than 6 percent of 1976 consolidated sales. These divisions were aggregated together and the eliminations were somewhat, but not completely, arbitrarily charged against this group of divisions since much of the intra-company activity was associated with a division included in this group. After subtracting the eliminations, this grouping represented 10 percent of 1976 sales. Another group of three divisions, with 23 percent of sales, operated in the consumer rather than industrial market and was aggregated together. Finally, three miscellaneous divisions, with collective sales of 10 percent of consolidated sales, were aggregated into a fourth group.

Clearly, the most important relationship was the cost of sales to sales equation. Exhibit 2 presents the regression results based on 1974-1976 data. As was true for the various subperiods examined for the consolidated cost of sales relationship, there is more stability

<table>
<thead>
<tr>
<th>Group</th>
<th>Constant</th>
<th>Slope</th>
<th>$R^2$</th>
<th>$\bar{y}/\bar{x}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4100</td>
<td>.775</td>
<td>.88</td>
<td>.87</td>
</tr>
<tr>
<td></td>
<td>(1.93)</td>
<td>(16.4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>-1000</td>
<td>.950</td>
<td>.84</td>
<td>.82</td>
</tr>
<tr>
<td></td>
<td>(1.81)</td>
<td>(13.4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>2200</td>
<td>.64</td>
<td>.98</td>
<td>.78</td>
</tr>
<tr>
<td></td>
<td>(8.27)</td>
<td>(40.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>400</td>
<td>.74</td>
<td>.76</td>
<td>.79</td>
</tr>
<tr>
<td></td>
<td>(.72)</td>
<td>(10.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consolidated</td>
<td>4900</td>
<td>.77</td>
<td>.88</td>
<td>.84</td>
</tr>
<tr>
<td></td>
<td>(1.36)</td>
<td>(16.1)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
across divisions in the average of cost of sales to sales ratio ($\bar{y}/\bar{x}$) than in the variable cost coefficient. Thus, when we assume a strictly proportional relationship, as we did at the consolidated level, there is not much difference in the divisional ratios from the consolidated ratio. Only in group 3 is there an extremely strong case for including a fixed component in cost of sales.

For depreciation expense, we obtained much more variation in the variable coefficient multiplying gross plant and equipment as shown in exhibit 3.

<table>
<thead>
<tr>
<th>Group</th>
<th>Constant</th>
<th>Slope *</th>
<th>$R^2$</th>
<th>$(\bar{y}/\bar{x})$ *</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-100</td>
<td>3.04</td>
<td>.54</td>
<td>2.75</td>
</tr>
<tr>
<td></td>
<td>(.62)</td>
<td>(.49)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>20</td>
<td>2.74</td>
<td>.27</td>
<td>3.51</td>
</tr>
<tr>
<td></td>
<td>(1.04)</td>
<td>(3.75)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>-10</td>
<td>4.72</td>
<td>.35</td>
<td>4.58</td>
</tr>
<tr>
<td></td>
<td>(0.13)</td>
<td>(4.46)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>-100</td>
<td>6.85</td>
<td>.42</td>
<td>4.61</td>
</tr>
<tr>
<td></td>
<td>(1.68)</td>
<td>(5.16)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consolidated</td>
<td>-270</td>
<td>3.77</td>
<td>.62</td>
<td>3.25</td>
</tr>
<tr>
<td></td>
<td>(1.09)</td>
<td>(7.62)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Both the variable cost coefficient and the ratio of depreciation expense to plant and equipment show considerable variation among the four groupings of divisions. As in the consolidated case, the fixed component of depreciation is, as expected, not important to consider. Therefore, prediction will be done using the average ratio ($\bar{y}/\bar{x}$). Even though the consolidated regression has a higher $R^2$ than any of the individual group regressions, this does not imply that the consolidated regression gives a better fit to the data than the aggregate of the individual regressions. A higher $R^2$ for the consolidated regression could be an artifact of the aggregation process (see Grunfeld and Griliches, 1960). More extensive tests of the residuals would need to be performed before concluding that the consolidated regression dominates the aggregation of the individual group regressions.

The regressions for commercial, administrative and research expenses, not presented here, proved less satisfactory at the group level than at the consolidated level. The coefficient on sales was occasionally negative and rarely significant. The coefficient on time trend varied greatly among the groups and was also negative occasionally.
Apparently there may be difficulty or at least variability over time in allocating these costs to the divisions.

Individual models for interest expense at the divisional level were not constructed. There seems little reason to allocate interest expense to divisions unless the borrowing rate on bank debt varies across divisions. We had no evidence that this was the situation for the company being modeled.

Income statement predictions for the first six months of 1977 were generated by aggregating the expense forecasts generated by the four individual groups. The results are summarized on a quarterly basis in exhibit 4.

Of the three expense accounts modeled on a divisional basis, improved accuracy was obtained only for depreciation expense. Apparently, the different groups have assets of differing average lifetimes so that tracking plant and equipment on a divisional basis and computing separate depreciation rates for each major group gives more accurate predictions than assuming a single depreciation rate for the entire firm. For the most critical account, cost of sales, the consolidated forecast was more accurate than the aggregation of individual forecasts. While the results are shown only on a quarterly basis, the consolidated forecast tended to have lower errors on the monthly forecasts too. Thus, for this company, where the ratio of cost of sales to sales is relatively constant across divisions, the accuracy of the overall predictions is not improved by building separate divisional models. Obviously, this finding is a function of the particular firm that was studied and is not expected to be a general result across many different firms.

**SUMMARY**

We have developed a financial model for a firm's income statement that can project monthly, quarterly and annual pretax income numbers once the actual monthly sales, plant and equipment, and interest-bearing liability accounts are known. The accuracy of the profit-before-tax income number is critically dependent on having an excellent predictor of cost of sales which averages 84 percent of sales. A small error in estimating cost of sales produces a large error in the profit-before-tax number. Monthly errors in the profit-before-tax number were as high as 3 percent of sales (which occasionally caused a profit to be projected when a loss actually occurred), but the highest quarterly error was 1.75 percent of sales. Annual pretax income was
### Exhibit 4. Comparison of Actual, Consolidated Forecast, and Aggregated Individual Forecasts; Quarters 3 and 4 1977

<table>
<thead>
<tr>
<th></th>
<th>Quarter 3 Actual</th>
<th>Quarter 3 Consolidated Forecast</th>
<th>Quarter 3 Aggregated Individual</th>
<th>Quarter 4 Actual</th>
<th>Quarter 4 Consolidated Forecast</th>
<th>Quarter 4 Aggregated Individual</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sales</td>
<td>2592 (100)</td>
<td>2582 (100)</td>
<td>2582 (100)</td>
<td>2595 (100)</td>
<td>2595 (100)</td>
<td>2595 (100)</td>
</tr>
<tr>
<td>Cost of sales</td>
<td>2174 (84.2)</td>
<td>2169 (84)</td>
<td>2153 (83.4)</td>
<td>2217 (85.4)</td>
<td>2180 (84)</td>
<td>2161 (83.3)</td>
</tr>
<tr>
<td>Depreciation</td>
<td>55 (2.1)</td>
<td>52 (2.0)</td>
<td>54 (2.1)</td>
<td>55 (2.1)</td>
<td>51 (2.0)</td>
<td>53 (2.0)</td>
</tr>
<tr>
<td>Gross profit</td>
<td>353 (13.7)</td>
<td>361 (14.0)</td>
<td>375 (14.5)</td>
<td>323 (12.5)</td>
<td>364 (14.0)</td>
<td>381 (14.7)</td>
</tr>
<tr>
<td>S A &amp; R expense</td>
<td>221 (8.6)</td>
<td>226 (8.8)</td>
<td>216 (8.4)</td>
<td>226 (8.7)</td>
<td>220 (8.5)</td>
<td>210 (8.1)</td>
</tr>
<tr>
<td>Operating profit</td>
<td>132 (5.1)</td>
<td>135 (5.2)</td>
<td>159 (6.9)</td>
<td>97 (3.7)</td>
<td>144 (5.6)</td>
<td>170 (6.6)</td>
</tr>
</tbody>
</table>
projected to within 10 percent of actual pretax income which represented an error of 0.3 percent of annual sales. Regression analysis produced equations for expense accounts that gave a good fit to historic data and enabled these accurate predictions on the holdout sample of 1977 data. The interest expense account was modeled directly, as a function of actual interest rates and composition of interest bearing liabilities, rather than estimated using regression analysis.

It proved more difficult to obtain good predictive models for the current asset and accounts payable accounts. Attempts were made to model directly the flow accounts of collections, purchases, and payments and relate these to the relevant balance sheet accounts. When predicting on 1977 data, however, it was hard to improve on the forecast accuracy of a simple auto-regressive model which predicts a month's account balance as the actual balance in the preceding month. Such an auto-regressive model, though, cannot explain changes in account balances in successive months in terms of variations in the firm's production or sales activities. Thus, we must admit to not being able to explain or predict variations in the firm's balance sheet accounts.

In principle, predictions should be improved by estimating separate divisional models and aggregating these into an overall forecast. In practice, except for the depreciation expense account, the accuracy of the expense account predictions decreased when we aggregated across divisional forecasts.

In practice, except for the depreciation expense account, the accuracy of the expense account predictions decreased when we aggregated across divisional forecasts.

How useful the constructed financial model can be to auditors is a question that must ultimately be answered by practicing auditors. The model was developed without use of advanced regression techniques but did require thought, judgment, and understanding of the strengths and limitations of statistical analysis (at least I would like to think it did). The accuracy of the expense account forecasts is gratifying, but whether it is good enough is a question that must be evaluated by further research and by the judgments of practicing auditors. Naturally, this exercise must be repeated on many more companies before we get a feeling for the general applicability of bringing together the forecasts of individual income and expense accounts into an overall forecasted income statement. The effort reported here should be viewed as an initial feasibility study and within that con-
text, I would judge it a success that warrants further investigation on a broader sample of companies.

REFERENCES


Simultaneous Equations Considerations
At the Illinois Conference, William Kinney raised the question of why I had not used simultaneous equation techniques for estimating the model. While I was generally aware that there were some interrelationships among the variables in the model, I had not previously thought that this was a crucial issue to consider. Bill's remark caused me to think more systematically about the interrelationships, and I now believe they should be explicitly recognized when developing and estimating a financial model. In retrospect, I am embarrassed that I had not developed the simultaneous equation system, but I am grateful for Kinney's comment that triggered my interest in doing the analysis. I also appreciate having the opportunity to add this short section to my paper. Professor Tim McGuire of CMU provided valuable counsel to me on the econometric issues.

The accounts we are interested in estimating are:

- Cost of Goods Sold \( (\text{Month } t) \) \( CGS_t \)
- Accounts Receivable \( (\text{Month } t) \) \( AR_t \)
- Inventory \( (\text{Month } t) \) \( I_t \)
- Accounts Payable \( (\text{Month } t) \) \( AP_t \)

This note, which treats part of the financial model as a simultaneous equation system, was stimulated by a remark by Professor William Kinney at the Illinois Symposium on Auditing Research (October 26, 1978). Professor Tim McGuire provided valuable assistance in clarifying the relevant econometric issues.
Explanatory or intermediate variables in the equations will be:

\[
\begin{align*}
\text{Sales} & \quad \text{(Month } t \text{)} \quad S_t \\
\text{Collections} & \quad \text{(Month } t \text{)} \quad \text{COL}_t \\
\text{Purchases} & \quad \text{(Month } t \text{)} \quad \text{PUR}_t \\
\text{Payments} & \quad \text{(Month } t \text{)} \quad \text{PAY}_t
\end{align*}
\]

There are four equations to be developed. For simplicity, the stochastic error term will be suppressed in the structural equations though, as we will see, the error structure is crucial when selecting an estimating procedure.

1. **Cost of Goods Sold**
   Using the strictly proportional relationship found for this firm,
   \[
   \text{CGS}_t = \alpha_1 S_t.
   \]
   (1)

2. **Accounts Receivable**
   We first estimate collections as a function of past sales:
   \[
   \text{COL}_t = \beta_1 S_t + \beta_2 S_{t-1} + (1 - \beta_1 - \beta_2) S_{t-2}.
   \]
   (2)
   Substituting this relationship into the accounting identity
   \[
   \text{AR}_t = \text{AR}_{t-1} + S_t - \text{COL}_t
   \]
   yields the equation:
   \[
   \text{AR}_t = \text{AR}_{t-1} + (1-\beta_1) S_t - \beta_2 S_{t-1} - (1-\beta_1-\beta_2) S_{t-2}.
   \]
   (4)

3. **Inventory**
   The empirical relationship estimates purchases as a function of sales:
   \[
   \text{PUR}_t = \gamma_1 S_t.
   \]
   (5)
   Substituting into the accounting identity,
   \[
   \text{I}_t = \text{I}_{t-1} + \text{PUR}_t - \text{CGS}_t,
   \]
   yields:
   \[
   \text{I}_t = \text{I}_{t-1} + \gamma_1 S_t - \text{CGS}_t.
   \]
   (7)

4. **Accounts Payable**
   We estimate payments as a function of last period’s accounts payable and this period’s purchases:
   \[
   \text{PAY}_t = \delta_0 + \delta_1 \text{AP}_{t-1} + \delta_2 \text{PUR}_t.
   \]
   (8)
   Using the accounting identity (6) we can rewrite equation (8) as
   \[
   \text{PAY}_t = \delta_0 + \delta_1 \text{AP}_{t-1} + \delta_2 (\text{I}_t - \text{I}_{t-1} + \text{CGS}_t).
   \]
   (9)
   With the third accounting identity,
   \[
   \text{AP}_t = \text{AP}_{t-1} + \text{PUR}_t - \text{PAY}_t,
   \]
   (10)
we can obtain the fourth structural equation:
\[ AP_t = \delta_0 + (1-\delta_1) AP_{t-1} + (1-\delta_2) (I_t - I_{t-1} + CGS_t). \] (11)

**Simultaneous System**
We collect equations (1), (4), (7) and (11) and write them, in matrix notation as:

\[
\begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
1 & 0 & 1 & 0 \\
-(1-\delta_2) & 0 & -(1-\delta_2) & 1
\end{pmatrix}
\begin{pmatrix}
CGS_t \\
AR_t \\
I_t \\
AP_t
\end{pmatrix} =
\begin{pmatrix}
\alpha_1 S_t \\
(1-\delta_1) S_t - \delta_2 S_{t-1} - (1-\delta_1-\delta_2) S_{t-2} + AR_{t-1} \\
\gamma_1 S_t + I_{t-1} \\
-\delta_0 + (1-\delta_1) AP_{t-1} - (1-\delta_2) I_{t-1}
\end{pmatrix}
\] (12)

The accounts receivable equation (4) is seen to be independent of the other three equations and can be estimated separately. Because the lagged sales variables appear as independent variables in equation (4), there could be correlated error terms in this equation. More efficient estimation could be performed by exploiting any first or second order auto-correlations in the error structure. This possibility will be investigated in subsequent work.

The other three equations form a recursive system since the structure is lower block triangular. If the error terms for these three equations are independent of each other (more formally, the covariance matrix must be diagonal), then the system is strongly recursive and ordinary least squares on each equation is an efficient estimation procedure. Thus, one can view the estimated equations presented in the first part of this paper as first and perhaps reasonable approximations to the simultaneous system which assume uncorrelated errors across the three equations. Subsequent research using two stage and three stage least squares will be done to test this assumption.

As a final note, it is easy to invert the three structural equation system:

\[
\begin{pmatrix}
1 & 0 & 0 \\
1 & 1 & 0 \\
-(1-\delta_2) & -(1-\delta_2) & 1
\end{pmatrix}
\begin{pmatrix}
CGS_t \\
I_t \\
AP_t
\end{pmatrix} =
\begin{pmatrix}
RHS
\end{pmatrix}
\] (13)
To obtain the reduced form equations:

\[
\begin{bmatrix}
CGS_t \\
L_t \\
AP_t
\end{bmatrix} =
\begin{bmatrix}
1 & 0 & 0 \\
-1 & 1 & 0 \\
0 & (1-\delta_2) & 1
\end{bmatrix}
\begin{bmatrix}
\text{RHS}
\end{bmatrix}
\]

which would be estimated using simultaneous equation techniques.
Audit work is analytical, whether it involves system flowcharting, sampling, or a host of other activities that comprise the audit process. Recent interest has focused on "analytical review" procedures, as the AICPA's Statement on Auditing Standards Number 23 (AICPA, 1978) demonstrates. Analytical review is viewed as that aspect of the audit process involving the estimation and interpretation of the relationships between variables. The variables may be of a financial or non-financial nature. They may be created within the enterprise, or they may arise in its environment. They may describe processes that occur over time, or they may describe the state of things at a point in time.

Fortunately, there are many tools at the auditor's disposal which permit him to study relationships ranging from simple ratios of two variables to elaborate expressions involving many variables. The possibilities for the construction of relationships are seemingly endless. Despite the obvious appeal of analytical review procedures, two key issues confront the auditor who desires to employ them as auditing tools.

First, the auditor should be sufficiently skilled in identifying relevant variables and in specifying expressions that describe their behavior. A skilled auditor will be parsimonious in his choice of relationships; he will use the simplest relationships possible, capturing the most information for the least effort. Kaplan does demonstrate the relative ease with which this issue can be resolved. Using examples,
he derives a variety of expressions from considerations of the time-series properties of accounting numbers, from the identities that arise in the double-entry conventions, and from existing vintaging and other causal relationships. None of the expressions he develops is so complicated as to be beyond the comprehension of an auditor who has a good understanding of algebra and the basics of regression analysis.

Nevertheless, Kaplan does appear to miss a few relationships that would interest the auditor. He makes no attempt to model relationships affecting key accounting estimates, notably the allowance for doubtful accounts. Whether or not bad debt expense is negligible, as he asserts, is open to question as a testable audit hypothesis. Consequently, some attempt should have been made to model the process governing write-offs, bad debts, and the allowance. Similarly, Kaplan makes no attempt to formulate the behavior of sales returns, discounts and other allowances, nor does he consider that inventory obsolescence is an audit concern worth modeling.

Moreover, Kaplan uses sales as his key input variable. While this may be a suitable choice for financial planning purposes, cash flow variables would be better choices for audit applications. Both cash receipts and cash disbursements data describe the most liquid assets of an enterprise. Because internal controls over cash flows are likely to be the most reliable that an auditor is likely to encounter, receipts and disbursements data lend themselves to precise measurement (which is not always the case with sales data). Furthermore, the development of models that use cash basis variables to explain accrual basis variables could prove to be the most fruitful approach to using quantitative methods as audit tools, particularly if Kaplan were to restrict the scope of his models to consideration of specific accounting cycles, rather than the entire financial accounting process. Thus, cash receipts would be the key explanatory variable for the revenue cycle, whereas cash disbursements could perform a similar role for the expenditure cycle. Formulation of the cycles in the manner suggested by Kaplan could also provide a suitable framework for the audit test process and would facilitate the auditor's choice of testing methodologies, whether they be tests of details or analytical review procedures.

Kaplan himself raises the second issue, which concerns the proper choice of decision strategies, as a problem that remains to be solved. Having developed the models, he prepares pro forma financial statements which are then compared with the reported results for the time
period under audit. What then? How is the auditor to interpret the inevitable differences between the two sets of numbers? Can the absence of significant deviations provide the auditor with any assurance other than cold comfort? Does the presence of significant deviations suggest the existence of misstatement, or could the models be insufficiently specified? What is the auditor to do in light of the information obtained from the analysis? As Kaplan observes, Stringer (1975) and Kinney (1977) provided some suggestions for integrating analytical review with other audit procedures. Kinney, however, introduced an approach to calculating the risk of committing an error of the second kind (beta risk) that is probably unique in the literature. Beta risk has never been a feature of the analysis of regression, and it remains to be seen if his formulation is effective in practice. Moreover, as Albrecht and McKeown (1977) observed, the auditor must take care that the audit use of these modeling techniques does not lead him along a circular path, especially if the variables have been “cleaned” in prior applications of analytical review procedures. Consider the possibility (perhaps extreme) that auditors will use these tools to assess the reasonableness of accounting data, and financial analysts will later employ the same devices to extract economic meaning from the same data. How happy the analyst will be to see the data conforming to expectation. If the tools of the financial analyst or planner are to be used by the auditor to serve a purpose for which they were not designed (to assess the accuracy of data, rather than to extract meaning from the data whose accuracy has been assessed), then considerable additional research is required. Otherwise, auditors may well find themselves “hoist by their own petard.”

REFERENCES

Discussant's Response to "Developing a Financial Planning Model for an Analytical Review"

RICHARD B. LEA

The perspective from which I review this paper is influenced by the following factors. First, my public accounting experience at Peat, Marwick, Mitchell & Co. was mainly in auditing research rather than in practice. Consequently, my remarks will focus more on conceptual issues than on implementation issues. Secondly, my auditing research experience included one project involving analytical review. In this project, we made some limited attempts at modeling certain aspects of selected clients' financial statements. Our modeling efforts were not very successful. Some of the problems we experienced were the following:

- Difficulties in getting clients to commit people for data gathering and model development;
- Difficulties in getting necessary data—for example, obtaining three years of monthly data for selected variables was often a significant problem;
- Difficulties in assessing the accuracy of the monthly data which were obtained.

I simply mention these difficulties because they have affected my expectations regarding what one can expect from the case study type of research that Kaplan is reporting. In effect, this type of research is difficult to perform not only because of data interpretation problems, but also because of constraints imposed by the client with respect to data access and availability of people.
The third and most important factor which influences my perspective is the fact that I was one of the reviewers of Kaplan's proposal underlying this research that was submitted to Peat, Marwick, Mitchell & Co. in conjunction with its program Research Opportunities in Auditing. I was very enthusiastic about his proposal at the time I reviewed it, and, thus I am happy to have the opportunity to review and discuss his research findings.

Given this perspective, let me turn to specific review comments which I will discuss under three major headings: basic modeling strategy, specifics of the various models, and interpretation of model output.

**BASIC MODELING STRATEGIES**

Although we do not have available in the literature an optimal solution for building a model, it seems to me that over the last twenty years or so, model builders (e.g., Morris, 1967) have developed some "generally accepted principles of modeling." Very briefly, these principles appear to be the following:

- Begin with a very simple model. If sufficient accuracy cannot be achieved, gradually refine and further develop the model.
- Begin with highly aggregated data and gradually disaggregate in areas which significantly affect model outputs.
- Begin with a very limited number of explanatory variables. Attempt to find one or a few explanatory variables that account for the major changes in the dependent variable of interest.
- Begin by postulating simple relationships among the independent variables.

The modeling strategy that I see reflected in Kaplan's paper would appear to follow these modeling principles very closely. Moreover, the principles themselves seem to make a great deal of sense. Essentially, the principles imply an evolutionary approach to model building which allows the model builder to periodically compare marginal improvements in accuracy with incremental costs of a more refined model.

Ideally, the accuracy requirements themselves would be derived from the decision model being fed by the prediction models. But, in analytical review, at least, we appear to be a long way from linking the accuracy of outputs of analytical review prediction models to the auditor's decision models.

In any event, I would conclude that Kaplan's overall modeling
strategy is basically sound. I therefore turn to some specific aspects of the model which he has developed.

**SPECIFICS OF THE VARIOUS MODELS**

A major aspect of several of Kaplan's models is the use of recorded sales as a key independent variable.

**Use of Recorded Sales**

Although the use of recorded sales has some advantages, these would appear to be more than offset by several disadvantages.

On the plus side, use of recorded sales obviously simplifies the initial model-building effort and therefore is consistent with the modeling principles outlined above. Also, in the particular case study presented by Kaplan, use of recorded sales results in accuracy of predictions in the income statement that were close to Kaplan's specified target of errors of 5 percent or less.

On the negative side, however, use of recorded sales may introduce an awfully big "IF" in the auditor's interpretation and use of analytical review results. In effect, all significant deviations between the client's recorded amounts for cost of goods sold, etc., and the auditor's predicted amounts derived from models that use recorded sales are conditional on the accuracy of the client's recorded sales figures. And, it is doubtful whether such "significant deviations" could be used in audit planning and audit resource allocation decisions unless or until this condition is removed. In many cases, of course, the auditor may have considerable confidence in the accuracy of the client's recorded sales figures, and therefore this conditional aspect would not be important. I suspect, however, that in a very large number of cases, the recorded sales figures, particularly monthly sales figures, will be known to be unreliable. Hence, the auditor will generally need predictions of sales that represent something more than just recorded figures. In effect, the auditor will need a model of the sales generating process itself.

It seems to me that modeling of a client's sales generating process has additional significant advantages. First, the sales generating process is typically a key aspect of a client's business, and modeling this process should therefore lead the auditor to a much better understanding of the business. I recognize that such a benefit is very intangible, yet it is nevertheless certainly real. Secondly, sales models typically include external economic factors as explanatory variables. In these
cases, the sales model provides predictions that are less susceptible to management manipulation. In effect, such models provide linkages with the external environment providing an external validity check on client figures which may appear to be otherwise internally consistent.

Finally, the sales amount typically involves a very large number of transactions. Accordingly, sales may require considerable detailed testing in order for the auditor to obtain reasonable assurance that the amount is not materially in error. Thus, sales represents a primary area in which analytical review results might be used as a basis for reducing detailed tests.

For these reasons, development of models for sales would appear to be a fruitful area for analytical review research. Incidentally, research has recently been reported regarding the modeling of sales for the purpose of preparing monthly financial statement forecasts. Elliott and Uphoff (1972) modeled a monthly income statement which was based on a sales forecast model. The sales model incorporated external variables such as population growth, market share, and an industrial production index.

The Interest Expense Model
If we viewed Kaplan’s model building process as a “statistical” process, we would certainly have to identify his interest expense model as an “outlier.” This model is unusual because it includes a relatively large number of explanatory variables—12 to be exact. It is not clear whether this degree of detail is commensurate with the relative importance of interest expense in the income statement—interest expense is approximately 2 percent of sales—particularly in light of the fact that miscellaneous income and expense—also 2 percent of sales—was not modeled because of its erratic nature.

In looking at the problem of modeling interest expense for analytical review purposes, I would question whether further developmental efforts in this area are a high priority for large industrial corporations. For these organizations, interest-bearing debt liability accounts generally involve relatively few transactions, yet each transaction may often be highly material. Hence, it is common—and seems to be cost effective—to verify each liability account transaction and then to relate interest expense on an item-by-item basis to the corresponding interest bearing debt.
The Cash Model

With respect to a large industrial corporation, it is not surprising to me that a simple model of cash does not perform very well in predicting the behavior of the monthly cash balance. For such companies, the cash balance typically reflects a relatively small difference between enormously large inflows and outflows. Moreover, the inflows and outflows reflect highly heterogeneous types of transactions. Indeed, the cash account may be reviewed as a key element of all the various transaction cycles found in a business.

In effect, we should expect a rather complex and costly model will need to be developed for predicting cash balances. And, here again, I doubt whether such a developmental effort is of high priority for analytical review. Traditional audit methods for cash, which involve confirmations, cutoff statements, counts, and reconciliations, would appear to be more cost effective compared with analytical review models.

The Accounts Receivable Model

Although the accounts receivable model developed by Kaplan did not perform very well, I believe that model development is moving in the right direction. As several writers (e.g., Stone, 1976; Lewellen and Edminster, 1973) recently have pointed out, traditional one-parameter forecasting models of accounts receivable, such as those involving accounts receivable to sales ratio, average days outstanding, or cash flow to accounts receivable ratio, generally perform poorly because they fail to deal separately with sales patterns and customer payment patterns. Models using aging percentages also suffer from this defect.

Earlier, I discussed the need for modeling sales patterns. With respect to customer payment patterns, improved forecasting may depend on introducing additional explanatory variables, such as the following:

- Differential credit terms by major product lines;
- Macro-economic variables related to customers' ability to pay
- Changes in a client's credit policies; and
- Changes in a client's collection policies.

This brief list of variables suggests that considerable disaggregation may be required in order to achieve desired forecasting accuracy. Indeed, one researcher (Stone, 1976) suggests that the logical focus for accounts receivable forecasting is at the product-line, credit-term level.
The Inventory Model

Kaplan explicitly recognizes the possibility of two different behavioral models underlying changes in inventory in relation to sales. He also recognizes the consequent difficulties involved in capturing both these behavioral models in one analytical review model. Because of those difficulties, Kaplan opted to leave it to the auditor to judge which behavioral model applies. On this point, my question to Kaplan is simply this: How will the auditor make this judgment? My own guess would be that the auditor would examine industry and macro-economic data in order to assess the direction in which sales were currently moving and expected to move in the near future. But this type of analysis would appear to be an essential part of analytical review, specifically that part of analytical review devoted to modeling and testing the reasonableness of sales. Accordingly, I would expect to see progress in inventory modeling tied closely to progress achieved in modeling a company’s sales generating process.

The Accounts Payable Model

The accounts payable model illustrates a tradeoff that seems to be particularly significant in auditing. The tradeoff involves the question of whether the auditor should use independent variables (e.g., recorded purchases) that lead to greater precision but that may be manipulated by management or use independent variables (e.g., predicted purchases derived, say, from predicted sales) that provide less precision but are impervious to management manipulation. As a general principle, the auditor should probably attempt to build analytical review models that, to the greatest extent possible, incorporate explanatory variables that cannot be easily manipulated by management.

Divisional Models

As Kaplan points out, in principle there should be gains in forecasting accuracy from estimating separate models for divisions and then aggregating. Indeed, this is the basic notion that underlies the growing demand for external reporting of segment information and the recent adoption by FASB of segment reporting requirements. Moreover, this notion has been empirically tested and positive results were obtained (Kinney, 1971; Collins, 1976). Thus, even though Kaplan’s results are somewhat negative, in light of other studies, it would appear that model development efforts should continue to be made at the divisional level.
INTERPRETATION OF MODEL OUTPUT

Two aspects of model output deserve comment.

Overall Test Reasonableness

Kaplan indicates that a primary objective in developing an overall financial statement model is to provide the auditor with an improved method in analytical review for arriving at a conclusion as to the overall reasonableness of a company's financial statements. Unfortunately, he does not develop this idea further in the paper. Some progress on this problem has been made recently, however, as evidenced in the recent paper by Albrecht (1976), in *The Accounting Review*.

Accuracy of Model Predictions

In general, one never really knows the "true" figures which forecast models are attempting to predict. Nevertheless, the choice of "benchmark" figures against which forecast results are to be evaluated is critical to the interpretation of forecast accuracy. Obviously, "significant deviations" could just as well reflect erroneous recorded amounts as forecast errors. Thus, some verification of the benchmark figures would be desirable; however, the case study is silent on this point.

SUMMARY

I believe that Kaplan's case study has generated numerous insights into problems confronting the auditor in modeling a client's financial statements. Moreover, I agree with Kaplan that reports of additional case studies are badly needed in order for the auditor to develop a better understanding of the strengths and limitations of the modeling approach.

REFERENCES


The Effect of Measurement Error on Regression Results in Analytical Review

WILLIAM R. KINNEY, JR. and GERALD L. SALAMON

Regression analysis has often been suggested as an analytical review technique and has been applied by some practicing CPAs. Several statistical and auditing problems may be associated with the use of regression in auditing including model misspecification, model misestimation, lack of sufficient precision in estimation and measurement error in the variables. In this paper, we consider the effect of errors in the independent and dependent variables on the application of regression in auditing. Specifically, we consider accounting errors in the dependent variable in the regression model estimation (base) period and random measurement error in the independent variable in the base and audit period.

The results should be of interest to practicing auditors since the presence of errors in past accounting data and the existence of measurement error in environmental data will influence the cost of implementing analytical review procedures. Furthermore, there is some evidence that data errors may have led to inappropriate reliance on analytical review in at least one audit. Accounting Series Release #241 [Securities and Exchange Commission, February 10, 1978] concerns Deloitte, Haskins & Sells and the firm of FISCO, Inc., a casualty insurer. To quote the release, "H & S relied extensively on an analytical test of the loss reserves to evaluate reserve adequacy. Although the methodology was basically sound, the results were inadequate because they were based on FISCO-supplied data which contained
numerous errors." While this case did not involve the use of the firm's STAR regression program, it does illustrate the practical problems associated with relying on potentially erroneous data in audit practice.

The analysis of random measurement error in the independent variable is of interest because its presence may so bias estimated regression coefficients that (1) the account is subjected to extensive audit tests when the account is not materially in error—type I error—or (2) the account is not extensively tested even though it is materially in error—type II error.

In the next section, we relate the use of regression as an analytical review tool to a simplified audit environment. The following two sections describe the procedure used to generate the various independent and dependent variable series analyzed and the specific operationalization of the regression approach to analytical review. The underlying correlation parameters for the simulated data are based on coefficients reported in the literature by practitioners using regression in auditing.

In the final section, we analyze results of the simulations. The results indicate that random measurement error in the independent variable does lead to increased type I and type II errors while uncorrected accounting errors in the base period leads to reduced type I errors but increased type II errors. The latter effect is particularly pronounced when there are several small uncorrected errors.

**ANALYTICAL REVIEW AND REGRESSION**

According to the AICPA's recent exposure draft on analytical review,

. . . A basic premise underlying the application of analytical review procedures is that known relationships may reasonably be expected by the auditor to exist and continue in the absence of known conditions to the contrary. The presence of those relationships provides the auditor with evidential matter required by the third standard of field work. [AICPA, 1978]

Thus, we see that the auditor must be able to determine relationships among data and extrapolate them into the audit period.

In an analytical review of account balances, the auditor compares the client's reported balance with the auditor's assessment of the likely true (audited) balance. This assessment or conditional expectation may incorporate the audited balances of the client for prior periods and structural or environmental data such as current and past data for the economy and industry, current unaudited client data from related accounts and "independent" internal records such as
production statistics. If client reported values are "close" to the
auditor's conditional expectation, the auditor's confidence in the
validity of the reported balance is increased. That is, the auditor
assesses a lower probability of the existence of a material accounting
error.

To formalize using regression as the basis for the analytical review,
consider a time series, \( y_1, y_2, \ldots, y_t, \ldots \) in which \( y_t \) is the audited
balance of an income statement account for month \( t \). The auditor
must express an opinion as to the fairness of the sum of twelve as yet
unaudited counterparts of \( y_t \) for the year under audit, say \( t = m + 1 \) to
\( m + 12 \). Assume that the auditor faces the following ideal situation: \( y_t \)
in the base or preaudit estimation period (say 1 to \( m \)) is linearly re-
late to an easily and inexpensively obtainable structural series \( x_t \) via:

\[
y_t = \alpha + \beta x_t + e_t \tag{1}\]

where \( \alpha \) and \( \beta \) are regression coefficients relating \( x_t \) to \( y_t \), and \( e_t \) is a
residual or unexplained part of \( y_t \) which is distributed NID \((0, \sigma^2)\).2

Under such conditions the auditor can estimate \( \alpha \) and \( \beta \) using
ordinary least squares regression. Assuming that the relationship
between \( x \) and \( y \) is expected to continue and \( x_t \) is available for \( t = m + 1 \)
through \( m + 12 \), the auditor can compute the conditional expectation
of \( \hat{y}_t \mid x_t \) (denoted \( \hat{y}_t \) ) via:

\[
\hat{y}_t = \hat{\alpha} + \hat{\beta} x_t. \tag{2}\]

By considering the related reported or book values of the audit period
for the account (denoted \( y_{rt} \)), the auditor can calculate the estimated
total accounting error (\( \hat{E} \)), the standard deviation of total accounting
error and the 1-\( A \) upper precision limit on total overstatement error via:

\[
\hat{E} = \sum_{t=m+1}^{m+12} (y_{rt} - \hat{y}_t), \tag{3a}\]

\[
\hat{\sigma}_E^2 = \hat{\sigma}^2 \sqrt{1 + \frac{144}{m + 144} (\overline{x} - \overline{E})^2 / \sum_{t=1}^{m} (x_t - \overline{x})^2} \tag{3b}\]

\[
\text{UPL}_{1-A}(E) = \hat{E} + t_{1-A,m-2} \cdot \hat{\sigma}_E, \tag{3c}\]

1The choice of time-series regression and an income statement account is made for
specificity and convenience. The regression approach can also be applied cross-
sectionally and time-series or cross-sectional regression can be applied to balance sheet
accounts.

2An accounting data model as simple as eq. 8 is probably rare. However, Stringer
[1975], Albrecht and McKeown [1976] and Kinney [1978] indicate that suitable trans-
formations of \( y_t \) and/or \( x_t \) usually allow the use of ordinary least squares estimation.
where \( \bar{x} \) denotes the audit period average and \( \bar{x} \) the base period average of the independent variables.\(^3\) If the 1-A upper precision limit calculated using \( 3c \) is less than a material amount (E*), the auditor will have evidence that accounting overstatement error is not material at the 1-A reliability level.

Even under these ideal conditions, we do not know the ability of the procedure to correctly identify accounting errors when they exist, nor do we know how often the analytical review results are negative (i.e., do not provide a basis for acceptance of book value) due to sampling problems associated with regression.

The ideal conditions outlined above will often not be met in practice, of course. Some of the potentially serious problems facing the auditor attempting to use regression in auditing are:

(a) \( y \) may be related to or caused by structural variables other than \( x \);
(b) the relationship between \( x \) and \( y \) may not be linear;
(c) the observed base period correlation between \( x \) and \( y \) may be spurious;
(d) the \( x,y \) relationship may change over time;
(e) the residual \( e \) may not be distributed NID \((0, \sigma^2)\);
(f) \( \sigma^2 \) may be too large to allow precise assessments of \( y \) and accounting error; and
(g) \( x \) and/or \( y \) may be observed or measured with error causing \( \alpha, \beta \) and/or \( \sigma^2 \) to be misestimated.

The nature of several of these problems is well known to econometricians and solutions or tests exist for some of them. For example, transformation of the data will often correct for violations related to \( b \) and \( e \), and \( d \) can be tested statistically. Economic theory as a basis for determining “logical relationships” can be used to reduce the risk of \( c \) and to indicate the need to consider expanding the set of potential structural variables as indicated in condition \( a \). Condition \( a \) is related to \( f \) in that the exclusion of relevant structural variables inflates the estimated \( \sigma^2 \) [see Johnston, p. 169]. Problem \( f \) may not be pervasive given continued use of regression in analytical review in audit practice. Thus problems \( a-f \) have well known and reasonably adequate remedies and do not present unique difficulties in audit applications.

Measurement error (condition \( g \), if random, is known to (1) inflate estimates of \( \sigma^2 \) (raising the UPL, achieved), when \( y \) is measured with error, and (2) bias the estimates of \( \beta \) downward (leading to nonrejec-

\(^3\)Since particular audit procedures for tests of details are usually effective on errors of only one direction, the auditor's interest in analytical review will often be in a one-sided confidence interval. We simplify our work by arbitrarily considering only overstatement errors.
tion of the hypothesis of $\beta = 0$), when $x$ is measured with error. Thus, either of these random measurement error problems may lead to extending audit tests even though no accounting error exists. Furthermore, due to sampling error, random measurement error may sufficiently influence estimated coefficients so that a client balance is inappropriately accepted. Thus, the potential practical importance of random measurement error in auditing and the lack of readily available and easily applied solutions to the problem caused us to consider the likely nature and extent of random measurement error in the independent variable in a simplified audit-like setting.

In addition to studying random measurement error in $x$, we will study base period accounting errors in $y$. We limit base period accounting errors to fixed amounts of overstatement. Study of other error patterns may also be informative but the constrained overstatement error case seems to be potentially the most dangerous for auditors.

The analysis of this paper is based on simulated time series with known parameters and correlation levels similar to those found in practice. The approach is analogous to that used by Neter and Loebbecke [AICPA, 1975] for statistical sampling. To determine the incidence of type I (a material accounting error signaled when none exists) and type II errors (no accounting error signaled with a material error exists), some replications of the simulated data will be “seeded” with several patterns of accounting error in the audit period.

**THE DATA**

The true relationship between $y_t$ and the particular structural series which actually affect $y_t$ are, of course, not known in practice and therefore one cannot separate errors arising from model specification, model estimation, random causes, or, indeed, accounting errors and irregularities. The approach taken in the current study is to use simulation techniques to develop data series with known parameters and known accounting errors and determine the ability of regression to correctly identify accounting errors.

The basic model which we use in the analysis is:

$$y_t = .33935 + .2x_t + e_t,$$

$$x_t \sim \text{NID}(4, .2606^2)$$

$$e_t \sim \text{NID}(0, .5^2).$$

---

*The McGill University Random Number program "SUPER-DUPER" was used to generate the $x_t$, $e_t$, $M_{Lt}$ and $M_{Ht}$ series.*
The model yields an expectation of $y_t$ of $8.3333$ and an audited value for the year of 12 • ($8.3333$) or $100$. The expected correlation between $x$ and $y$ is approximately .95 which corresponds to the 67th percentile in Deloitte, Haskins & Sells audit experience with regression [Stringer, 1975, p. 8].

Forty-eight observations of $x$, $e$ and $y$ were generated and the process repeated 200 times to provide 200 simulated audit years. The forty-eight months allows a thirty-six-month base period which has been found useful in audit practice [Stringer, 1975] and other research in the use of regression [Deakin and Granof, 1974; Kinney, 1978].

Eight additional series were generated to simulate the error portion of the variables and were added to the $x$ and $y$ series in various combinations. Since the same underlying $x$, $y$ and $e$ series were used throughout, the observed differences in results are attributable to the error seedings.

To set random measurement error in the $x_t$ series, we generated two series which correspond to "low" and "high" measurement error in $x$. The resulting "low measurement error" case ($x_{Lt}$) was $x_t + M_{Lt}$ where $M_{Lt} \sim \text{NID}(0, .2920^2)$ and the high error case ($x_{Ht}$) was $x_t + M_{Ht}$ where $M_{Ht} \sim \text{NID}(0, .3397^2)$. The $\sigma_M^2$ terms yield expected correlations with $y$ of approximately .90 and .85 respectively. These correlations correspond to the 50th and 33rd percentiles in Deloitte, Haskins and Sells experience. The extent of actual measurement error in the Deloitte, Haskins and Sells experience statistics is, of course, not known. However, the levels used here seem reasonable at least for exploratory purposes in that one reason for low correlation between the independent and dependent variables is measurement error in the independent variable.

To simulate the effect of uncorrected accounting overstatement errors in the base period, we created three additional dependent variable series by adding one of three sets of accounting error ($a_t$) to the base period $y_t$. A material accounting error ($E^*$) was somewhat arbitrarily set at $2 or 2$ percent of the expected audited balance of the account. The first additional base period series was created by concentrating a material error of $2 in a single base period month, the second series was created by distributing $.6667$ to each of three months, and the third series was created by distributing $.1667$ to each of twelve months. The months so seeded were chosen at random.

Finally, to determine the ability of the regression approach to give correct signals to the auditor, we examined four conditions of ac-
counting error in the audit period. The first was no accounting error, the second, one error of $2, the third, three errors of $.6667, and the fourth, twelve errors of $.1667. We now describe the nature of the analytical review procedure used in the analyses.

THE ANALYTICAL REVIEW PROCEDURE
The analytical review procedure used has five stages. After each stage, the analysis is either continued or terminated depending on the test condition at the end of the stage. In the review procedure and the analyses of results below, we do not consider what the auditor might do when the analytical review is terminated. For example, when the assumptions underlying regression are apparently well met but the upper precision limit exceeds $E^*$, the auditor might continue with an analytical investigation [AICPA, 1978] or extend the planned substantive tests of details in the areas indicated by the regression results. Since such further analysis necessarily involves many additional assumptions, we truncate our work at the “termination of analytical review” point. Thus, if the analytical review is negative for any reason and material accounting error is present in the audit period, we count it as a correct decision.

The analytical review procedure begins with ordinary least squares estimation using $y$ and the assumed $x$ structural series ($x_L$ or $x_H$). For simplicity we omit distributional tests for normality and serial independence of the residuals. In practice such tests are often not met for the original variables but appropriate transformations usually allow acceptance of the assumption of normality and serial independence of the transformed variables [Albrecht and McKeown, 1977; Kinney, 1978].

The first decision point involves testing whether $\beta = 0$ for the base period. If this condition is met, then the structural series is of no significant value in explaining or justifying the $y_t$ observed and thus does not provide a basis for reliance. Given a significant $\beta$, the second stage involves scanning the base period $y$ for the presence of “outliers.” Outliers are estimated residuals so large that the probability of their coming from the same population as the other observations is small. Their inclusion in the estimation period may reflect non-recurring conditions and bias the estimated coefficients under more normal conditions. Operationally, a base period outlier is defined as $|e_t| \geq 3.6 \hat{\sigma}_t$. Under a normal distribution the probability of observing a residual greater than or equal to $3.6 \sigma$ is about .001 and the probability
of observing one or more such residuals in thirty-six random "draws" from such a population is approximately 1 - .99986 or about .04.

Given a significant base period estimate of $\beta$ and no base period outliers, the Chow test [Chow, 1960] is conducted to determine whether the audit period book value observations ($y_{rt}$, for $t > 36$) seem to follow the same relation as the base period. We conduct the Chow test at the .05 level.

If these three tests are passed, we assume that the auditor concludes that the regression parameters are acceptable, and we proceed with an outlier scan for the audit period. The logic of this scan is that if exceptionally large residuals exist for the book values of the audit period, there is a good chance that the results may be due to the presence of accounting error. For this outlier search, $2.44\hat{\sigma}_t$ is used which corresponds to .11 probability for finding at least one outlier in twelve observations, given that no accounting errors are present.

Finally, if no individually large residuals are detected in the audit period, we calculate the upper precision limit on overstatement error at the .37 risk level. If $\text{UPL}_{1-A}(E)$ is less than $E^*$, the auditor concludes with .63 reliability that the balance for the year is not materially overstated. Figure 1 summarizes our analytical review and outcome classification procedure.

RESULTS
Four sets of results are presented in exhibits 1 through 4. These correspond to the four accounting error conditions in the audit period (AP). The AP conditions are (1) no accounting error, (2) one accounting error equal to a material amount $E^*$, (3) three errors of $E^*/3$, and (4) twelve errors of $E^*/12$. Thus, rejection under the first condition is a type I error, and acceptance under the latter conditions is a type II error.

In each of the exhibits, panel (1) presents results for a base period (BP) with no accounting overstatement error. Panel (b) presents results for a base period with one accounting error of $E^*$ while panels (c) and (d) present results for the less concentrated BP accounting errors of 3 of $E^*/3$ and 12 of $E^*/12$ respectively.

No AP Error
Line one of a panel (a) in exhibit 1 shows that for no errors other than the random residual, $\epsilon$, there are 65 type I errors or a rate of 32.5 percent. This rate is due to 4 percent which failed the Chow test, 6.5
FIGURE 1. Analytical Review and Outcome Classification Procedure

For given $x_1, x_2, M_1^{(1)}, M_{1}^{(2)}$, and $u$ series:

a. compute $x_1, x_2, M_1^{(1)}, M_{1}^{(2)}, u$ series

b. estimate $\alpha, \beta$, and $\delta^*$ via OLS

can $H_0: \delta^* = 0$
be rejected?

Yes

are there outliers in the base period?

No

are there outliers in the audit period?

No

are all audit period $u_i = 0$?

Yes

No

correct decision to Not investigate

Type II error

are all audit period $u_i = 0$?

Yes

No

correct decision to Investigate

Type I error

57
<table>
<thead>
<tr>
<th></th>
<th>Test Causing Rejection</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\beta = 0$</td>
<td>BP outliers</td>
</tr>
<tr>
<td>(a) No BP accounting errors:</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>No meas. error in x</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Low meas. error in x</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>High meas. error in x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(b) 1 BP accounting error in $E^*$:</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>No meas. error in x</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Low meas. error in x</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>High meas. error in x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(c) 3 BP accounting errors of $E^*/3$:</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>No meas. error in x</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Low meas. error in x</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>High meas. error in x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(d) 12 BP accounting errors of $E^*/12$:</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>No meas. error in x</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Low meas. error in x</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>High meas. error in x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test Causing Rejection</td>
<td>Summary</td>
<td></td>
</tr>
<tr>
<td>-----------------------</td>
<td>---------</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\beta = 0$</td>
<td>BP outliers</td>
</tr>
<tr>
<td>(a) No BP accounting errors:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No meas. error in x</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Low meas. error in x</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>High meas. error in x</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>(b) 1 BP accounting error of $E^*$:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No meas. error in x</td>
<td>0</td>
<td>31</td>
</tr>
<tr>
<td>Low meas. error in x</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>High meas. error in x</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>(c) 3 BP accounting errors of $E^*/3$:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No meas. error in x</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Low meas. error in x</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>High meas. error in x</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>(d) 12 BP accounting errors of $E^*/12$:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No meas. error in x</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Low meas. error in x</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>High meas. error in x</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>
### Exhibit 3. Simulation Results for 200 Audit Years With Three Audit Period Accounting Errors of $E^*$/3

<table>
<thead>
<tr>
<th>Test Causing Rejection</th>
<th>$\beta = 0$</th>
<th>$\text{BP outliers}$</th>
<th>$\text{Chow outliers}$</th>
<th>$\text{AP outliers}$</th>
<th>$\text{UPL}$</th>
<th>Correct decision to reject</th>
<th>Type II error</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>(a) No BP accounting errors:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No meas. error in x</td>
<td>0</td>
<td>0</td>
<td>35</td>
<td>46</td>
<td>65</td>
<td>146</td>
<td>54</td>
</tr>
<tr>
<td>Low meas. error in x</td>
<td>0</td>
<td>1</td>
<td>20</td>
<td>35</td>
<td>84</td>
<td>140</td>
<td>60</td>
</tr>
<tr>
<td>High meas. error in x</td>
<td>0</td>
<td>2</td>
<td>18</td>
<td>28</td>
<td>86</td>
<td>134</td>
<td>66</td>
</tr>
<tr>
<td><strong>(b) 1 BP accounting error of $E^*$:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No meas. error in x</td>
<td>0</td>
<td>31</td>
<td>1</td>
<td>13</td>
<td>81</td>
<td>126</td>
<td>74</td>
</tr>
<tr>
<td>Low meas. error in x</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>28</td>
<td>91</td>
<td>125</td>
<td>75</td>
</tr>
<tr>
<td>High meas. error in x</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>25</td>
<td>89</td>
<td>120</td>
<td>80</td>
</tr>
<tr>
<td><strong>(c) 3 BP accounting errors of $E^*/3$:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No meas. error in x</td>
<td>0</td>
<td>0</td>
<td>17</td>
<td>33</td>
<td>68</td>
<td>118</td>
<td>82</td>
</tr>
<tr>
<td>Low meas. error in x</td>
<td>0</td>
<td>1</td>
<td>14</td>
<td>26</td>
<td>84</td>
<td>125</td>
<td>75</td>
</tr>
<tr>
<td>High meas. error in x</td>
<td>0</td>
<td>2</td>
<td>12</td>
<td>25</td>
<td>82</td>
<td>121</td>
<td>79</td>
</tr>
<tr>
<td><strong>(d) 12 BP accounting errors of $E^*/12$:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No meas. error in x</td>
<td>0</td>
<td>0</td>
<td>28</td>
<td>38</td>
<td>58</td>
<td>124</td>
<td>76</td>
</tr>
<tr>
<td>Low meas. error in x</td>
<td>0</td>
<td>2</td>
<td>16</td>
<td>27</td>
<td>78</td>
<td>123</td>
<td>77</td>
</tr>
<tr>
<td>High meas. error in x</td>
<td>0</td>
<td>2</td>
<td>14</td>
<td>26</td>
<td>80</td>
<td>122</td>
<td>78</td>
</tr>
<tr>
<td></td>
<td>Test Causing Rejection</td>
<td>Summary</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------</td>
<td>------------------------</td>
<td>---------</td>
<td>----------------</td>
<td>----------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\beta = 0$</td>
<td>BP</td>
<td>AP</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Outliers</td>
<td>outliers</td>
<td>UPL</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(a) No BP</td>
<td></td>
<td></td>
<td>92</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>accounting</td>
<td></td>
<td></td>
<td>98</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>errors:</td>
<td></td>
<td></td>
<td>97</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No meas.</td>
<td>0</td>
<td>0</td>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>error in x</td>
<td></td>
<td></td>
<td>28</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low meas.</td>
<td>0</td>
<td>1</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>error in x</td>
<td></td>
<td></td>
<td>23</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High meas.</td>
<td>0</td>
<td>2</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>error in x</td>
<td></td>
<td></td>
<td>27</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(b) 1 BP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>accounting</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>errors in $E^*$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No meas.</td>
<td>0</td>
<td>31</td>
<td>0</td>
<td>91</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>error in x</td>
<td></td>
<td></td>
<td>1</td>
<td>123</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low meas.</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>102</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>error in x</td>
<td></td>
<td></td>
<td>14</td>
<td>119</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High meas.</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>98</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>error in x</td>
<td></td>
<td></td>
<td>12</td>
<td>115</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(c) 3 BP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>accounting</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>errors in $E^*/3$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No meas.</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>error in x</td>
<td></td>
<td></td>
<td>90</td>
<td>109</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low meas.</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>96</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>error in x</td>
<td></td>
<td></td>
<td>18</td>
<td>117</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High meas.</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>94</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>error in x</td>
<td></td>
<td></td>
<td>19</td>
<td>118</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(d) 12 BP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>accounting</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>errors of $E^*/12$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No meas.</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>18</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>error in x</td>
<td></td>
<td></td>
<td>81</td>
<td>108</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low meas.</td>
<td>0</td>
<td>2</td>
<td>7</td>
<td>17</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>error in x</td>
<td></td>
<td></td>
<td>91</td>
<td>117</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High meas.</td>
<td>0</td>
<td>2</td>
<td>7</td>
<td>91</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>error in x</td>
<td></td>
<td></td>
<td>21</td>
<td>121</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
percent which passed the Chow test but exhibited AP outliers, and 22 percent which passed all four regression-related tests but had upper precision limits (UPLs) greater than $E^*$. The effect of measurement error in the $x$ series (lines 2, 3) is an increase in the overall rejection rate of 42.5 and 44 percent respectfully. Thus measurement error in the independent variable results in increased type I errors and deprives the auditor of some of the potential benefits of regression.

Base period overstatement errors in $y$ yield a dramatic increase (0 to 15.5 percent) in rejection due to BP outliers for concentrated error. For the two less concentrated error cases, however, results are similar to those with no BP accounting overstatement error. In all cases with uncorrected BP overstatement errors, the total rejection rate (and thus the type I error rate) goes down. The reduction is caused by inclusion of the effect of overstatement errors in the parameter estimates which in turn include some overstatement in the "predictions" for the AP. The result is higher predictions for AP $y$, smaller $E$, and a lower UPL. The reduced type I error is, of course, little comfort to the auditor since the uncorrected BP errors lead to increased type II errors when AP accounting error is present as discussed below.

One AP Error of $E^*$

For concentrated AP accounting errors with no random measurement error in $x$ (exhibit 2, panel 1, line 1) regression is quite useful. The Chow test is failed in 60.5 percent of the years, and the AP outlier test is failed in another 32 percent. Of the remaining fifteen years, the UPL test is failed in four or 2 percent of the years, making a total rejection rate of 94.5 percent. Measurement error in $x$ reduces the discriminating power of the Chow test and this reduction is only partially compensated by increased rejection rates for the AP outlier and UPL tests. The result is a 3 and 4.5 fold increase in type II errors.

Introduction of concentrated BP overstatement errors in $y$ yields 260 percent, 55 percent and 33 percent increases in type II error when compared to no BP $y$ errors. For the less concentrated BP error cases, type II errors are less than for the concentrated errors due primarily to increased effectiveness of the Chow test. The addition of three BP overstatement errors of $E^*/3$ or twelve errors of $E^*/12$ compared with one AP error of $E^*$ allows frequent rejection of the constant regression parameters hypothesis.

Note that the highest type II error rate in exhibit 2 is 33.5 percent which is less than the nominal risk level of .37 indicated in the EPL.
test. This overall observed error rate is the result of a combination of
tests conducted at several risk levels. However, it is some comfort to
see that even with the measurement and BP overstatement errors
included, the observed error rate is less than the nominal rate for
applications assuming no measurement errors.

Less Concentrated Errors
For material AP accounting errors allocated $E^*/8$ and $E^*/12$, type II
error rates rise to 27 percent and 33.5 percent when there is no
measurement error in x. Even with measurement error in x, the rate is
well below the .37 level. However, the addition of BP errors in y,
whether or not x is measured with error, yields type II error rates at or
above the nominal level of the UPL test.

Comparison of exhibit 3 with exhibit 2 reveals the importance of
correcting small accounting errors if such series are to be used in
analytical review. Comparison with exhibit 4 shows that more nu-
merous smaller errors are of even more concern. Correction by month
of small errors discovered in an audit is not important for the year
under audit. However, the cumulative effect of leaving the errors
uncorrected can be the failure to reject future balances which contain
material accounting error in the same direction. These findings
should be of particular concern to auditors using regression for the
first time on a particular client. Part of the startup cost for a new
application is the cost of obtaining a reasonably correct monthly
series on which the projections for the current audit year can be based.

CONCLUSION
The simulations and tests presented here are based on a fixed and
simplified statistical model. Audit applications will often require
more effort in developing an adequate model on which audit period
projections can be based. However, the simplified setting can be
useful in assessing the marginal impact of data errors.

For the simplified setting, we find that random measurement error
in the independent variable does lead to an increase in type I and type
II errors. The errors are increased as the measurement error increases.
For randomly occurring base period overstatement errors of fixed
amounts in the dependent variable, type I error rates are reduced due
to the overestimation of audited value. Type II errors are considerably
increased for the same reason. Furthermore, the presence of some
patterns of data errors can lead to type II error rates much higher than
the nominal rate in the UPL tests. We conclude that auditors considering the application of regression should consider testing the accuracy of the data on which the analysis is to be based. Especially important is the correction of small accounting errors in the base period.

More complex (realistic) settings can also be simulated and the effect of measurement error assessed. Some settings of interest are models with

a. two or more parameter values over the base-audit period,
b. several independent variables—some functionally related to y and some not,
c. random measurement error in y, and
d. negatively correlated errors in y (to simulate the effect of poor “cutoffs”).

Determination of the behavior of the regression approach under these conditions will allow a more complete understanding of its reliability in analytical review.

REFERENCES


THE USE OF REGRESSION ANALYSIS ON AUDITS

Before turning to the paper, I would like to make a few remarks on the use of multiple regression in the audit process. Auditors have traditionally used analytical techniques during the audit process, and these included both time series (on both flows and stocks) and cross-sectional techniques. Multiple regression represents an improvement over informal methods for several reasons, including the following: regression techniques can handle a large number of variables simultaneously (whereas informal techniques can generally consider only two variables simultaneously) and regression techniques result in the calculation of confidence intervals, meaning that risk levels can be controlled explicitly and the significance of deviations can be assessed. Possible uses of regression techniques (or analytical techniques in general) in the audit include the following:

- As an attention director at the outset of the audit—outliers (either time series or cross-sectional) are identified for more intensive auditing.
- As a substitute for all or part of a test of details—if regression analysis shows a balance to be substantially correct, the substantive test of details can be reduced or eliminated.
- Roll forward from interim audit to financial statement date—where internal controls are reliable and the auditor’s report is required soon after the financial statement date, the substantive
audit may take place at an interim date (as of one to several months before the year end); intervening transactions must be reviewed so that many unusual patterns that develop can be audited.

- Final consistency check—at the end of the audit, after all audit adjustments (if any) have been recognized, the financial statements to be issued can be reviewed to test their reasonableness, internal consistency, etc.
- Quarterly reviews—quarterly financial information is not audited, but is reviewed; regression offers improved review techniques.

Additional uses may develop (such as analysis of market data to evaluate the materiality measure for a client or analysis of an audit firm’s client portfolio to identify outlier clients that require more intensive audit effort).

Since the above applications are important and intuitively appealing, a natural question arises: why is regression analysis not commonly used in audits? The answer includes three general considerations: high costs, low benefits, and technical validity questions.

**HIGH COSTS**

There are three key factors making regression analysis a costly procedure:

- Training—auditors using regression analysis must be adequately trained, and training in regression to develop usable skills is quite time consuming and expensive (it appears to me that regression training for audit use would be even more expensive than statistical sampling training, which is itself quite expensive); valid regression modeling depends on meeting a number of technical regression assumptions that require considerable skill to evaluate.
- Model building—a great deal of time, effort, and skill is required to build models that reflect the underlying economic relationships for each client.
- Data acquisition—often data must be disaggregated, recast, cleaned up, or otherwise processed before they are useful for the regression model; useable data are seldom conveniently available at negligible cost.

(Computation costs are typically negligible and need not be considered, due to the cheap availability of a variety of computerized regression packages.)
Certain techniques may help to reduce these costs, but must be dealt with carefully:

- Computerized rules to test whether the data meet the regression assumptions (e.g., linearity, normality and independence of residuals, etc.) and to transform the data to "fix" any deficiencies—mechanical fixes may simply obscure the problems, not fix them.

- Step-wise regression to reduce the cost of model building—the problem here is that a step-wise regression procedure is certain to develop a high $R^2$ when operating on data series that are highly correlated in the first place, as accounting data are. Once, to amuse myself, I generated a $10 \times 20$ array of random numbers (simulating 20 variables for 10 periods) and developed, almost immediately, through step-wise regression, an equation that "explained" 95 percent of the variance.

A more appropriate technique, but at the cost of specificity, would be to develop classes of models by industry, client type, etc., that would fairly generally capture the underlying economic relationships. This would presumably lower both the educational and model-building costs, while marginally reducing the explanatory power of the model.

**LOW BENEFITS**
Regression models typically have low resolving power compared with audit precision requirements. In a typical audit situation, a company may return, say, 10 percent on assets and have a 50 percent marginal tax rate. If net income is to be correct to $\pm$ 5 percent, assets must be correct to $\pm$ 1 percent (ignoring the effect of liabilities). Even a regression model with a very high $R$ is unlikely to get anywhere close to that precision requirement (i.e., has wide prediction intervals). If a regression model tells an auditor that accounts receivable are off no more than $2,000,000 when he must have assurance they are off no more than $100,000 to sign the audit opinion, he cannot reduce his other substantive procedures very much in reliance upon the regression results. (Informal analytical models have even lower resolving power, but then, they are much cheaper.)

**TECHNICAL VALIDITY**
Auditors are subject to high standards of professional care and assume onerous legal burdens with substantial potential liability in the case of substandard work. In this environment, auditors are quite concerned about the technical validity of new procedures that may be
adopted. In the case of regression models, auditors are particularly concerned about the robustness of the technique in the face of such problems as the following:

- nonlinear relationships;
- multicollinearity;
- singularity;
- autocorrelation;
- heteroscedasticity;
- nonnormality; and
- erroneous data (independent and dependent variables).

Auditors are also worried about spurious correlations and whether, after dredging the data, any degrees of freedom are left for inference.

The Kinney-Salamon (K-S) paper addresses one of these technical issues—errors in the data—and so makes a contribution to the use of regression in auditing.

**KINNEY-SALAMON PAPER**

The K-S paper is a straightforward investigation of certain aspects of data error in regression. It is not an auditing paper (since the issues it explores are not auditing, but statistical, issues), but should be of interest to auditors who use or intend to use regression. The paper states a simple problem, models the problem using Monte Carlo simulation, states the results, and interprets the results. Although the problem could be dealt with analytically, the use of Monte Carlo methods is appropriate and valid.

The conclusion is unremarkable: introduction of noise into data series degrades the explanatory power of regressions on those data series. The value of the paper is that it quantifies this effect under certain limited assumptions. These limitations, which I hope the authors will relax in future work, include the following:

- simple regression (only one independent variable);
- thirty six base periods and twelve current periods (applying primarily to flow accounts, not stock accounts or cross sections); and
- no secular trend to simulated data (applying primarily to differenced data).

On this last point, a further word is in order. The error-free data in the simulation model are distributed with normal density both horizontally (x values) and vertically (error terms). The result is that the isoprobability on \([E(x), E(y)]\). Audit-period, error-free realizations

68
will be contained in precisely the same ellipses. The addition of random overstatement errors in the base period will have the effect of widening and raising and/or slightly rotating the regression confidence band, all of which will result in keeping most points in the ellipse still within the confidence band. If, on the other hand, the data had a secular trend, rotation of the confidence band would tend to exclude more audit period realizations. Since accounting data usually have a secular trend, and since differencing the data is not always appropriate, the simulation study should be extended.

The logic path in figure 1 has not been justified by appeal either to theory or practice. In fact, I have problems with the top two diamonds: in the first, if β = 0, the auditor does not reject the audit period, he looks for a better model; in the second, outliers in the base period should not cause rejection of the audit period, but investigation of the outliers to see whether the model is misspecified.

The next three diamonds seem reasonable, but what theory justifies their use, either separately or in series?

The results in the exhibits show how many cases are rejected by each test given that the prior tests did not cause rejection. Thus, I cannot evaluate what results the simulation would have assuming that one or more logic tests were eliminated. If the data in the exhibits were presented raw with a column to subtract duplication and arrive at the net rejections, alternate evaluations would be facilitated.

Although the simulations can be viewed as testing certain hypotheses about rejection rates, I think it is more important to look at them as quantifying the magnitude of the effects. Therefore, the paper should make some attempt to quantify the achieved precision of the simulation results (as opposed to simply testing hypotheses about whether the noted differences are significant).

In general, I regard the paper as shedding some light on a problem of interest, but that extensions would be needed (especially modeling more, and more realistic, situations) in order for policy conclusions to be based on the results.

**SUMMARY**

In order for regression analysis to become a generally useful tool, three types of research are needed: research to

1. reduce the cost of using regression,
2. improve the predictive power of models, and
3. establish the conditions under which use of models is valid.

The K-S paper makes a useful contribution to meeting the third need.
Discussant's Response to "The Effect of Measurement Error on Regression Results in Analytical Review"

WANDA A. WALLACE

Professors Kinney and Salamon are to be complimented for directing the profession's attention to the existence of measurement errors and their effects on the tool of regression analysis as currently applied in analytical procedures. Since measurement errors are likely to result from negligence, intentional misstatement, random error, and/or the common practice of auditors of not requiring clients to record "immaterial" journal entries, the investigation of the effect of such errors on statistical tests and the level of error in audit decisions is of interest to the profession.

Kinney and Salamon's conclusion that increased Type II errors are observed when several small uncorrected errors are imputed into the base period has implications for practitioners. Before utilizing past audited data for regression analysis, "passed" or "unbooked" adjusting journal entries proposed by the auditors during their examinations in prior periods should be corrected. The practitioner should likewise be careful in utilizing unaudited data of new clients in the analytical review procedure. The AICPA's recent exposure draft on analytical review suggests "The auditor should consider that financial information might not be reliable..." (AICPA, 1978, p. 6). Without prior examinations, it would be difficult to preclude even large measurement error from clients' data. It may be preferable to utilize industry data and externally generated statistics similar to those discussed in Neumann (1974), since the nature of these statistics'
measurement error may be better understood, more easily anticipated, or at least less likely to have a bias in a particular client's favor.

Although Kinney and Salamon have provided some insight on the effect of measurement error, there are two major conceptual problems with their study. In order to suggest any policy implications or normative prescriptions, it is necessary to positively discuss all of the dimensions of the problem. Yet, the authors do not discuss the major incentives faced by the actors in the audit setting nor the loss function implicitly assumed by the methodology applied. The second major issue addresses the question: Are the techniques being applied those having the greatest comparative advantage? Finally, some specific comments on details of the paper and possible extensions are discussed.

**THE LOSS FUNCTION**

Before addressing the question of measurement error's effect on the discovery of misstatements and the tool of regression analysis for analytical review, it is important to examine the incentives of all the parties involved in the audit setting.

In addition to concern for negligence and random error effects, a relevant question in approaching the analytical review process is what are the incentives of the firm's management and employees to misstate account balances? Incentives are provided in the form of profit-sharing plans, performance evaluation measures, and bond covenant requirements which are frequently tied to accounting numbers. These incentives, a firm's tax situation, anti-trust considerations, the weaknesses in control over assets which provide opportunity for stealing, and similar cost/benefit issues which relate to the firm's motivation for misstatements should be discussed in the paper. These issues are relevant in determining the types of errors most likely to occur, the symmetry of the expected errors, and whether intentional errors are more likely to be accounted for “bottom line” directed. These incentives may well affect the selection of an analytical review technique and the specification of statistical levels of significance.

A loss function for the auditor is necessarily assumed in specifying significance tests of hypotheses, and it may well be that the expected value of the sum of squared deviations—utilized by Kinney and Salamon—does not exactly reflect the auditor's actual penalty from errors in estimation.

The cost of rejecting a correct account balance is the average loss
from additional sampling and investigation costs. The cost of accepting an incorrect account balance is primarily the cost of liability cases and related penalties. Since these costs are likely to exceed additional sampling and investigation costs, the important question concerns the probability of a liability suit. For reasons not fully understood, court awarded damages are more frequent if reported values are ex post “overstated” rather than ex post “understated.” Kellogg (1978) reports that out of 148 cases sampled, only 14 were seller suits (suits by sellers of securities alleging they sold based on statements reporting understated values), whereas 134 cases were purchaser suits (suits by purchasers alleging financial statements reported overstated values). Given this apparent asymmetry of suits and awards, the use of mean square error or absolute error in hypotheses testing may be inappropriate.

The loss function of the auditor and management should be discussed to provide a positive formulation of the problem which can be referred to in the decision process of specifying acceptable levels of errors (methods of mathematically incorporating loss functions are discussed in Hogg and Craig, 1970). The reasonableness of decision rules in light of linear, nonlinear, quadratic, and various other forms of loss functions warrants investigation.

Kinney and Salamon compare the trade-off of Type I and Type II errors in the simulation results involving dependent variable measurement error. The reasonableness of any such comparison rests on the symmetry of the loss function.

In addition to the criticism that the relevant cost and benefit dimensions of the problem are not explicitly considered, the question arises as to whether simulation analysis is the best approach to the problem addressed by Kinney and Salamon.

**SIMULATION VERSUS ALGEBRA**

A more useful way to consider measurement error’s effect in simple regression analysis is through algebraic analysis. Simulation analysis is not required to generate the results reported. Most of the results

---

1As an example, the Texas Gulf Sulphur (TGS) Corporation case involved the courts’ establishing the possibility that eventual legal damages to be paid by the corporation resulting from disclosure issues would exceed the net worth of TGS by $150,000,000. (All cases related to the TGS case were settled on December 30, 1971, for $2,700,000.) (Kellogg, 1978) Similarly, audit firms have incurred substantial court costs and settlements.
Exhibit I. Mathematical Analysis of the Effect of Measurement Error in the Dependent Variable on Regression Results in Analytical Review

I. Model Without Measurement Errors

\[ y_t = \alpha + \beta x_t + e_t \]

\[ \hat{\beta} = \frac{\sum(x_t - \bar{x})(y_t - \bar{y})}{\sum(x_t - \bar{x})^2} \]

\[ \hat{\alpha} = y - \hat{\beta} \bar{x} \]

II. Determine \( \hat{\beta} \) and \( \hat{\alpha} \) values, given errors \( c_t \) in the base period \( y_t^* = y_t + c_t \)

\[ b^* = \frac{\sum(x_t - \bar{x})(y_t + c_t - (\bar{y} + c))}{\sum(x_t - \bar{x})^2} = b + \frac{\sum(x_t - \bar{x})(c_t - c)}{\sum(x_t - \bar{x})^2} = b + d \]

\[ a^* = \frac{(\bar{y} + c) - \bar{x}(b + d)}{b^*} = a + \frac{(c - d)(c)}{b^*} \]

III. Analyze the Effect of the Above Measurement Error in the Base Period’s Dependent Variable on the Analytical Review Filter Rules

a. Can \( H_0: \beta = 0 \) be rejected?
   If \( c_t \) and \( x_t \) are uncorrelated, \( b^* = b \), hence this filter rule is unaffected.

b. Are there outliers in the base period?
   \[ \hat{c}_t^* = (y_t^* - \hat{y}_t^*) = (y_t + c_t - (\hat{y}_t + c)) \]
   \[ = e_t + (c_t - c) + d(x_t - \bar{x}) = e_t + (c_t - c), \text{ if } d = 0 \]

If \( c_t \) and \( x_t \) are uncorrelated, the above description of the residual term suggests \( |\hat{c}_t^*| > |e_t| \). Hence, the number of base period outliers which will be observed increases.

c. Can \( H_0: (\alpha, \hat{\beta}) \) Base Period = (\( \alpha, \beta \)) Audit Period be rejected?
   The Chow Test compares SEPARATE REGRESSIONS (\( R_1 \)):
   Residuals Sum of Squares = \[ \sum_{(base)} (y_t - \hat{y}_t)^2 \]
   \[ + \sum_{(audit)} (y_t - \hat{y}_t)^2 \text{ with } (n_1 - 2) + (n_2 - 2) = (n-4) \text{ d.f.} \]

and COMBINED REGRESSION (\( R_2 \)):
   Residuals Sum of Squares = \[ \sum_{1}^{n} (y_t - \hat{y}_t)^2 \text{ with } (n-2) \text{ d.f.} \]

by means of the \( F \)-Statistic

\[ F_{2, n-4} = \frac{(R_2 - R_1)/2}{R_1/(n-4)} \]

Graphically the three regressions are being compared to determine whether I and II are significantly different, i.e., \( H_0: \alpha_1 = \alpha_2 = \alpha; \beta_1 = \beta_2 = \beta \).
Total $SS^*_{1} = \frac{1}{N} \sum_{i=1}^{N} (y_i - \bar{y})^2 = \frac{1}{N} \sum_{i=1}^{N} [y_i + c_i - (\bar{y} + \bar{c})]^2$

$= \frac{1}{N} \sum_{i=1}^{N} (y_i - \bar{y})^2 + \frac{1}{N} \sum_{i=1}^{N} (c_i - \bar{c})^2 + 2 \frac{1}{N} \sum_{i=1}^{N} (y_i - \bar{y})(c_i - \bar{c})$

$= S_y^2 + S_c^2 + 2S_{yc}$

Note: $1$ = Base Period; $S_{yc}$ = covariance of $y$ and $c$ during base period.

**Separate Regression ($R_1$)**

$b^* = \frac{\frac{1}{N} \sum_{i=1}^{N} (x_i - \bar{x})(y_i^* - \bar{y}^*)}{\frac{1}{N} \sum_{i=1}^{N} (x_i - \bar{x})^2}$

Note: $\frac{1}{N} \sum_{i=1}^{N} (x_i - \bar{x})(y_i^* - \bar{y}^*) = \frac{1}{N} \sum_{i=1}^{N} (x_i - \bar{x})(y_i - \bar{y}) + \frac{1}{N} \sum_{i=1}^{N} (x_i - \bar{x})(c_i - \bar{c})$

$= S_{xy} + S_{xc}$

$b_1^* = \frac{S_{xy} + S_{xc}}{S_{x}^2} = b_1 + t_1$

Note: $t_1$ is the slope of $c$ on $x$ period one $= \frac{S_{xc}}{S_{x}^2}$

$\gamma_1 + t_1 x$

$\begin{align*}
(\text{Regression } SS1)^* &= (b_1^*)^2 \sum_{i=1}^{N} (x_i - \bar{x})^2 \\
&= b_1^2 S_{x}^2 + (t_1^2 + 2b_1 t_1)S_{x}^2 \\
&= \text{Reg } SS1 + (t_1^2 + 2b_1 t_1)S_{x}^2
\end{align*}$
\[ (\text{Residual SS1})^* = (\text{Total SS1})^* - (\text{Regression SS1})^* \]

\[ = \text{Residual SS1} + S_c^2 + 2S_{ycl} \]

\[ = (t_1^2 + 2b_1t_1)S_{x_1}^2 \]

\[ = \text{Residual SS1} + (S_c^2 - t_1^2S_{x_1}^2) + 2(S_{ycl} - b_1S_{xcl}) \]

Regression SS2* = Reg SS2
Residual SS2* = Residual SS2

\[ \text{Combined Regression (R}_2\text{)} \quad \text{(Note: } \bar{x}, \bar{y}, \text{and } \bar{c} \text{ are pooled values)} \]

\[ b^* = -\frac{1}{\sum_{t=1}^{n}(x_t - \bar{x})(y_t - \bar{y})} + \frac{1}{\sum_{t=1}^{n}(x_t - \bar{x})(c_t - \bar{c})} \]

\[ + \frac{1}{\sum_{t=1}^{n}(x_t - \bar{x})} \]

\[ b^* = b + 1 \frac{\sum_{t=1}^{n}(x_t - \bar{x})(c_t - \bar{c})}{\sum_{t=1}^{n}(x_t - \bar{x})^2} \]

\[ = b + T_1 \]

Note: \( T_1 \) is the slope of \( c \) on \( x \) for the combined regression = \( \frac{S_{xc}}{S_x^2} \).

[Regression SS (1 & 2) combined]*

\[ = (b + T_1)^2 \frac{1}{\sum_{t=1}^{n}(x_t - \bar{x})^2} \]

\[ = b^2 \sum_{t=1}^{n}(x_t - \bar{x})^2 + (T_1^2 + 2bT_1) \sum_{t=1}^{n}(x_t - \bar{x})^2 \]

\[ = \text{Reg SS} + (T_1^2 + 2bT_1) \sum_{t=1}^{n}(x_t - \bar{x})^2 \]

(DEVIATIONS)* = (Reg. SS combined)* - (Separate Regression SS or Regression SS1* plus Regression SS2*)

\[ = \text{Reg. SS} + (T_1^2 + 2bT_1) \sum_{t=1}^{n}(x_t - \bar{x})^2 \]

\[ - \text{Reg. SS1} - (t_1^2 + 2b_1t_1)S_{x_1}^2 - \text{Reg. SS2} \]
\[\begin{align*}
&= \text{Deviations} + (T_1^2 + 2bT_1) \sum_1^N (x_i - \bar{x})^2 \\
&- (t_1^2 + 2b_1t_1)S_{x1}^2 \\
F^* &= \frac{(\text{Deviations})^*/n-4}{(\text{Residual SS}(R_1))^*/(n-4)} \\
2F^* &= \frac{\text{Deviations} + (T_1^2 + 2bT_1)S_x^2 - (t_1^2 + 2b_1t_1)S_{x1}^2}{(\text{Residual SS}(R_1))^* + \text{(Separate Error SS)}} + (S_{c1}^2 - t_1^2 S_{x1}^2) + 2(S_{ycl}^2 - b_1S_{xcl}^2)
\end{align*}\]

Given that \(c_t\) and \(x_t\) are uncorrelated
\[
\frac{2F^*}{n-4} = \left(\frac{F}{S_c^2}\right) \frac{2}{(n-4)}
\]

Hence, the F Statistic is understated and a decreased number of rejections will occur for the Chow test rule.

d. Are there outliers in the audit period? (Note: \(\gamma = \text{standard deviation}\))
\[
\gamma_t^2 = V(y_t + c_t) = V(y_t) + V(c_t) + 2 \text{ cov}(y_t, c_t)
\]

Note: \(V = \text{variance};\ V(y_t) = V(e_t) = \gamma_t^2\)

Given that \(c_t\) and \(x_t\) are uncorrelated
\[
\gamma_t^2 = V(y_t) + V(c_t)
\]

and \(\hat{\gamma}_E^2\) is increased. Since the rejection rule, i.e., \(\hat{\epsilon}_t \geq 2.44 \hat{\gamma}_E^2\), as \(\hat{\gamma}_E^2\) increases, it becomes more difficult to find \(\hat{\epsilon}_t \geq 2.44 \hat{\gamma}_E^2\) leading to a decreased number of audit period outliers.

Note: \(\hat{\gamma}_E\) is the product of \(\hat{\gamma}_t\) and a positive quantity described by Kinney and Salamon, p. 52.

e. Is \(UPL_{1-A}(E) \geq E^*\)

(Note: \(UPL = \text{Upper Precision Limit}\))
\[
\hat{E} = \frac{36 + 12}{36 + 1} \sum_{t=1}^{\hat{y}_{rt}} (y_{rt} - \hat{y}_t)
\]

If it is assumed \(b^* = b\) and \(\alpha^*\) is overstated (as given in the simulation) then
\[
\hat{y}_t^* = \hat{\alpha}^* + \hat{\beta}_0 + \hat{\beta}_1 x_t \\
= (\hat{\alpha} + b\hat{\beta}_1) + \hat{c} \\
= \hat{y}_t + \hat{c}
\]
The result that \(\hat{y}_t^*\) is overstated leads to
\[
y_{rt} - \hat{y}_t^* = c_t - \bar{c}
\]
which means that \( \hat{E} \) will decrease. However as described earlier in section d, \( \gamma_t^2 \) is increased by the presence of dependent variable error in the base period.

Since

\[
UPL_{1,A}(E) = \hat{E} + t_{\text{statistic}_{1,A}, 36-2} \cdot \hat{\gamma}_E^2
\]

and \( \hat{E} \) is decreased while \( \gamma_t^2 \) is increased, the change in rejection rates for this filter rule rests upon the relative magnitude of changes in \( \hat{E} \) and changes in \( \gamma_t^2 \). In the simulation, the parameters tested led to domination by \( \hat{E} \), resulting in a lower UPL_{1,A} level.

obtained can be predicted by applying algebra, and those which cannot be directionally anticipated are largely a function of the numbers and assumptions presented and cannot be generalized. In other words, closed form expressions for the results can be obtained by algebra and the simulation results not clearly predicted are strictly due to the parameters chosen and cannot be generalized. As an example, exhibit 1 presents the algebraic analysis of the effect of measurement error in the dependent variable for the base period on each of the filter rules applied in the analytical review procedure assumed by Kinney and Salamon.

Algebraic derivation of the dependent variable's measurement error for the base period's effect on the slope and intercept of simple regression facilitates analysis of the error's effect on each of the test statistics utilized during the analytical review process. Unambiguous predictions that beta significance for the base period is unaffected, the base period outliers are increased in number, the Chow test has an understated F-statistic, and the number of audit period outliers are decreased in number can be obtained. Analysis of the upper precision limit indicates that the direction of measurement error effect cannot be predicted without knowing the relative magnitude of the difference in book values and predicted audit values and the size of the standard deviation of the error (the "closeness" of residual values and the measurement error in the dependent variable is also relevant).

A similar analysis is possible for measurement errors in the independent variable. Analysis of the simple regression equations predict the performance of filter rules, depending on correlation and magnitude assumptions, and the subsequent effect in terms of increased or decreased Type I and Type II errors.

It would appear that a paper mathematically anticipating the effect of measurement errors and audit implications is preferable to a simulation analysis of simple regression. The results can then be generalized and the reason for bias in the filter rules can be better understood.
METHODOLOGICAL PROBLEMS WITH THE SIMULATION
Not only does simulation analysis lack the comparative advantage for use in this analysis, but, in addition, there are some problems with the technique being applied. The authors state that they omitted distributional tests for normality and serial independence of the residuals, suggesting that transformation of variables are available which usually allow acceptance of these assumptions. Without testing the residuals, it is impossible to attribute the results reported entirely to changes in measurement error. Since the filter rules rely on t and F tests which assume uncorrelated residuals, it is possible that the hypothesis tests are inapplicable to the simulation model. This is especially true in light of the fairly small samples being examined (36 and 12 sample points for the base and audit period, respectively).

Similarly, it is assumed that the error terms are uncorrelated with y's and x's; although these expectations may hold for large samples, given random generation of the errors, it is unlikely that the correlation is zero for the small sample sizes being examined. Since many of the effects of measurement error depend on the existence and direction of such correlation, a scatter diagram for a random sample of the forty-eight month periods examined of c against x and c against y would have been helpful. While it is probable that the 200 periods examined eliminated any consistent effect of correlation, Kinney and Salamon should recognize that a problem with correlation of errors exists per audit period examined.

Finally, there is some question as to why the statistical levels of significance are varied for the different filter rules, as opposed to specifying a significance level (p value) and constantly applying that cut-off per analytical procedure. No explanation for the .04 to .10 range is offered by the authors. Similarly, there is no rationale provided for the .87 upper precision level (UPL). Is 63 percent reliability generally accepted in practice for analytical review? Will the comparison of errors observed relative to the UPL yield different results if the amounts of errors, frequency of errors, direction of errors, and/or the reliability demanded by the auditor is varied?

Any simulation analysis is limited in the sense that conclusions are only applicable to parameters tested, and there is a lack of a test for internal validity, with the exception of observing results consistent with mathematically predictable effects. It is especially important to recognize that certain effects on the filter rules reported by Kinney and Salamon are due entirely to the assumption that only overstatement errors exist in the dependent variable. While Kinney and Salamon
state that the constrained overstatement error case seems to be potentially the most dangerous for auditors, the effect of understatements of expense and liability accounts appear to be just as dangerous. The effect of both types of measurement errors should be examined.

In addition to these problems with the simulation analysis applied, the study would be improved if available alternatives to ordinary least squares were explored.

**ALTERNATIVE ESTIMATORS**

Since measurement error in the independent variables poses a serious estimation problem and alternative estimators like instrumental variables and other maximum likelihood methods are available, it would have been interesting had the alternative method(s) been applied to determine whether a significant improvement over ordinary least squares is observed.

These estimators are discussed in Johnston (1972), Lee (1975), Cochran (1968), and Madansky (1959). Based on a brief review of this literature, it would appear that an easy instrumental variable adjustment to apply in auditing to obtain a consistent estimator, given that measurement errors in the independent variables are known to exist, is the grouping method. Wald (1940) describes this method which results in a consistent estimator; however, this estimator is likely to have a large sampling variance. Bartlett (1949) suggests a three-way grouping, but due to auditing problems with sample size, it appears to be inapplicable to analytical review analysis. When multiple regression analysis is applied, Durbin’s (1954) method utilizing instrumental variables is reported to show efficiency for large samples, relative to least-squares, of 96 percent and for samples of size twenty, 86 percent. None of these methods require identification of the measurement error variance.

Madansky (1959) suggests that knowledge of the range and standard deviation of errors, observations of a third set of variables related to the original variables of interest, and replications of observations on each independent variable are helpful in dealing with the prob-

---

1 The method requires that the independent variable be ranked, if n is odd the central observation be omitted, and the sample be split into two equal groups. Beta is estimated by computing the difference in the mean Y and mean X values.

2 Durbin’s method requires that X’s be ranked in order; then an instrumental variable is defined as the rank order for each explanatory variable.
lems of measurement error. Since the maximum likelihood methods available are buttressed with fairly strong assumptions about the covariance matrix of the measurement errors, the properties of variables commonly analyzed during the analytical review process and their probable correlation with errors warrant investigation.

A PROPOSED APPLICATION OF SIMULATION ANALYSIS
While simulation analysis lacks the comparative advantage relative to algebraic derivation in the Kinney and Salamon study, simulation analysis can be a useful methodology for investigating complex measurement error situations. The two major criticisms discussed earlier could combine to create a useful setting for simulation analysis, i.e., a nonlinear loss function and the "mixed" effect of a measurement error can best be examined by systematic variation of inputs to the analysis and observation of the effect on the audit decision and the level of Type I and Type II errors. In general, simulation analysis has advantages when complex interaction effects exist or multiple sets of equations are being examined, the combination of which may be nonlinear. For the auditor's analytical review situation, initially the effect of measurement errors in simple regression and multiple regression should be mathematically derived (as illustrated in Lee, 1975). Based on these predictions, a systematic testing of various levels of correlations, measurement errors, autocorrelations, reliability measures, and significance tests based on known relationships in audited data would provide insight into the substantive effect of measurement errors on ordinary least squares.

Multiple regression analysis appears to be much more applicable than simple regression to analytical review, given seasonality factors and the known condition that "Year-end adjustments resulting from accounting policy changes, errors in cost estimations and allocations, and write-offs, of facilities or inventories may make previously reported interim data inappropriate as bases for comparisons" (Edwards, et al., 1972, p. 221). This same idea is expressed by Kaplan (1978) when he suggests that fiscal end-of-year months be excluded in regression analysis of monthly accounting data. These effects may cause differing fluctuations in monthly measures across years (such issues have been raised in the recent FASB Discussion Memorandum, "Interim Financial Accounting and Reporting"). Adjustment for such fluctuations may require the use of multiple regression. Since interaction effects are complicated with multiple independent variables, simulation analysis becomes an appropriate technique.
CONCLUSION
The effect of measurement error on regression analysis is an area that warrants further investigation as suggested in the proposals for extensions of simulation analysis discussed herein and those suggested by Kinney and Salamon. A more thorough investigation of different sizes of correlation, autocorrelation, levels of significance, assumed loss functions, available methods of correction, and properties of the measurement errors commonly observed would provide a basis to evaluate the propriety of techniques being utilized. It would also facilitate the understanding of measurement errors’ effect so that the significance levels of the filter rules applied in analytical review could be adjusted to reflect the mathematically known overstatement or understatement bias in the test statistic being applied.

REFERENCES

Bartlett, M. S., “Fitting a Straight Line When Both Variables are Subject to Error,” Biometrics (Vol. 5, 1949), pp. 207-12.
Wald, A., “The Fitting of Straight Lines if Both Variables are Subject to Error,” Annals of Mathematical Statistics (Vol. 11, 1940), pp. 284-300.
Session

THREE
Audit Tests for Internal Control Reliance

WILLIAM L. FELIX, JR. and JAMES L. GOODFELLOW

INTRODUCTION
The purpose of this paper is threefold:
1. To review and summarize some of the existing professional literature on tests of internal control reliability;
2. To point out some of the confusion that presently exists with respect to such tests in both the literature and in practice; and
3. To explore the context and the nature of certain choices which must be addressed in performing tests of internal control reliability.

Note that our objectives are limited. We intend to identify issues and, hopefully, stimulate further discussion and analysis.

The paper is organized as follows:
Summary of Professional Literature;
Frequency and Monetary Measures of Transactions Containing Procedural Deviations;
A Description of Sample Selection Methods;
Analysis of Sampling Plans;
Example of Various Alternatives;
Conclusions; and
References.

CATEGORIES OF TESTS OF COMPLIANCE
In both the United States and Canada, compliance tests are divided into two categories based on whether documentary evidence is avail-
able. The following quotation from the CICA Handbook (CICA, 1978) illustrates this distinction:

Compliance Procedures are generally of two types:
(a) tests requiring inspection of documents supporting transactions to gain evidence (such as signatures and initials) that controls such as checking and authorizing have been properly performed. Such a compliance procedure could involve the auditor not only in verifying that the document has been initialed to indicate that a calculation was checked, but also in re-checking the calculation in an attempt to verify that the calculation was in fact properly checked;
(b) procedures requiring enquiry about and observation of controls which leave no audit trail (for example, with respect to segregation of incompatible functions, to determine who actually discharges each function, not merely who is supposed to discharge it). (5215.14)

It should be noted that the Canadian pronouncement acknowledges that the auditor might “re-check” the calculation or procedure which supports the signature, initial or audit stamp whereas the AICPA pronouncement is silent on this issue and seems to imply that such “re-checking” is not necessary.

A SUMMARY OF PROFESSIONAL LITERATURE
We shall begin with a brief review of the nature and objectives of compliance tests as found in the professional literature in the United States and Canada.

Purpose of Tests of Compliance
The term “tests of compliance” first appeared in the professional literature in Statements on Auditing Procedure No. 54 (now section 320 of Statement on Auditing Standards No. 1) which states that “The purpose of tests of compliance is to provide reasonable assurance that the accounting control procedures are being applied as prescribed.” (520.55)

The equivalent Canadian pronouncement is found in section 5200 of the CICA Handbook. Although this section was not issued until June of 1977, the standards for compliance testing are generally the same as those contained in Statement on Auditing Standards No. 1, although some differences in terminology exist. For example, tests of compliance are referred to in the Canadian material as “compliance procedures.” The following quotation illustrates the Canadian position on the objectives of compliance procedures:

The purpose of performing compliance procedures is to gain a sufficient degree of assurance that the internal controls as described to the auditor exist
and are applied effectively throughout the period of intended reliance on the
controls. (5215.13)

The question in the United States is as follows:

Some aspects of accounting require procedures that are not necessarily re-
quired for the execution of transactions. This class of procedures includes the
approval or checking of documents evidencing transactions. Tests of such
procedures require inspection of the related documents to obtain evidence in
the form of signatures, initials, audit stamps, and the like, to indicate whether
and by whom they were performed and to permit an evaluation of the
propriety of their performance. (320.58)

The issue of what constitutes acceptable audit evidence in such
circumstances (i.e., does the auditor need to reperform the control
procedure to evaluate the propriety of its performance?) is not only a
matter of current professional debate, but also an important factor to
be considered when deciding on the sampling strategy to be used in a
compliance test. For example, if such “reperformance of control
procedures” provides the auditor with substantive evidence (as is
usually the case), then it may be advantageous for the auditor to select
a sampling plan that can accurately measure the substantive con-
tribution of this “dual purpose test.”

The category of compliance tests that relates to controls that “leave
no audit trail of documentary evidence” raises many practical and
theoretical issues involving timing, extent and evaluation of evidence
obtained. These problems are worthy of study but are beyond the
scope of this paper. Instead we shall devote our attention to the types
of compliance tests that are designed when an “audit trail of docu-
mentary evidence” is available.

**Requirements for Tests of Compliance**

The necessity for performing compliance tests when the auditor relies
on internal control is also clearly stated in both countries:

Such tests are necessary if the prescribed procedures are to be relied upon in
determining the nature, timing, or extent of substantive tests of particular
classes of transaction or balances. . . . (320.55)

The auditor should conduct compliance procedures in respect of those in-
ternal controls on which, following his preliminary evaluation, he still
intends to rely in determining the nature, extent and timing of substantive
auditing procedures. (5220.12)

**Sampling Issues**

Finally, both countries take a neutral position as to whether statis-
tical sampling is superior or preferred to judgement sampling in tests
of compliance. However, both countries do state a preference to some sort of representative sampling from the transactions executed throughout the period under audit.

As to accounting control procedures that leave an audit trail of documentary evidence of compliance, tests of compliance as described in paragraph 58 ideally should be applied to transactions executed throughout the period under audit because of the general sampling concept that the items to be examined should be selected from the entire set of data to which the resulting conclusions are to be applied. (320.61)

Such a representative sample is best obtained by random sampling. The equivalent Canadian pronouncement is worded similarly to that of the AICPA except that the modifier “ideally” is replaced by “normally.” Both countries acknowledge that such tests are often performed at an interim date and that the “application of such tests throughout the remaining period may not be necessary.” (320.61) The factors to be considered where internal control reliability tests may be applied only through the date of interim tests are set forth as follows:

... (a) the results of the tests during the interim period, (b) responses to inquiries concerning the remaining period, (c) the length of the remaining period, (d) the nature and amount of the transactions or balances involved, (e) evidence of compliance within the remaining period that may be obtained from substantive tests performed by the independent auditor or from tests performed by internal auditors, and (f) other matters the auditor considers relevant in the circumstances. (320.61)

The key point of these factors is the obtaining of evidence that the potentially reliable controls were in effective operation from the interim date through the end of the accounting period.

**FREQUENCY AND MONETARY MEASURES OF TRANSACTIONS CONTAINING PROCEDURAL DEVIATIONS**

**Planning Internal Control Reliability Tests**

In capsule form, the professional literature of both countries can be summarized as follows:

1. The auditor may choose to rely on internal control and thereby reduce the scope or change the nature or timing of his substantive tests.
2. Such reliance must be supported by compliance tests that yield satisfactory results.

Satisfactory results in this context could be interpreted subjectively.
if a judgemental approach to sampling the population were used. If statistical methods are used, explicit interpretation of satisfactory results in terms of the risk of sampling error is appropriate. The confidence level necessary for reliance on the internal controls is often established by policy, and precision is then allowed to vary depending on the audit circumstances (an euphemism for audit risks and materiality). This common choice is arbitrary since confidence and precision are inversely related as far as the statistical estimation is concerned.

Relating the precision of the estimate resulting from internal control reliability tests to materiality presents a problem. In planning the acceptable level of procedural deviations, the auditor must consider the relationship between the procedural deviations detected in the compliance tests and the monetary errors which might exist in the accounting records. Again the literature is somewhat sparse, although paragraph 19 of Appendix B of section 320 does provide some insight:

While procedural deviations increase the risk of material errors and irregularities in the accounting records, such errors and irregularities do not necessarily result from procedural deviations. For example, a disbursement that does not show evidence of required approval may nevertheless be a valid transaction that was properly recorded. Procedural deviations would result in errors or irregularities remaining undetected in the accounting records to be audited only if the procedural deviations and the errors or irregularities occurred on the same transactions. Consequently, procedural deviations of any given percentage ordinarily would not be expected to result in substantive errors or irregularities of the same magnitude in the accounting records. (our emphasis) (320.B.19)

Since procedural deviations are likely to occur at frequencies (and magnitudes) that are different from the frequencies and magnitudes of monetary errors in the accounting records, the auditor is forced to establish different decision rules for evaluating compliance and substantive tests. For substantive tests, the decision rules usually employ financial statement materiality. However, the relationship between materiality as it effects tests of compliance and materiality as it effects the financial statements is confused and not discussed in detail in the professional literature.

**Frequency Measures of Compliance Deviations**

The following three quotations from SAS 1 provide an example of the level of guidance that is provided in SAS 1 and an indication that the AICPA Committee contemplates that the auditor would evaluate
compliance tests in terms of the frequency of compliance deviations (i.e., the use of Attribute Sampling).

(a) The committee believes that samples taken for this purpose should be evaluated in terms of the frequency and nature of deviations from any procedures the auditor considers essential in his preliminary evaluation of internal control, and that their influence on his final evaluation of internal control should be based on his judgment as to the effect of such deviations on the risk of material errors in the financial statements... (320.A.22)

(b) The precision limits discussed in this paragraph for compliance tests relate only to deviations from pertinent procedures, which may or may not result in substantive errors in the accounting records (see paragraph .19)... and

(c) Based on consideration of the general matters discussed in paragraphs 19 through 21 and of the specific factors mentioned in this paragraph, an auditor may decide for example that an upper precision limit of 10% for compliance tests would be reasonable; if substantial reliance is to be placed upon the procedures, he may decide, for example, that a limit of 5 percent or possibly lower would be reasonable. (320.B.22)

At this time there are no professional pronouncements concerning "Precision and Reliability for Statistical Sampling in Auditing" in Canada.

Monetary Measures of Procedural Deviations

A recent text book, The External Audit (Anderson, 1977) is an attempt to shed new light on the purpose of compliance procedures:

The objective of a compliance test is not to search for monetary errors (the sample size would generally be insufficient to be sure of detecting these even if a material total were present) but rather to search for critical compliance deviations (likely to be present in greater quantities than actual monetary errors). (p. 322)

Anderson (1977, p. 230) explains his rationale as follows:

Since the relationship of monetary errors to materiality is the ultimate consideration, it is logical that the relationship between the frequency of critical compliance deviations and some multiple of materiality (say 3 or 4 times; at any rate, not 50 times) should be of significance. In this book a multiple of 3 times is suggested as a reasonable (and probably conservative) working rule for the normal case.

Consequently, Anderson's approach is somewhat different than that of the AICPA Committee since he argues that a measure of the dollar value of transactions containing procedural deviations is likely to be a better indicator of material errors in the population than a projected frequency of transactions containing procedural deviations. To illustrate the difference between the two approaches, con-
sider a sample of disbursement voucher files that is to be tested to identify the reliability of several different internal control procedures that leave "tracks" in the document file. Using the AICPA approach, the number of procedural deviations is divided by the sample size which in this case is the number of transactions. A sample of one hundred transactions where three transactions are missing required receiving reports results in a 3 percent error rate per transaction. The confidence interval statement would be, for example, that the upper error limit at 95 percent confidence is 7.6 percent. Using Anderson's approach in the same situation, the number of procedural deviations again would be divided by the sample size. However, in this case the sample size is the number of transaction dollars in the sample. The above example in this approach for a sample of one hundred transaction dollars would result in a 3 percent error rate per transaction dollar. The confidence interval statement in this case would be an upper error limit at 95 percent confidence of 7.6 percent per transaction dollar. Anderson would state this result as a 95 percent confidence that not more than 7.6 percent of the transaction dollars have been exposed to procedural deviations. Note that the sample of one hundred physical units would not be the same sample as the one hundred dollar units.

These two apparently divergent approaches to internal control reliability tests raise an important set of issues:

1. Is one method clearly superior and therefore, should be uniformly adopted in practice, or
2. Do audit circumstances determine the choice between the two methods either at the practice unit policy level or on the individual examination level?

A major purpose of the following sections of this paper is to address these issues.

A DESCRIPTION OF SAMPLE SELECTION METHODS
A basic component of the differences between the two approaches is the method of selecting the sample that is to be used for the compliance test. Essentially, they are:

(a) A random sample of transactions, and
(b) a value-oriented selection of transactions (normally with probabilities proportionate to the dollar value of the individual transaction).
Random Selection of Transactions

Simple random sampling of transactions in practice is usually made with assistance of special random number generating programs available on computer time-sharing systems. This enables random samples of physical units to be generated quickly and extremely inexpensively. When computer time sharing is not available, the auditor must use random number tables or an alternative selection method such as systematic selection. In these situations, the costs and effectiveness of the technique may change. Random selection of transactions for compliance tests ignores the dollar size of the transactions and clearly focuses on the application of internal controls uniformly across all transactions. If the size of the transaction is expected to influence the application of controls or the auditor has important substantive objectives, this approach may not be ideal.

Value-Oriented Selection of Transactions

Value-oriented selection may be much more cumbersome than a simple random selection when the complete transaction stream is not available on computer files. In these cases the auditor must either add the complete population of transactions processed during the period under audit or resort to using page totals or some other “subtotals” available in the population as a basis for making a selection. In any event, the selection costs are likely to be considerably higher than those incurred when a simple random sample is selected by utilizing a computer time-sharing package. A more serious concern is that this method of sampling may be less effective when internal controls vary with the size of the transaction (e.g., all purchases over $1,000 require the approval of the controller). It is possible both that specified controls are added and that comprehensive controls are more carefully applied to large dollar transactions. In this circumstance, the pps or value-oriented selection is biased against selection of the transactions most likely to contain procedural errors. If, as Anderson suggests, there is a strong link between the size of the transaction processed and the size of possible errors, the described problem may not matter. However, in an environment where the inverse relationship is expected and a value-oriented test of internal controls is used, we believe the auditor should carefully assess the degree of reliance placed on transaction system controls, especially when rolling forward interim balances to year-end and in testing smaller balances at year end.
Finally, since value-oriented selection methods require populations that are expressed in dollars, the method is not appropriate for testing all controls. For example, value-oriented selection methods would normally be appropriate for testing controls that ensure that all goods invoiced are shipped, but they may not be appropriate for testing controls that ensure that all "goods shipped are invoiced."

However, value-oriented selection does possess some distinct advantages such as making it easier to "sub-sample." For example, suppose each disbursement is supported by several invoices, that the auditor is interested in ensuring that all invoices are properly approved for payment, and that selection must be made from the cheque register. If a simple random sample is selected of cheques, then the auditor must either examine all invoices which support the disbursement selected or use rather complex multi-stage sampling methods. If, on the other hand, a value-oriented selection is made using the "dollar unit sampling technique," then the invoice containing the dollar selected would be examined for approval, not the other invoices contained as support for the disbursement. Consequently, the amount of audit work performed per sample item might be substantially reduced when a dollar unit sampling approach is used.

ANALYSIS OF SAMPLING PLANS
The following diagram summarizes the alternatives which might be used in a test of internal control reliability. Each of the four alternatives is summarized briefly in the following section and then those worth pursuing are considered in some detail in the following case.

Alternative 1
This alternative employs value-oriented selection with frequency estimation. Essentially this is the use of attribute sampling applied to a sample that is selected with probabilities proportionate to size. In our previous example, we obtained a result of a 95 percent upper error limit of 7.6 percent per transaction dollar. This method is not widely used in practice, but contains the same type of information as classical attributes estimation. Characteristics of this method that might be viewed as advantages include:

1. Larger transactions (in terms of dollar value) are emphasized.
2. A control reliance decision based on frequency of compliance errors rather than dollar magnitudes is facilitated.
<table>
<thead>
<tr>
<th>Selection Methods</th>
<th>Monetary</th>
<th>Alternative 2</th>
<th>Alternative 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
<td></td>
<td>Alternative 3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Value-oriented</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(probabilities</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>proportionate</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>to size)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Neutral (simple random)</td>
<td></td>
</tr>
</tbody>
</table>
3. When a dual purpose objective exists, this approach facilitates both an error rate and an error amount evaluation with reasonable sample sizes.

Characteristics of this method that might be viewed as disadvantages include:

1. It may be expensive to use due to the method of selection.
2. Sample sizes tend to become fixed from engagement to engagement and from year to year because the planning criteria are only indirectly related to materiality.
3. The method’s bias toward transactions may not meet the auditor’s objectives when internal control reliability is expected or known to be inversely related to transaction size.
4. The problems of relating the acceptable level of compliance deviations to materiality is present.
5. A population expressed in monetary amounts is required.

**Alternative 2**

This alternative combines value-oriented selection with estimation of magnitudes. In the above example, we were 95 percent confident that no more than 7.6 percent of the transaction dollars were exposed to internal control errors. If the dollar population was $1,000,000, then our upper limit would be $76,000. Note that the only difference between alternatives one and two is the way the evidence is stated (and, possibly, interpreted). Anderson has suggested that compliance tests should be planned on the basis using 80 percent confidence, precision equal to three times materiality and an expectation of detecting at least one compliance deviation.

Possible advantages of this method are:

1. The sample size calculation may be conveniently linked to the materiality level that is used in planning the substantive audit procedures. As a result, fixed sample sizes will not become prevalent and extent of compliance testing will differ from engagement to engagement and from year to year.
2. The selection method emphasizes large dollar value transactions which are usually of greater concern to the auditor.
3. The method can be interpreted using a procedural error rate approach.
4. This method shares with alternative 1 an efficient use of substantive evidence in the sample.
Possible disadvantages of this alternative are:

1. Since internal control systems often operate in different manners with respect to the value of transactions (e.g., dual signatures are required for all disbursements over $1,000, or special approval must be obtained for disbursements over a certain magnitude), then a selection which is based on probabilities proportionate to size tends to emphasize larger transactions and may not be as effective for observing internal control in the medium to small transaction categories.

2. The selection method used by this alternative is often difficult and expensive to apply in practice when the systems are either not computerized or utilize equipment that is not compatible with the software packages available to the auditor.

3. This method demands a selection from a recorded population and therefore it is not likely to be effective in testing internal controls over "completeness" or detecting compliance deviations that relate to "omission."

**Alternative 3**

This alternative combines simple random sampling with frequency estimation and is described in the auditing literature as attribute sampling. We believe attribute sampling to be the most common form of compliance testing at the present time when statistical sampling is used. As referred to earlier in this paper, upper precision limits are usually selected between 1 and 5 percent and in some cases up to 10 percent. Reliability levels usually range from 90 to 95 percent.

The possible advantages of this technique are:

1. The design and selection costs are minimal especially when computer time sharing is available to assist in the sample selection.

2. All transactions processed by the accounting system during the period under examination have an equal chance of being selected and therefore this plan may better suit the auditor's objective of observing the effectiveness of the control as it operates on all transactions regardless of dollar values.

The possible disadvantages are:

1. It is difficult to relate the frequency of compliance deviations that is tolerable to materiality in the financial statements. As a result, "rules-of-thumb" are widespread (e.g., 95 percent reliability with a 5 percent upper precision limit and a zero expected error rate). Consequently sample sizes tend to become fixed and not vary from en-
Engagement to engagement or change when financial statement materiality changes.

2. This procedure can be costly when the documents selected for examination are supported by several other documents.

3. When a transaction stream contains transactions that are processed by several different systems, computation of sample size and selection of the sample can be a significant problem.

Alternative 4

This alternative combines a magnitude measure of transactions coupled with simple random sampling. The magnitude measure could be either an estimate of the transaction dollars exposed to control errors or of the substantive errors in the transactions. We believe that only the former would be responsive to a compliance objectives. These methods are not common in practice to our knowledge and essentially would represent the application of variable estimation techniques for compliance deviations. Those methods would seem to have several limitations which include:

1. Controlling the precision to within even three times materiality would require a very large sample size if simple random were used.

2. These procedures would require the use of computer software in addition to computer time sharing and would therefore appear to be extremely expensive for use in practice.

We shall not discuss alternative 4 further. In the next section we demonstrate how alternatives 1, 2 and 3 compare in a hypothetical case study.

EXAMPLE OF VARIOUS ALTERNATIVES

Introduction

The following example is intended to illustrate the differences between three audit strategies that place reliance on internal control (i.e., a system-based approach) and three audit strategies that do not (i.e., a substantive approach). Although the example is extremely simple, we believe the principles developed are appropriate for more complex situations.

Background Information

The client is a construction company that is working on one major project which is a "cost plus" contract, and several other minor fixed
lee contracts. All expenses are charged either to the major project (Account A) or to a general project expenses control account (Account B).

Assumptions
This example assumes the following:
1. The feature of interest is the audit of the cash disbursements system.
2. The auditor’s planned level of overall assurance requires a confidence level of 95 percent.
3. The transaction population comprises about 2,200 small transactions (all transactions are less than $3,000), totalling $1,000,000.
4. At year end, the balance on Account A is $800,000 and on Account B is $200,000. Beginning balances were zero.
5. Materiality has been set at $10,000, based on a pretax income of $200,000.
6. No reliance is to be placed on analytical review procedures (a simplifying assumption).
7. The auditor’s preliminary evaluation of internal accounting control is good and the likelihood of monetary errors is considered to be low.

System Description
The client’s cash disbursements system is depicted in the flowchart on the opposite page.

Possible Audit Strategies
There are at least six possible audit strategies to obtain sufficient competent evidence on Accounts A and B.
1. Place reliance on internal control, perform compliance tests using attribute sampling with simple random selection, and substantive tests of balances.
2. Place reliance on internal control, perform compliance tests using attribute sampling with value-oriented selection and substantive tests of balances.
3. Place reliance on internal control, perform compliance tests using a value-oriented selection and substantive tests of transactions and balances.
4. Perform substantive tests of balances on Accounts A and B only.
Checks over $1,000 have to be counter-signed by Warren, the President.

De Account A or B
Cr Cash

- Zim prepares bank reconciliation and accounts for accuracy sequence of recorded checks.
5. Perform substantive test of transactions only.
6. Perform a combination of substantive tests of transactions and substantive tests of balances.

Note that in this simplified situation, direct tests of the balances do not differ in content from substantive tests of the transactions. Since in more complex situations they would be different in content, we will behave as though they are different here.

**Strategy 1**

The compliance tests are designed to use attribute sampling with a 95 percent reliability level, a zero expected error rate and an upper precision limit of 5 percent. We do not believe that a zero expected error rate is necessarily reasonable but use it as an illustration of a common practice. Based on an assumption of favorable results from the compliance tests, a 70 percent reliability level was used for the substantive tests.

The key element of control in this system is the checking activities performed by Joe and the review by Fred. To test compliance, a sample of 60 cheques was selected to:

(a) observe whether the initials by Joe and the signature by Fred (and Norton’s if appropriate) exist on the appropriate documents as evidence of compliance; and to
(b) decide whether the evidence of control (initials and signatures) is reliable by reperforming the various control activities.

If errors are not discovered in this test, the control will be considered effective and the substantive tests reduced.

**Strategy 2**

This strategy employs the use of the “rough rule-of-thumb sample size” formula proposed by Anderson [net population dollar value / materiality (in dollars) coupled with value-oriented selection]. Note that the sample subjectivity determined error rate used in the physical unit sample would have been used here.

**Strategy 3**

This strategy explicitly makes substantive use of the transaction system evidence obtained for internal control reliance purposes in Strategy 2. Note that the sample size does not change since the same sample is used for both types of objectives. The cost of testing the sample could increase under this strategy. Based on the expectation of
favorable results from the tests of transactions, a further reduction of substantive tests of balances was made by using a Reliability Factor of one (about 63 percent confidence level).

**Strategy 4**
In this case a reliability level of 95 percent was used for each account balance. Transactions were selected from the entries in each account and supporting documents examined to determine the validity and propriety of the accounting entries.

**Strategy 5**
In this case a reliability level of 95 percent was used for the transaction stream. Transactions were selected from the cash disbursements book; supporting documents and the posting in either Account A or B were checked to determine the validity and propriety of the accounting treatment.

**Strategy 6**
In this case a reliability level of 80 percent was used for the substantive tests of the transaction stream and 75 percent for the substantive tests of the account balances, so that the combined sampling risk of both the substantive tests of transactions (.20) and the substantive tests of balances (.25) failing to detect a material error would be limited to 5 percent (i.e., .2 × .25 = .05 or 5 percent).

**Sample Sizes**
The following table represents the sample sizes used in each of these strategies. The details of the tests, and the calculations of sample sizes, are presented in the last section of this example.

<table>
<thead>
<tr>
<th>Audit Test</th>
<th>Strategy 1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compliance Test of Transactions</td>
<td>60</td>
<td>100</td>
<td>100</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Substantive Test of Transactions</td>
<td>—</td>
<td>—</td>
<td>*</td>
<td>—</td>
<td>—</td>
<td>300 160</td>
</tr>
<tr>
<td>Substantive Tests of Balances Account A</td>
<td>96</td>
<td>96</td>
<td>80</td>
<td>240</td>
<td>—</td>
<td>112</td>
</tr>
<tr>
<td>Account B</td>
<td>24</td>
<td>24</td>
<td>20</td>
<td>60</td>
<td>—</td>
<td>28</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>180</td>
<td>220</td>
<td>200</td>
<td>300</td>
<td>300</td>
<td>300</td>
</tr>
</tbody>
</table>

*Same sample used for substantive objective*
It should be noted that if it became evident that the controls were not effective and monetary errors were likely, then the sample sizes for the substantive tests would increase, based on the amount of monetary error expected. A great deal of care is necessary in viewing the changes in sample size across the strategies. Ideally we would have held overall audit risk constant across all strategies. We do not know enough about the differences between some of the strategies (such as numbers one and two) to accomplish this. Therefore, the reader should view the illustration as a rough sketch of the differences and carefully consider the relevant information going into the decisions. For example, the implication that Strategy 1 results in smaller sample sizes is probably an artifact of our decision to use a "common practice."

Sample Size Calculations
Sample sizes for the substantive tests were calculated using the following formula for Dollar Unit Sampling.

\[
\text{Sample Size} = \frac{\text{Population Value} \times \text{Reliability Factor}}{\text{Materiality}}
\]

<table>
<thead>
<tr>
<th>Reliability Level</th>
<th>Reliability Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>70</td>
<td>1.2</td>
</tr>
<tr>
<td>75</td>
<td>1.4</td>
</tr>
<tr>
<td>80</td>
<td>1.6</td>
</tr>
<tr>
<td>95</td>
<td>3.0</td>
</tr>
</tbody>
</table>

In all cases, a materiality of $10,000 was used. Samples size for the attribute sample was calculated using conventional binomial tables.

A Summary of the Case Illustration
This example illustrates that:
1. The amount of audit verification in total may be reduced by placing reliance on internal control.
2. The only difference between Strategies 1 and 2 is the method of selection and the criteria for establishing precision.
3. Strategies 2 and 3 employ sample size calculations that directly utilize materiality which in this case produces a sample size larger than the conventional attribute sampling method.
4. Since Strategies 4, 5 and 6 do not place reliance on internal control, the substantive tests must be designed to obtain reasonable assurance.
(in this case 95 percent) that the amount of error does not exceed materiality.

5. The use of substantive tests of transactions may be useful in shifting some work from the year end to an interim date (hence reducing the amount of year-end procedures).

6. Strategies 4, 5 and 6 treat the accounting system as a “black box” and effectively ignore it. In the Canadian computer auditing literature, this has been called a “data-oriented approach” to distinguish it from the “systems approach.”

7. The choice between Strategies 4, 5 and 6 must be based on the ease of performing the test and the quality of the expected evidence, since the extent does not change. The only difference is the “direction of the tests,” not the amount of work involved.

CONCLUSIONS
A number of issues have been raised in our analysis regarding the auditor’s reliance on and testing of internal controls as a part of tests of transactions. In our summary we would like to focus on the issues pointed out previously as to (1) whether a physical unit or value unit approach should always be chosen or (2) whether there are criteria that should lead an auditor to choose one method in some circumstances and another method in other circumstances. It is our view that the quality of evidence on the reliability of internal controls does not vary between the two methods. An error rate per transaction and an error rate per transaction dollar convey essentially the same message: either the control is operating with an acceptable rate of frequency or it is not. The decision parameters of confidence level and precision differ only in the definition of precision (the acceptable upper error limit or rate of procedural deviation). It seems to us that in terms of assessing the reliable operation of an internal control that it does not matter whether or not this rate is stated in terms of deviations per transaction or deviations per transaction dollar. The argument that the latter is conceptually easier to relate to materiality is really a matter of tastes and preferences rather than concept. In both cases it is still necessary to recognize that a procedural deviation does not cause a monetary error. Anderson’s smoke and fire ratio subsumes this point by assuming that procedural deviations are (approximately) equally likely across all transaction dollars (regardless of the dollar size of the transaction) and that one dollar of monetary error will occur for every three transaction dollars exposed to monetary error.
These assumptions are really empirical questions that should be tested. We suggest that the approach is very reasonable in many cases, but that in some audit circumstances an auditor may wish to choose other approaches.

An important practical advantage of the value-oriented approach to internal control reliability tests is that the resulting sample can also be used efficiently for a substantive test objective in the transactions population. Note that in the above example in Strategy 3, the same sample was used for both the compliance and the substantive objective. This approach is not as feasible when a physical unit sample is selected because of the impact of the variability of the dollar sizes of the transactions on the traditional variables sampling estimators. Typically, the needed substantive sample size would be a very large multiple of the compliance sample size. We suspect that many practice units use the AICPA approach to tests of transactions even when a substantive objective exists. Their evaluation of the substantive evidence in this setting would have to be subjective.

Our analysis should not be viewed as complete. However, based on what we have accomplished to this point, the following recommendations are reasonable:

1. A value-oriented approach to tests of transactions is reasonable and may be the most efficient choice in the presence of dual purpose objective biased toward the monetary error end of the spectrum.
2. A value-oriented approach is not clearly superior in all audit circumstances. The auditor should recognize the cost and benefits of using each approach in the context of the risks and record-keeping characteristics of a specific client setting.

REFERENCES


I am happy to be here to discuss the paper by Bill Felix and Jim Goodfellow. I would also like to thank the Price Waterhouse Foundation for making this conference possible. First, I should explain my orientation and sources of possible bias on this issue. We have extensively used random selection for attributes testing for compliance. We are just beginning to use PPS for substantive testing (under fairly strict guidelines). We have not used PPS for compliance testing. As a second source of possible bias, Bill Felix is consulting with us on the implementation of PPS for substantive testing. I shall try not to let these sources of possible bias influence my discussion.

It is difficult to discuss a paper which concludes that either a value-oriented approach or random approach is reasonable and that one is not clearly superior to the other. Although I would like to agree with the conclusion, I believe that the conclusion has not been proven.

I should like to discuss certain points in the paper and indicate my thoughts on them. First, the paper indicates that the Canadian literature acknowledges that the auditor might recheck calculations or procedures which support signatures, initials or audit stamps, whereas the AICPA pronouncement is silent and might imply that such rechecking is not necessary. In practice, rechecking is usually performed. The auditor cannot evaluate the propriety of performance unless he is satisfied that the procedure was performed correctly. While some would think of the correctness as a substantive test, I view
it as part of the compliance test. Just looking at initials is a meaningless procedure unless the auditor is satisfied that the procedure was actually performed. This is an area where U.S. professional literature needs clarification.

Page 88 indicates that both countries indicate a preference for some sort of "representative" sampling from the transactions executed throughout the period under audit. The authors indicate that a representative sample is best obtained by random selection. It should first be noted that a random (or representative) selection is not called for by the SAS, but only that tests be applied to transactions throughout the period. Thus, for example, a sample of five transactions per month, selected in a nonrepresentative manner, would meet the criteria of the SAS. I believe that in some firms (probably not the larger firms) there exists a great deal of block sampling for compliance testing. In some cases, only one or two blocks are chosen, i.e., the "test week," the "test month," or even the "test day." Further, even in the large firms, when it comes to testing payroll transactions, the one week payroll test is still quite common. At least in the United States, and especially in the smaller firms, the word "ideally" has given this paragraph a meaning of "not required." Thus, although SAS 1, Sec. 820.61 can be read to state a preference for some sort of representative sampling, there are many practitioners who do not read it that way or who do not use random selections for compliance testing.

The paper then goes on to discuss the use of frequency and monetary measures of transactions containing procedural deviations. The paper indicates quite correctly that the outlook of SAS 1 is towards attributes sampling for compliance testing. However, procedural deviations of any given percentage ordinarily would not be expected to result in substantive errors or irregularities of the same magnitude in the accounting records. Reasoning from this, one can state that if the objective is to determine the error rate in dollars, there is no benefit in knowing the error rate in physical units. Or, put another way, if you need the error rate in dollars, you should not perform a compliance test; you should perform a substantive test. Or, put a third way, since the auditor cannot reason directly from compliance error rates to dollar error rates, there is really no reason to ever perform compliance testing. It would be very useful for some research to be done to prove whether good internal control leads to correct financial statements (before audit adjustments). For example, one could take samples of financial statements before and after adjustments and correlate them with internal control evaluations or management
comments and determine whether good internal control really leads to good financial statements.

The paper then discusses each sampling method and how it works. In practice, random selection of transactions is an efficient method. This is typically because timesharing programs are almost always used. A PPS selection of transactions tends to be difficult to implement because one must add through the file or perform other clerical methods throughout the audit period. Even if computers are available, it means all the tapes for the year must be retained by the client and run through the audit software package.

I disagree with the conclusion that the value-oriented selection possesses advantages when it is necessary to subsample. When the auditor needs to sample disbursements attached to vouchers, he can make a random selection of invoices either by using as the sampling unit line items on an invoice register or invoice edit runs or batch control runs.

The real question is which transactions are more likely to contain the type of error that the auditor is interested in finding. In most cases, the auditor does not know. If he did, he could design his test to attack those problem areas. Thus, he should first analyze the system of internal control and determine where, if anywhere, errors are likely. However, if errors were likely, he might not perform compliance testing but substantive testing. Since most good systems should generate few errors, the question of likelihood is difficult to answer. However, with no clear answer as to where errors are likely, the random selection seems just as good as any other selection.

The paper then discusses the advantages and disadvantages of each of the four alternatives. A major disadvantage of alternatives 1 and 2 is mentioned in item 3 on page 96 but not in any detail. This disadvantage results because the PPS method tests only for compliance deviations related to overstatements. It has less chance of finding compliance deviations which cause understatements. For example, if there were a pricing error on a sales invoice which caused the invoice to be billed for $1 instead of $10,000, and pricing was inadvertent not checked by the client, the auditor would have very little chance of finding it. On the other hand, he would have a better chance of finding the error in a random selection.

What we really obtain with a PPS sample is not a statement that the upper limit of compliance error is 7.6 percent of the population dollars. Rather, we obtain a statement that the upper bound of
compliance errors which could result in overstatement is 7.6 percent. Thus, every PPS selection should contain tests that the population is not understated. These tests may be from another population such as shipping documents or may be based on a random selection or an analytical review. This additional selection increases the cost of a PPS sample (and should have been considered in the author’s analysis of required sample sizes on page 101). On the other hand, if a random selection is made, it is still necessary to see that the population is complete. But this can simply be done by testing the beginning and ending check numbers and using the sample itself to see that numerical control was maintained.

Page 96 discusses the disadvantages of the random selection method. The second disadvantage is that random selection can be costly when the documents selected for examination are supported by several other documents. This only results when the random selection is made from the wrong documents. For example, if a voucher is supported by ten invoices, the sampling unit should be the invoice, not the voucher. If the invoices are on the computer file, it is easy to make a selection of invoices. Also, a PPS selection will require manual procedures to calculate the proper invoice for selection. I assume this disadvantage is not referring to the situation where a voucher is supported by an invoice, a receiving report, a purchase order and a requisition. In that case, both samples would have to look at all of the documents if the error being tested required that all of the documents be present.

The third disadvantage of random selection given is that when a transaction stream contains transactions that are processed by several different systems, computation of sample size and selection of the sample can be a significant problem. I do not see computation of sample size to be difficult because that is still a simple attribute computation and the population size is usually not relevant. What is important is that the auditor correctly define the population of interest. For example, assume that a cash disbursement system is used to make payments for purchases, payroll, and other expenses. The other expenses are subject to different approval procedures than purchases and payroll. In this situation, the auditor would have to make a decision as to whether he was testing compliance with one system or three. This has nothing to do with the question of whether a PPS or random sample is used. If he thought that there were three systems because he is relying on the controls for payroll, purchasing and
expenses differently, he might make three selections using either PPS or random selection. For a random selection, he would use pages and lines in an appropriate register rather than the check numbers since the check numbers are all in sequence. Alternatively, he might use an overall random selection of checks, picking enough extra random numbers to get appropriate sample for each group. Note also that the sample size for each group might not be the same.

Alternative 4 involves a direct application of variables estimation technique. As the authors state, it is rarely used for compliance testing; however, in some applications it may be necessary to switch from an attributes to a variables measurement. For example, assume that in testing compliance with perpetual inventory records some errors are found. The auditor may wish to evaluate the magnitude of the errors. Of course, one would rarely go the other way. Having concluded that records are not materially in error, it is rarely necessary to make an inference about the system.

Page 97 presents an example of various alternatives. While I do not quarrel too much with the example, it may be too simple for practical purposes. That is, the results of the transactions can be easily traced to two accounts and the results clearly related to materiality for those two accounts. In practice, transactions involve many accounts and errors in compliance found in one account may imply monetary error in different accounts. Also, the example does not consider the costs of selection. Nor does it consider the additional costs which are required to test the defined population for understatement by selecting from another ("reciprocal") population. Since a PPS sample is often more costly to obtain, the auditor might choose a large random selection rather than a PSS selection (if he knew they both provided equivalent evidence). The paper assumes that each of the six strategies provide equivalent audit evidence. This assumption is really what the authors are trying to prove by the paper.

Strategy 1 says perform a compliance test at 95 percent reliability, zero expected rate of occurrence, and upper precision limit of 5 percent. While the authors note that the zero expected error rate may not be reasonable, it is a common practice in some firms and illustrates the dangers of applying statistical sampling by rote. To me, the zero expected rate of occurrence does not seem to be a good strategy because the chance of failure is quite high. With this plan, finding one error would cause the auditor not to rely on the system. Since this system is a manual system, controlled by just a few people, it is quite
likely that that auditor would find one error. In fact, the actual error rate would have to be much lower than 5 percent for the auditor to accept this population most of the time. The following table from the January 1976 article entitled “Statistical Sampling and Internal Control” by Dr. Arkin in The CPA Journal contains the probability of finding one or more errors in a sample of sixty.

<table>
<thead>
<tr>
<th>Actual Population Error Rate</th>
<th>Probability of Some Errors in Sample of 60</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>0.5</td>
<td>26.0</td>
</tr>
<tr>
<td>1.0</td>
<td>45.3</td>
</tr>
<tr>
<td>1.5</td>
<td>59.6</td>
</tr>
<tr>
<td>2.0</td>
<td>70.2</td>
</tr>
<tr>
<td>2.5</td>
<td>78.1</td>
</tr>
<tr>
<td>3.0</td>
<td>83.9</td>
</tr>
<tr>
<td>3.5</td>
<td>88.2</td>
</tr>
<tr>
<td>4.0</td>
<td>91.4</td>
</tr>
<tr>
<td>4.5</td>
<td>94.7</td>
</tr>
</tbody>
</table>

Many auditors would not wish to rely on a system such as this because of the possibility of management override (there are only two accounting people involved plus the president). It would take only a few items to be shifted between Account A and Account B to seriously impact the earning of the company. Thus, it is hard to justify a 70 percent reliability level for substantive tests, even if the “paper processing” appeared to be reasonable. Because of this possibility of management override, it would seem that a strategy which relies heavily on substantive testing would be most appropriate. Thus, I would probably prefer Strategy 4 for testing each balance account or perhaps the sum of the two balance accounts as one population. Whether this should be done as a PPS sample or a random sample (for ratio and difference estimation) depends on the error rate expected. The auditor should also test that transactions go only to these two accounts and this can be done by review of the general ledger. Strategy 5 also makes some sense since it tells the auditor that transactions are properly recorded. He would also perform an analytical review and a review for collectability of the accounts.
The other strategies make much less sense. Strategies 2 and 3 require a reliance on internal control which may not be appropriate because of the possibility of management override. Strategy 6 seems to be duplication in that one should test either transactions or balances but not both since one leads to the other.

Strategies 2, 3 and 6 seem to be geared to finding no errors, yet it is probable in a system such as this that some errors will be found. The formulas used to calculate sample size should consider the expected rate of occurrence. Also, if the auditor thinks adjustments will be required, he should use an approach which will permit an adjustment. Probably Strategy 4 with a random or stratified selection will be useful for producing an adjustment. Strategies 3 and 6 are not likely to produce useful adjustments.

The authors' conclusions are stated on pages 103 and 104. They conclude that an error rate per transaction and an error rate per transaction dollar convey essentially the same message. Since these items have different statistical meanings, they should present different meanings to auditors. The auditor should be clear as to what he needs in a given situation and why. Auditors often perform tests which they call compliance tests, which are really substantive tests in disguise. The example presented by the authors is typical. If what the auditor is really doing is a disguised substantive test, certainly a value-oriented approach is reasonable, especially when audit software is available (using a PPS approach if few errors are expected and a stratified approach if more than a few errors are expected). If what the auditor is doing is a true compliance test, then I would think a random selection is just as good, if not superior.

Unfortunately, the recommendations on page 104 are not fully supported by this paper. If one assumes the exhibit on page 101 to be a choice of equivalently effective strategies (and forgetting the above analysis), then Strategy 1 appears most efficient. While the auditor should go through this sample size analysis, it is also important to figure out which method provides the best evidence. The paper has addressed the efficiency issues; the paper does not answer the effectiveness question.

This paper did not intend to answer all questions but to identify some problems for the profession. Clearly, it has fulfilled that role. It is hoped that the AICPA will consider the problems raised by this paper and provide practical guidelines.
Discussant’s Response to “Audit Tests For Internal Control Reliance”

WILLIAM R. SCOTT

It seems clearly established in auditing theory and practice that the better the state of internal control in the client’s accounting system, the more the auditor may reduce the extent of his substantive tests.

This raises at least three subsidiary issues, however, should the auditor decide to rely on internal control. First, how can the state of internal control be measured? Second, what reliability and precision limits should be required for compliance tests? Third, how do the compliance results affect the substantive tests?

The professional auditing literature leaves these matters largely to the auditor’s judgement. Anderson (1977, p. 316), for example, suggests confidence for substantive tests of as low as 80 percent if internal control risk is “low or very low” and up to 99 percent if it is “high.” But individual judgements may differ. Even if they do not, an auditor who, say, is arguing with a client with a poor internal control system over the extent of his substantive tests, or perhaps defending himself in court, will be in a stronger position, it seems to me, to the extent he can demonstrate analytically rather than judgementally, how he measured the internal control risk as “high” or “low” and how this affected his substantive work. I expect most auditors at this point would say “amen” but then point out that a workable, rigorous link between the quality of internal control and the probability of material error in the final balances has not yet been, and may never be, developed. I would have to agree, at least with the “not yet” part, and I
do not want to argue that auditor's judgement should be dispensed with. Rather, I think the questions are the extent to which statistical techniques can assist the auditor's judgement and, as I will come back to later, on how this judgement can best be brought to bear.

Consequently, I think the Felix and Goodfellow paper addresses an important set of issues, since it deals with the choice of statistical sampling techniques for internal control evaluation.

To return to the three issues I mentioned above, the paper says little about the third issue, other than illustrating audit strategies that do and do not rely on internal control. I would like to comment on how it deals with the first two issues, however.

With respect to the first, Felix and Goodfellow suggest in Part B that internal control quality can be measured in terms of an error rate per transaction or an error rate per transaction dollar. They equate the error rate per transaction approach with the AICPA, SAS 1, and the error rate per transaction dollar approach with Anderson, and indicate that a major purpose of the paper is to address whether or not one or the other of these two approaches dominates. Their conclusion in Part F is that an "error rate per transaction and an error rate per transaction dollar convey essentially the same message." I agree. But I would also have agreed back in Part B. I do not see any real difference since it should always be possible to convert from one to the other. I am puzzled as to what it is in their intervening discussion that leads them to this conclusion.

The major difference between their "frequency" and the "monetary" approaches, however, is not whether the error rate is measured in transactions or dollars. Rather, it is whether the sample should be chosen to give judgementally specified confidence and precision or whether it should be tied to materiality through some rule or model linking compliance errors with final balance errors. I think this could have been made clearer in the paper.

Under this interpretation, however, the choice of selection method (pps or simple random) is invariant with respect to the choice between the "frequency" and "monetary" approaches. This is why there is such an overlap between the pros and cons of the four alternatives discussed in Part D. I think that the cross-classification of Part D could usefully be eliminated, with the discussion (perhaps expanded) going directly to the strategies in Part E.

Also, I think that a comparative discussion of the statistical characteristics of the estimators of the various strategies should be given.
Then, the practical pro and con judgement factors which they list can be interpreted with these statistical considerations in mind—I would be more willing to accept a strategy that was, say, expensive to use if I knew its estimator was very efficient, for example.

With respect to the second issue, of reliability and precision limits, I will discuss this in the context of the six possible audit strategies in Part E.

I like the suggestion, which I believe comes from Anderson, that the upper error rate limit be related to materiality through a "smoke to fire" ratio. Thus, if it takes $3 of compliance errors to produce $1 of final balance errors, the upper error rate limit is that rate that if applied to the transaction dollar population will yield a total error of three times materiality. Provided a reasonable estimate of the smoke to fire ratio can be obtained empirically, this reduces by one the number of required judgement inputs into the audit.

Let me apply this smoke to fire approach to their Strategy 1 example. If the smoke to fire ratio is 3:1, then a 5 percent error rate per transaction dollar will just produce a material error ($10,000) in the final balance. Their sample size of sixty was selected to provide a 85 percent reliability that if the true error rate was 5 percent or more the sample would have picked up one or more compliance errors. From this it follows that there is about an 81 percent reliability that if the true error rate was 3 percent or more the sample would have picked up one or more errors. Thus, assuming the sample reveals zero errors, there is a 19 percent risk that this population contains errors greater than materiality.

A 70 percent reliability level was used for the substantive tests, implying a 30 percent risk of not picking up a material error. If we define the audit risk as the product of the compliance and substantive risks, we get .19 x .30 = .057 as audit risk for Strategy 1.

Now look at Strategy 2. Clearly the audit risk will be lower since the compliance sample size is larger with no change in the substantive sample. In fact, the audit risk here works out to .0148.

What I am saying, then, is that I cannot compare these strategies because they are not standardized for audit risk. Either the audit risk should be held constant and the audit sizes compared or the audit sizes should be held constant and the audit risks compared before useful comparisons can be made.

I have two other comments on these strategies. First, I assume the control procedure in Strategies 3, 5 and 6 would be reperformed by the
auditor but I do not know about Strategies 1 or 2. Perhaps this could be clarified. Second, the authors state that the amount of substantive testing must be the same in Strategies 4, 5 and 6. Their example has the same dollar total for transactions and final balances and it is not clear to me that this statement would continue to hold if these were not equal.

Let me now make two final and more general points. First, I think Felix and Goodfellow might usefully have referenced some of the academic (as opposed to professional) literature in this area. As I read the paper, I found myself searching for a model which might be useful in pinning down more precisely the process of give and take between internal control evaluation and substantive tests.

An example of the sort of model I have in mind is one of Bill Kinney's (1975). There are a host of modelling difficulties in this area, of course, and I do not propose to get into these here. What does strike me about Kinney's model is how it provides a focus for the auditor's subjective judgements about system reliability and compliance rate—given these the rest follows analytically. When I compare this with the bewildering variety of judgment calls in the more conventional procedure as outlined in Felix and Goodfellow, I think it is an open question as to which method is potentially more defensible.

Second, and finally, I think that auditors should consider statistical sampling of the internal control system in a multivariate context. There may be problems in observing a vector of transactions from different systems simultaneously, but to the extent this can be done there could be benefits. One possible benefit would be the applicability of more powerful estimation techniques, such as Stein estimation (Ijiri and Leitch 1978). Another would be the information content of the covariance structure of the errors in the various systems. A few years ago, I did some work with a multivariate audit model. I had expected that covariance between errors in different systems would reduce audit size because, in effect, more would be learned from each individual transaction observation. What I found was that positive covariance tended to increase audit size and negative covariance to decrease it. The reason turned out to be that positive covariance made it more likely that errors in the systems would be of the same sign, and of opposite sign if negative covariance. Since errors of the same sign combine in their impact on net income while errors of opposite sign cancel out, the audit size was responding to the greater risk of overall material error that positive covariance creates. I cite this
just as an example of the sort of insight that can be gained from a multivariate model but not from a univariate one.

In summary, I think the Felix and Goodfellow paper addresses some interesting and important issues. Its main usefulness is in illustrating and discussing a variety of routes to the auditor's final opinion. However, I do not think it tells us very much about the conditions under which one or the other of these routes should be followed. Whether it could, or even should, tell us this seems to involve some basic questions about the tradeoff between judgement and formal analysis in the audit process.

REFERENCES

Session

FOUR
"Subject To" Audit Opinions: A Preliminary Investigation of Statement Users' and Statement Issuers' Perceptions

JOHN K. SHANK and JESSE F. DILLARD

The primary product of an audit by an independent certified public accountant is an opinion regarding the financial statements being audited. The opinion has many forms, depending on the message the CPA wants to convey about the statements.

In recent years there has been a large increase in the number of audit opinions which involve a qualification ("subject to") regarding some major uncertainty facing the firm. In 1967, in the sample of 600 firms surveyed each year by Accounting Trends and Techniques, twenty-one "subject to" opinions were noted (3.5 percent). By comparison, seventy-six were noted (12.7 percent) in 1975. For the approximately 9500 firms which must file a 10K Report each year, over 1400 uncertainty qualifications (15 percent) were issued in 1975. This represents nearly a four-fold increase in the number of "subject to" opinions between 1967 and 1975. To some extent, this increase could be due to an increasing level of business complexity and uncertainty facing publicly-held firms. However, a recent survey conducted by The Corporate Communications Report** (July, 1977) indicated five other reasons for the rapid increase in uncertainty qualifications:

*The authors wish to thank the Peat Marwick Mitchell Foundation for its financial support which made this study possible.

**The Corporate Communication Report is a bi-monthly investor relations newsletter published by Corpcom, Inc.
1. A rise in private and regulatory litigation against companies, which has created a new wave of corporate uncertainty;
2. An increasing reluctance by attorneys to state that their clients have good defenses against major lawsuits;
3. An increased vulnerability of accounts to litigation and their resulting desire to take protective steps;
4. The inability of the accounting and legal professions to come to a satisfactory agreement as to how lawyers can communicate to auditors the existence, status and likely outcome of a major lawsuit against a client company; and
5. SEC pressure for accountants to be more “hardnosed” toward their clients.

We share the views expressed in this survey that the last decade has exhibited a rising concern within the public accounting profession about being particularly careful that the financial statements, footnotes, and related audit opinion communicate as fully as possible the major uncertainties faced by the firm. A “subject to” opinion, in principle, does imply a higher level of uncertainty than does an unqualified opinion. Also, the use of “subject to” opinions implies that the responsibility for communicating this uncertainty is shared by management and the auditor.

There is no systematic empirical evidence of which we are aware regarding the extent to which issuers or readers of financial statements do interpret a “subject to” opinion to convey a message which is significantly different from that in an unqualified opinion. The efficacy of the significant increase in the use of “subject to” opinions depends on an untested assumption that statement readers do interpret such opinions as useful messages from the auditor which convey significantly more uncertainty about a firm’s financial position.

The AICPA’s Auditing Standards Executive Committee (AudSEC) tentatively concluded recently (AudSEC, 1977) that the “subject to” opinion does not represent a valid extension of the attest function. Following a recommendation by the CAR (AICPA Commission on Auditor’s Responsibilities, 1978), AudSEC proposed that “subject to” qualifications be eliminated. Although their exposure draft has now been tabled pending a broader review of the reporting responsibilities of management and auditors, AudSEC has not withdrawn its recommendation. Although they presented no systematic empirical evidence to support their view, the CAR concluded that “subject to” opinions can be confusing to statement readers and are not always consistently used by auditors. Since the auditor cannot evaluate the contingency, the Commission argued that the qualification in the
opinion provides no valid information beyond what should be conveyed in the footnotes. Apparently, a majority of the members of AudSEC concur in this view.

There is some evidence to support the belief that statement readers do not fully understand the different messages which an auditor can present. For example, public accountants have been forced to stop the practice of issuing unaudited statements on CPA stationery because many readers, including some bankers, were not sensitive to the difference between audited and unaudited statements. On the other hand, Libby (1977) found that bankers' and auditors' perceptions of the similarities among various forms of the audit opinion are not significantly different. In a capital market setting, investors seem to attach as much significance to unaudited quarterly earnings reports (May, 1971) as they do to audited annual earnings reports (Ball and Brown, 1968). May (1971) cited this observation as evidence that investors as a whole may not really be sensitive to the extra assurances represented by an audit opinion. Kaplan (1975) argued instead that this evidence suggests that an audit may have only marginal significance in changing investor expectations. In a somewhat related vein, Baskin (1972) found no significant security market reaction to the message conveyed by a consistency exception in the auditor's report. Also, Estes and Reiner (1977) found that an "except for" opinion regarding a violation of GAAP did not change bank lending officers' decisions on a loan application.

This study provides some preliminary empirical evidence concerning statement issuers' and statement users' perceptions of and attitudes "subject to" opinions. Specifically, we investigated how a sample of corporate financial executives and financial analysts believe a specific set of contingencies should be disclosed in the annual report and how they believe the contingencies would be disclosed. We think that the views of these groups, and others, constitute important evidence in evaluating the role "subject to" opinions do play in financial reporting and the role they should play.

RESEARCH DESIGN

The Instrument
The research instrument used was a self-administered questionnaire, attached here as Appendix A. The questionnaire consists of eight caselets representing actual contingencies selected from an extensive
review of published financial statements and the Disclosure Journal. The cases chosen span the following sets of issues:

(a) Four caselets involving an asset realization issue and four involving a litigation issue;
(b) Four caselets involving a contingency of "going concern" magnitude and four involving contingencies of major, but not drastic, proportions; and
(c) Four caselets in which a "subject to" opinion was actually given by the auditor and four in which an unqualified opinion was given. Although this can be thought of as a 2 x 2 x 2 repeated-measures design, an ANOVA model is not used to evaluate the results since no experimental "treatment" was administered. Exhibit 1 represents an overview of the issues incorporated in the questionnaire, with each specific corresponding caselet.

<table>
<thead>
<tr>
<th>Exhibit 1. Classification of Questionnaire Caselets</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Scope of the</strong></td>
</tr>
<tr>
<td><strong>contingency</strong></td>
</tr>
<tr>
<td><strong>in the annual report</strong></td>
</tr>
<tr>
<td>Going-Concern Major, but not drastic Case III</td>
</tr>
<tr>
<td>Case I-A</td>
</tr>
</tbody>
</table>

For each case, the questionnaire included the following five disclosure alternatives:

(1) No specific reference to this situation in the annual report.
(2) Disclosure only in the unaudited section of the annual report.
(3) Disclosure in a footnote to the financial statements, but no further mention.
(4) Disclosure in the footnotes, plus a three-paragraph "subject to" audit report in which the middle paragraph provides an explanation of the contingency and its anticipated resolution.
(5) A "disclaimer of opinion" by the auditor, along with the auditor's explanation for the disclaimer.

The third disclosure alternative above represents the recommenda-
tion of the CAR that major contingencies be disclosed by management in a separate footnote to the financial statements, like the accounting policies footnote, but not mentioned in the auditor's report.

The fourth disclosure alternative is designed to reflect current practice regarding contingency disclosure. In this regard, a question we faced was whether or not to include the last phrase, "and its anticipated resolution." If this statement, "and its anticipated resolution," is interpreted as the auditor's own forecast of the resolution of the contingency, the questionnaire results are clearly biased. There is little doubt that respondents would like someone to resolve business risks for them, if at all possible. However, after considering this problem, we chose to include the phrase in question anyway because current practice does usually include a reference to management's view regarding the resolution of the contingency. Consider, for example, the following paragraph taken from the "subject to" opinion in the 1974 Lockheed Annual Report.

The Company's studies indicate that the carrying value of its L-1011 TriStar inventories will be recovered and that gross profit will be realized over the remainder of the program. As discussed in Note 5, future sales and cost of sales of the L-1011 TriStar program will be affected by a number of factors, the effects of which have been estimated by the Company in accounting for the program. We believe these estimates are reasonable; however, because of uncertainties inherent in such estimates, the ultimate impact of the factors referred to above cannot be presently determined.

Such wording is typical over the hundreds of audit opinions we read in designing the questionnaire. We thus do not believe that the disclosure option corresponding to current practice would be fairly presented without this concluding phrase. Whether, in fact, such a phrase does encourage the statement reader to infer more about the auditor's own conclusions than the auditor explicitly states is an interesting question which is beyond the scope of this study.

Since federal law does not permit a disclaimer of opinion for SEC registrants, the fifth disclosure option is rarely seen in practice. It is tantamount to the auditor's resignation from the engagement. Even though it is used only rarely, we included it in the set for purposes of completeness.

The respondents were asked to rate each of the caselets in terms of: (1) the disclosure level they feel should be chosen to best communicate the contingency described ("subject to"), and (2) the disclosure
level they believe the company’s annual report actually would contain (‘would be’).

In evaluating each situation, the respondents were told to assume that the company’s financial statements had been audited by an independent CPA in accordance with generally accepted auditing standards; and, that, except for the specific material contingencies described in the cases, the financial statements fairly presented the financial position of the company, in accordance with generally accepted accounting principles. Further, whenever “disclosure” was mentioned, the respondents were instructed to assume that the description was full and forthright.

The problem of the auditor’s legal liability was also addressed in the questionnaire. The respondents were asked to rate the extent to which they felt the issuance of a “subject to” opinion should affect the auditor’s defensive posture if the contingency were ultimately resolved unfavorably for the company and a lawsuit resulted. One asset realization case and one litigation case were included in this test. The ratings for this question were made on a five-point “verbally anchored” scale. The “verbal anchors” were: (1) significantly adverse effect; (2) moderately adverse effect; (3) no effect; (4) moderately favorable effect, and (5) significantly favorable effect.

Pretest

A preliminary draft of the questionnaire was administered in person to three participants: a bank trust officer, a brokerage firm account executive, and a CPA. Each case was discussed with the respondents in an attempt to clarify confusing statements and determine the adequacy of the information presented. One respondent was interviewed while completing the questionnaire. The other two respondents were interviewed after their completion of the questionnaire. The questionnaire was then revised based upon our findings.

The revised questionnaire was administered to five CPAs and five bank loan officers in a second pre-test. Although direct interviews were not feasible, the participants were asked to note any problems encountered, as well as the time each required to complete the questionnaire. This version was also circulated within the research group at the executive office of a major public accounting firm. Appropriate final modifications were made based on these pre-test results. The estimated completion time of the revised questionnaire was approximately twenty-five minutes.
The Sample
Two groups were chosen for this study: corporate financial executives and chartered financial analysts. As a representative group of statement issuers, we decided to survey the top financial officers in major corporations. The actual sample was selected from U.S. firms which are among the top 500 in terms of either total assets, total stock market value, total sales, or net income. A listing of these firms was obtained from Forbes, May 15, 1976, "The Dimensions of American Business: A Roster of the U.S.'s Biggest Corporations." This listing was alphabetical and containing 823 companies. A random sample of 500 companies was chosen. The name and address of the chief financial executive for each of these 500 firms was obtained from Standard and Poor's Million Dollar Directory. As a representative group of knowledgeable financial statement users, we decided to survey experienced financial analysts. The Institute of Chartered Financial Analysts supplied their mailing list from which we selected a random sample of 800 names.

A larger initial sample was selected for this group because we anticipated a lower response rate. Since financial analysts represent the most "questionnaired" group in the country, it is difficult to achieve high response rates with them. Our goal was at least 150 usable responses from each sample. Assuming a 30 percent response rate for the financial executives and a 20 percent rate for the financial analysts, samples of 500 and 800, respectively, were drawn.

Questionnaire Administration
Questionnaire packets were mailed during the summer of 1977 directly to each person in both samples. The packet contained: (1) a cover letter explaining the purpose of the study, identifying the supporting organizations, and assuring the confidentiality of individual responses, (2) the eight-page questionnaire, and (3) a stamped, addressed returned envelope. A second mailing was made to all non-respondents approximately four weeks after the initial mailing. The only change in the packet was the cover letter. Response rates from the two mailings are summarized in Exhibit 2.

We consider the overall response rate of 40.8 percent to be excellent for a study of this type. The response rate of 61.5 percent for the financial executives is a strong indication of the high level of interest in the topic. A rate of 27.7 percent for chartered financial analysts is also higher than normal for this group.
Exhibit 2. Summary of Response Rates

<table>
<thead>
<tr>
<th>Sample group</th>
<th>Questionnaires mailed</th>
<th>Usable responses received</th>
<th>1st Mailing</th>
<th>2nd Mailing</th>
<th>Total</th>
<th>Response rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Financial executives</td>
<td>500 499</td>
<td></td>
<td>182</td>
<td>125</td>
<td>307</td>
<td>61.5</td>
</tr>
<tr>
<td>Financial analysis</td>
<td>800 790</td>
<td></td>
<td>118</td>
<td>101</td>
<td>219</td>
<td>27.7</td>
</tr>
<tr>
<td>Total</td>
<td>1300 1289</td>
<td></td>
<td>300</td>
<td>226</td>
<td>526</td>
<td>46.8</td>
</tr>
</tbody>
</table>

Response Bias
Although a nonresponse bias is of little concern with a 61.5 percent response rate, it is potentially troublesome with only a 27.7 percent response rate. For each of the samples, responses on the first mailing were compared with responses on the second mailing. Using a non-parametric chi-square test, no significant differences were found. To the extent that late responders are a surrogate for nonresponders, this test indicates that nonresponse bias is not a major problem in this study.

FINDINGS
For each of the two samples, data are presented regarding four issues:
(1) Perceptions about the disclosure option the respondents consider to be most appropriate for each of the questionnaire caselets. The issue here is whether a footnote disclosure alone is considered adequate;
(2) Perceptions about the disclosure option the respondents believe the annual report would contain. The issue here is the respondents' views of the way "subject to" opinions are used in practice,
(3) Predictive accuracy among the respondents. Since the disclosure option actually used in the annual report is known for all the cases, a direct comparison of the "would be" responses with the actual results can be made; and
(4) Attitudes of the respondents toward the legal liability issue.
For each of these four issues, responses for the two samples are also compared.
Desired Disclosure Levels

Exhibit 3 summarizes the comparative data regarding the "should be" responses. For all the cases combined, the most frequently chosen category is the "subject to" opinion which represents 40.4 percent of the financial executives' responses and 50.9 percent of the financial analysts' responses. The financial executives indicated that a "subject to" opinion was preferred for the litigation cases (45.7 percent) and the going-concern cases (52.8 percent), as well as for those cases in which the auditor did issue a qualified opinion (48.9 percent). This group indicated that footnote disclosure alone is adequate for the asset realization cases (46.0 percent), the major but not disastrous cases (52.6 percent), and for the cases in which the annual report did in fact show only footnote disclosure (31.9 percent). The financial analysts felt that a "subject to" opinion was the preferred disclosure in all cases. The response percentages for the "subject to" opinion category ranged from a high of 59.5 percent for going-concern contingencies to a low of 42.4 percent for the major but not drastic contingencies. Congruence between the two samples regarding the disclosure level considered to be most appropriate is measured by comparing the response frequencies using a nonparametric chi-square test. The chi-square values, as reported in exhibit 3, are significant for all seven comparisons. This indicates that financial executives and financial analysts do have different perceptions regarding how the contingencies can best be disclosed.

It is interesting that only 7.2 percent of the overall responses for the financial executives and 5.2 percent for the financial analysts indicate that the contingencies need not be disclosed in the audited section of the annual report. Further, the two groups differ little in the percentage of responses calling for a disclaimer of opinion. The major differences related to footnote disclosure alone, versus the "subject to" opinion alternative. The financial executives do not favor the "subject to" opinion option as strongly as the financial analysts.

Making comparisons across related issues, litigation contingencies generate noticeably more support for a "subject to" opinion than do asset realization contingencies for both groups. This can be interpreted as one justification for the emerging tendency among CPA firms to devote comparatively more audit attention to litigation problems than to asset realization problems.

Another issue of concern to AudSEC is whether the "subject to" opinion should be reserved for use only in situations involving a
<table>
<thead>
<tr>
<th></th>
<th>No reference needed</th>
<th>Disclosure in unaudited section only</th>
<th>Footnotes (only)</th>
<th>&quot;Subject to&quot; opinion</th>
<th>Disclaimer of opinion</th>
<th>Chi-square value†</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Asset realization cases</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Financial executives</td>
<td>4.7</td>
<td>5.0</td>
<td>46.0</td>
<td>35.2</td>
<td>9.1</td>
<td>53.25*</td>
</tr>
<tr>
<td>Financial analysts</td>
<td>1.5</td>
<td>4.6</td>
<td>34.4</td>
<td>48.1</td>
<td>11.4</td>
<td></td>
</tr>
<tr>
<td><strong>Litigation cases</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Financial executives</td>
<td>2.4</td>
<td>2.4</td>
<td>27.5</td>
<td>45.7</td>
<td>22.0</td>
<td>27.41*</td>
</tr>
<tr>
<td>Financial analysts</td>
<td>0.7</td>
<td>3.7</td>
<td>20.1</td>
<td>53.6</td>
<td>21.9</td>
<td></td>
</tr>
<tr>
<td><strong>Going-concern cases</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Financial executives</td>
<td>2.4</td>
<td>4.0</td>
<td>20.6</td>
<td>52.8</td>
<td>19.9</td>
<td>29.87*</td>
</tr>
<tr>
<td>Financial analysts</td>
<td>0.9</td>
<td>4.2</td>
<td>12.7</td>
<td>59.5</td>
<td>22.7</td>
<td></td>
</tr>
<tr>
<td><strong>Major but not drastic cases</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Financial executives</td>
<td>4.7</td>
<td>3.4</td>
<td>52.6</td>
<td>28.1</td>
<td>11.2</td>
<td>58.34*</td>
</tr>
<tr>
<td>Financial analysts</td>
<td>1.3</td>
<td>4.2</td>
<td>41.8</td>
<td>42.2</td>
<td>10.5</td>
<td></td>
</tr>
<tr>
<td><strong>&quot;Subject to&quot; opinion cases</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Financial executives</td>
<td>2.5</td>
<td>3.2</td>
<td>26.3</td>
<td>48.9</td>
<td>19.1</td>
<td>16.44*</td>
</tr>
<tr>
<td>Financial analysts</td>
<td>0.8</td>
<td>3.7</td>
<td>20.9</td>
<td>54.7</td>
<td>19.9</td>
<td></td>
</tr>
<tr>
<td><strong>Unqualified opinion cases</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Financial executives</td>
<td>4.5</td>
<td>4.2</td>
<td>47.3</td>
<td>31.9</td>
<td>12.1</td>
<td>66.22*</td>
</tr>
<tr>
<td>Financial analysts</td>
<td>1.4</td>
<td>4.6</td>
<td>38.6</td>
<td>47.0</td>
<td>13.3</td>
<td></td>
</tr>
<tr>
<td><strong>All cases combined</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Financial executives</td>
<td>3.5</td>
<td>3.7</td>
<td>36.7</td>
<td>40.4</td>
<td>15.7</td>
<td>74.33*</td>
</tr>
<tr>
<td>Financial analysts</td>
<td>1.1</td>
<td>4.1</td>
<td>27.2</td>
<td>50.9</td>
<td>16.7</td>
<td></td>
</tr>
</tbody>
</table>

†Chi-square values are calculated from the underlying frequency distributions for the two samples.

*P < .05.
The problem of "going-concern" magnitude. From exhibit 3, support for the use of a "subject to" opinion is noticeably higher in both samples for the going-concern cases than for the less drastic cases. Nevertheless, the support for at least "subject to" disclosure is still quite high in both samples for these less severe cases (39.3 percent of the responses for the financial executives and 52.7 percent for the financial analysts). These data indicate that neither sample group would favor reserving "subject to" opinions for use only in going-concern situations.

**Anticipated Disclosure Levels**

Exhibit 4 summarizes the comparative data regarding the "would be" responses. In general, the financial executives anticipate noticeably higher levels of disclosure than the financial analysts. The distributions are significantly different under the chi-square test for all seven of the comparisons. For all cases combined, 49.7 percent of the financial executives' responses indicate that the auditor would issue at least a "subject to" opinion versus only 28.9 percent of the financial analysts' responses. Further, only 11.4 percent of the financial executives' responses anticipate that there would be no mention of the contingency in the audited portion of the annual report, whereas 24.2 percent of the financial analysts' responses anticipate that this would be the case.

This difference in the anticipated posture of the auditor was surprising to us. It may reflect the differing contexts in which the two groups observe the audit function. The financial analysts do not actually observe auditors at work. Their perceptions may be heavily influenced by the widespread publicity accorded alleged breakdowns in auditing. Financial executives, on the other hand, do observe at least one audit firm at work each year. Furthermore, many financial executives have worked in public accounting earlier in their careers. Their perceptions of the auditor's posture may be influenced by their own struggles with their company's auditor or by their recollections of such struggles when they themselves were auditors. For whatever reasons, our data do seem to indicate that financial executives and financial analysts differ markedly in their anticipation of the aggressive auditors will exhibit in regard to the disclosure of material contingencies.

The lack of congruence between the expectations of the two groups can also be interpreted as evidence supporting the contention of the
### Exhibit 4. Comparative Perceptions Regarding the Disclosure Level Most Likely to be Used in the Annual Report

<table>
<thead>
<tr>
<th></th>
<th>No reference needed</th>
<th>Disclosure in unaudited section only</th>
<th>Footnotes (only)</th>
<th>&quot;Subject to&quot; opinion</th>
<th>Disclaimer of opinion</th>
<th>Chi-square value†</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Asset realization cases</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Financial executives</td>
<td>7.5</td>
<td>7.7</td>
<td>46.6</td>
<td>33.2</td>
<td>5.0</td>
<td></td>
</tr>
<tr>
<td>Financial analysts</td>
<td>13.4</td>
<td>15.4</td>
<td>51.3</td>
<td>17.4</td>
<td>2.5</td>
<td>97.19*</td>
</tr>
<tr>
<td><strong>Litigation cases</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Financial executives</td>
<td>3.5</td>
<td>3.9</td>
<td>31.6</td>
<td>47.8</td>
<td>13.4</td>
<td></td>
</tr>
<tr>
<td>Financial analysts</td>
<td>5.1</td>
<td>14.6</td>
<td>42.4</td>
<td>29.8</td>
<td>8.6</td>
<td>185.00*</td>
</tr>
<tr>
<td><strong>Going-concern cases</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Financial executives</td>
<td>3.6</td>
<td>6.9</td>
<td>24.6</td>
<td>54.1</td>
<td>10.8</td>
<td></td>
</tr>
<tr>
<td>Financial analysts</td>
<td>10.5</td>
<td>15.6</td>
<td>36.5</td>
<td>31.5</td>
<td>6.0</td>
<td>142.95*</td>
</tr>
<tr>
<td><strong>Major but not drastic cases</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Financial executives</td>
<td>7.3</td>
<td>4.7</td>
<td>53.7</td>
<td>26.8</td>
<td>7.5</td>
<td></td>
</tr>
<tr>
<td>Financial analysts</td>
<td>7.9</td>
<td>14.5</td>
<td>57.2</td>
<td>15.2</td>
<td>5.1</td>
<td>88.74*</td>
</tr>
<tr>
<td><strong>&quot;Subject to&quot; opinion cases</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Financial executives</td>
<td>4.7</td>
<td>5.4</td>
<td>80.2</td>
<td>47.3</td>
<td>12.3</td>
<td></td>
</tr>
<tr>
<td>Financial analysts</td>
<td>10.9</td>
<td>14.1</td>
<td>41.4</td>
<td>25.7</td>
<td>7.8</td>
<td>153.22*</td>
</tr>
<tr>
<td><strong>Unqualified opinion cases</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Financial executives</td>
<td>6.2</td>
<td>6.3</td>
<td>48.1</td>
<td>33.7</td>
<td>5.9</td>
<td></td>
</tr>
<tr>
<td>Financial analysts</td>
<td>7.5</td>
<td>15.9</td>
<td>52.3</td>
<td>20.9</td>
<td>3.4</td>
<td>81.10*</td>
</tr>
<tr>
<td><strong>All cases combined</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Financial executives</td>
<td>5.4</td>
<td>5.8</td>
<td>39.1</td>
<td>40.5</td>
<td>9.2</td>
<td></td>
</tr>
<tr>
<td>Financial analysts</td>
<td>9.2</td>
<td>15.0</td>
<td>46.9</td>
<td>23.8</td>
<td>5.6</td>
<td>215.37*</td>
</tr>
</tbody>
</table>

†Chi-square values are calculated from the underlying frequency distributions for the two samples.

*P < .05.
CAR that widespread confusion exists regarding the way “subject to” opinions are used in practice. As reflected in exhibit 4, both groups cannot be correct in their assessment of the way auditors use “subject to” opinions.

**Accuracy in Predicting Disclosure Levels**

Exhibit 5 summarizes the comparative data regarding predictive accuracy for the two groups. In this table, responses regarding the anticipated disclosure level are matched with the actual disclosure level in the annual report. A “hit” is recorded when the anticipated disclosure alternative is the same as the disclosure level actually used. A “high” is recorded when the anticipated disclosure level is higher than the actual level. A “low” is recorded when the anticipated disclosure level is lower than the level actually used. Overall, the rates are quite low for both samples. The financial executives’ hit rates range from 57.2 percent down to 38.2 percent across the test caselets. The financial analysts’ hit rates range from 52.3 percent down to a very low 25.7 percent. For all six comparisons considered, the responses of the two groups are significantly different under the nonparametric chi-square test. In general, the financial executives exhibit slightly higher predictive accuracy than the financial analysts. The financial analysts tend toward a consistently large percentage of low responses, whereas the misses for the financial executives are fairly evenly balanced between highs and lows.

In the preceding section, we observed that the disclosure levels anticipated by the financial executives were consistently noticeably higher than the levels anticipated by the financial analysts. We noted that both groups could not be correct in their assessment of the auditor’s posture. The results of the tests reported in this section suggest that neither group’s predictions are very accurate. The financial analysts tend to underestimate the actual disclosure level used. The financial executives fluctuate back and forth from underestimation to overestimation. This low level of predictive accuracy across both samples can be interpreted as further evidence in support of the contention by the CAR that widespread confusion exists regarding the way CPAs use “subject to” opinions.

**Defensive Posture**

Another issue probed in this study was the respondents’ views about how the issuance of a “subject to” audit opinion should affect the
Exhibit 5. Comparative Accuracy in Predicting Disclosure Levels

<table>
<thead>
<tr>
<th></th>
<th>Hits</th>
<th>Highs</th>
<th>Lows</th>
<th>Chi-square value†</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Asset realization cases</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Financial executives</td>
<td>42.1%</td>
<td>21.3</td>
<td>56.6</td>
<td></td>
</tr>
<tr>
<td>Financial analysts</td>
<td>52.3</td>
<td>14.4</td>
<td>53.2</td>
<td>55.67†</td>
</tr>
<tr>
<td><strong>Litigation cases</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Financial executives</td>
<td>53.2</td>
<td>30.6</td>
<td>16.2</td>
<td></td>
</tr>
<tr>
<td>Financial analysts</td>
<td>45.7</td>
<td>17.6</td>
<td>56.6</td>
<td>119.35†</td>
</tr>
<tr>
<td><strong>Going-concern cases</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Financial executives</td>
<td>38.2</td>
<td>39.0</td>
<td>22.8</td>
<td></td>
</tr>
<tr>
<td>Financial analysts</td>
<td>33.7</td>
<td>22.5</td>
<td>43.8</td>
<td>90.64†</td>
</tr>
<tr>
<td><strong>Major but not drastic cases</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Financial executives</td>
<td>57.2</td>
<td>12.9</td>
<td>30.0</td>
<td></td>
</tr>
<tr>
<td>Financial analysts</td>
<td>44.3</td>
<td>9.6</td>
<td>46.1</td>
<td>55.31†</td>
</tr>
<tr>
<td><strong>“Subject to” opinion cases</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Financial executives</td>
<td>47.3</td>
<td>12.3</td>
<td>40.4</td>
<td></td>
</tr>
<tr>
<td>Financial analysts</td>
<td>25.7</td>
<td>7.8</td>
<td>66.5</td>
<td>146.13†</td>
</tr>
<tr>
<td><strong>Unqualified opinion cases</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Financial executives</td>
<td>48.0</td>
<td>39.6</td>
<td>12.4</td>
<td></td>
</tr>
<tr>
<td>Financial analysts</td>
<td>52.3</td>
<td>24.3</td>
<td>23.4</td>
<td>71.77†</td>
</tr>
<tr>
<td><strong>All cases combined</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Financial executives</td>
<td>47.0</td>
<td>26.0</td>
<td>26.4</td>
<td></td>
</tr>
<tr>
<td>Financial analysts</td>
<td>39.0</td>
<td>16.1</td>
<td>44.9</td>
<td>159.54†</td>
</tr>
</tbody>
</table>

†Chi-square values are calculated from the underlying frequency distributions for two samples.
* p < .05.

The auditor’s defensive posture in situations where a contingency is ultimately resolved unfavorably for the client and litigation ensues. A summary of the findings is shown in exhibit 6.

A total of 79.6 percent of the responses for the two samples, combined, indicated that a “subject to” opinion should, at least, moderately improve the auditor’s defensive posture. The favorable responses are fairly evenly divided between “moderately favorable effect” and “significantly favorable effect.” The total percentage of favorable responses is only slightly higher for the litigation cases (82.8 percent) than for the asset realization cases (76.4 percent). The results are also consistent across the two samples. This suggests to us a fairly
high degree of consensus among the respondents that the auditor's legal obligation for full and fair disclosure is more clearly met by issuing a "subject to" opinion in situations such as those in the questionnaire.

The CAR cited two actual court cases as evidence to support their contention that issuing a "subject to" opinion does not improve the auditor's defensive posture. To the extent that the law changes over time to reflect widely held social mores, the data cited here could be interpreted as evidence that the contention by the CAR is not indisputable. Financial analysts and financial executives are not a direct surrogate for judges and juries, but their attitudes do reflect some of the issues which judges and juries must consider in such cases. In this regard, it is interesting to note that virtually identical findings were also obtained when the questionnaire was administered to a sample of senior lending officers in large banks and a sample of partners in major public accounting firms (Shank, Dillard and Murdock, 1978 (a) and (b)).

**Perceived Usefulness of "Subject to" Opinions**

The report of the CAR carefully avoids any direct reference to the "usefulness" of "subject to" opinions. A perception that such opinions are useful might be predicated on a misperception of the role a contingency qualification is intended to serve. According to the CAR, since the auditor cannot resolve the contingency, the "subject to" qualification is inappropriate for at least three reasons:

1. it adds nothing substantive to the disclosure beyond what is reported in the footnotes;
2. it encourages the reader to infer that the auditor concurs in management's statements about the anticipated resolution of the contingency; and
3. it can foster the misconception that an unqualified opinion implies there are no significant unresolved contingencies.

We believe it is fair to conclude that the CAR recommends eliminating "subject to" opinions because they believe such opinions do not serve a valid useful purpose. At least two of the major public accounting firms also support this view.

There is another widely held viewpoint. According to this view, which is also espoused by at least two of the major public accounting firms, the "subject to" opinion serves as a red flag to alert statement readers to major contingencies which the auditor believes are particu-
### Exhibit 6. Defensive Posture Impact

<table>
<thead>
<tr>
<th></th>
<th>Percentage of responses</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Significantly adverse effect</td>
</tr>
<tr>
<td><strong>Asset realization issue</strong></td>
<td></td>
</tr>
<tr>
<td>Financial executives</td>
<td>1.0</td>
</tr>
<tr>
<td>Financial analysts</td>
<td>2.4</td>
</tr>
<tr>
<td><strong>Litigation issue</strong></td>
<td></td>
</tr>
<tr>
<td>Financial executives</td>
<td>1.7</td>
</tr>
<tr>
<td>Financial analysts</td>
<td>1.5</td>
</tr>
<tr>
<td><strong>Combined</strong></td>
<td></td>
</tr>
<tr>
<td>Financial executives</td>
<td>1.4</td>
</tr>
<tr>
<td>Financial analysts</td>
<td>2.0</td>
</tr>
</tbody>
</table>
larly significant in interpreting the financial statements. Under this view, even though the auditor cannot resolve the contingency, an acknowledgement in the audit opinion of its existence does communicate useful information to the reader. Proponents of this view would dispute the contention that the opinion qualification adds nothing substantive to the disclosure. They would also dispute the contention that statement readers naively assume unqualified opinions imply no unresolved contingencies.

We believe this is an issue which can and should be investigated empirically before drawing a definitive conclusion. The data presented in exhibit 3 can be interpreted as bearing at least indirectly on this issue. The analysis assumes that if a respondent specifies a "subject to" opinion or a disclaimer as the preferred disclosure for the contingency, he or she perceives opinion qualifications to represent useful extensions of the attest function. Such an analysis is predicated on the view that the respondents do not interpret the qualification as a surrogate for the anticipated resolution of the conflict. Since the data do not control for the mind set of respondents, they can only be interpreted as an indirect, preliminary test of the "valid use" question. Exhibit 7 presents the data extracted from exhibit 3.

Approximately 62 percent of the total responses specify that a "subject to" opinion or disclaimer is preferred to adequately communicate the significance of the contingency. Overall, 56.1 percent of the financial executives' responses and 69.5 percent of the financial analysts' responses fall in this range. The financial executives' response rates range from a high of 72.7 percent for going-concern contingencies to a low of 39.4 percent for the major but not drastic contingencies. The financial analysts' response rates range from a high of 82.2 percent for going-concern contingencies to a low of 52.7 percent for the major but not drastic contingencies. Based on these results, we would argue that more research is needed before concluding that contingency qualifications serve no valid, useful role.

**SUMMARY AND CONCLUSIONS**

The incidence of "subject to" opinions has increased dramatically in the past ten years. The efficacy of this extension of the audit communication process is now being questioned by many people. The CAR has recommended that the responsibility for disclosing all major contingencies fall exclusively on corporate management. In their view, annual reports should contain a separate footnote, similar
### Exhibit 7. Preferred Disclosure Levels—Use of Qualified Opinions

<table>
<thead>
<tr>
<th>Case Type</th>
<th>Percentage of responses preferring level 4 or 5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Asset realization cases</strong></td>
<td></td>
</tr>
<tr>
<td>Financial executives</td>
<td>44.2</td>
</tr>
<tr>
<td>Financial analysts</td>
<td>59.5</td>
</tr>
<tr>
<td><strong>Litigation cases</strong></td>
<td></td>
</tr>
<tr>
<td>Financial executives</td>
<td>67.7</td>
</tr>
<tr>
<td>Financial analysts</td>
<td>75.5</td>
</tr>
<tr>
<td><strong>Going-concern cases</strong></td>
<td></td>
</tr>
<tr>
<td>Financial executives</td>
<td>72.2</td>
</tr>
<tr>
<td>Financial analysts</td>
<td>82.2</td>
</tr>
<tr>
<td><strong>Major but not drastic cases</strong></td>
<td></td>
</tr>
<tr>
<td>Financial executives</td>
<td>39.4</td>
</tr>
<tr>
<td>Financial analysts</td>
<td>52.7</td>
</tr>
<tr>
<td><strong>“Subject to” opinion cases</strong></td>
<td></td>
</tr>
<tr>
<td>Financial executives</td>
<td>68.0</td>
</tr>
<tr>
<td>Financial analysts</td>
<td>74.6</td>
</tr>
<tr>
<td><strong>Unqualified opinion cases</strong></td>
<td></td>
</tr>
<tr>
<td>Financial executives</td>
<td>44.0</td>
</tr>
<tr>
<td>Financial analysts</td>
<td>60.3</td>
</tr>
<tr>
<td><strong>All cases combined</strong></td>
<td></td>
</tr>
<tr>
<td>Financial executives</td>
<td>56.1</td>
</tr>
<tr>
<td>Financial analysts</td>
<td>69.5</td>
</tr>
</tbody>
</table>

to that on accounting policies, describing the circumstances surrounding each major uncertainty and management's assumptions about the range of possible outcomes and their potential effects. The sole responsibility of the auditor in this regard would be to assure full and fair disclosure by management. Auditors would no longer be responsible for communicating their interpretations of the potential seriousness of one contingency versus another.

Unfortunately, very little systematic evidence exists regarding this issue. This study provides some preliminary empirical evidence regarding statement issuers' and statement readers' perceptions of and attitudes toward “subject to” opinions. We believe that such evidence is one important input into the policy-making process.

This study is designed to provide empirical evidence regarding
the attitudes of sophisticated financial statement readers and issuers toward the use by auditors of "subject to" audit opinions.

Five issues are addressed specifically:

1. **Perceived usefulness**—do sophisticated readers and issuers of financial statements perceive "subject to" audit opinions to represent a useful extension of the auditor's response in situations involving material contingencies?

2. **Predictive accuracy**—how well do sophisticated statement issuers and readers understand the way generally accepted audit standards (GAAS) relating to "subject to" opinions are actually applied in practice?

3. **Defensive posture**—what impact do sophisticated statement issuers and readers believe a "subject to" opinion should have on the auditor's legal liability for full and fair disclosure?

4. **Litigation versus asset realization contingencies**—do sophisticated statement readers and issuers perceive litigation issues to be sufficiently more important than asset realization issues to warrant the significant additional audit time devoted to concern over litigation disclosures?

5. **Going-Concern issues versus those of less drastic proportions**—is there evidence in our sample that sophisticated statement readers and issuers believe "subject to" opinions should be reserved for those situations involving contingencies of "going concern" proportions?

On these issues, the findings can be summarized as follows:

1. The respondents to our survey do consider a "subject to" audit opinion to provide useful information for cases such as those in this study. Over all eight of the cases, about 62 percent of the responses indicated that a qualified opinion was the preferred disclosure for the contingency described;

2. Perceptions about the anticipated posture of the auditor regarding the disclosure level in the annual report differed markedly between the statement issuers and the statement users. The financial analysts expected much lower disclosure levels than did the financial executives;

3. For the eight case situations in the questionnaire, neither the financial executives nor the financial analysts we surveyed were very good predictors of the actual treatment accorded the contingency in the annual report. Only about 44 percent of the responses correctly identified the actual disclosure used. There does appear to be substan-
tial confusion about how contingency qualification standards are being applied under current GAAS;
(4) Our findings indicate a high degree of consensus that a "subject to" opinion should improve the auditor's defensive position if the contingency is resolved unfavorably for the company and the adequacy of disclosure is subsequently questioned in court. About 80 percent of the responses for both samples indicated that the qualified opinion should have a favorable effect upon the court. More specifically, 37 percent of the responses indicated it should have a significantly favorable effect and 43 percent a moderately favorable effect;
(5) The study also attempted to determine whether the respondents differ in their attitude toward asset realization contingencies and litigation contingencies. Although both sets of respondents are better predictors of litigation disclosure than asset realization disclosure, neither category yielded a correct response rate higher than 1 out of 2 (53 percent and 45 percent versus 42 percent and 52 percent). When asked to indicate how the asset realization cases should be disclosed, about 52 percent of the responses indicated that the situations required at least a qualified opinion. However, over 71 percent of the responses indicated that the litigation cases warrant at least a qualified opinion.

This can be interpreted as evidence that the special audit attention typically accorded litigation matters is justified in terms of the respondents' perceptions of relative importance; and
(6) On another dimension, about 77 percent of the responses indicated that at least a qualified opinion is warranted for the going-concern contingencies. However, about 46 percent of the responses indicated that this level of disclosure is warranted for the less severe contingencies as well. There is thus no noticeable support for limiting qualifications to only drastic contingencies as tentatively proposed by AudSEC.

In interpreting these findings, the following limitations must be kept in mind. The five disclosure level alternatives are assumed to represent the full range of disclosure available and to contain all reasonable disclosure options. The caselets included in the questionnaire are assumed to be representative of the broad issues addressed. Further, for each case, it is assumed that the most appropriate disclosure level was chosen by the company's auditor, in conformity with GAAS, and that the key, relevant information leading to this decision is presented in the questionnaire caselet. The generaliza-
bility of the study's findings is limited by the extent to which these assumptions are valid.

APPENDIX A

Qualified Audit Opinions. A Questionnaire

This questionnaire comprises eight specific fact situations which you are asked to judge. All eight cases are taken from actual annual reports and each should be considered separately.

In each case the financial statements have been audited by an independent CPA in accordance with generally accepted auditing standards. Except for the specific material uncertainty described in the case, the financial statements are fairly presented, in accordance with generally accepted accounting principles.

Depending on your view of the function of the auditor’s opinion and your perception of the severity of the fact situation, you are asked to judge which of the following five disclosure options would be most appropriate:

A. No specific reference to this situation in the annual report.
B. Disclosure only in the unaudited section of the annual report.
C. Disclosure in a footnote to the financial statements, but no further mention.
D. Disclosure in the footnotes, plus a three-paragraph “subject to” audit report in which the middle paragraph provides an explanation of the contingency and its anticipated resolution.
E. A “disclaimer of opinion” by the auditor, along with the auditor’s explanation for the disclaimer.

In addition to judging which option you believe to be most appropriate in each case, you are also asked which option you think would actually be used in the annual report.

Wherever “disclosure” is mentioned, you may assume that the description will be complete and forthright. This questionnaire deals with the way the uncertainty should be disclosed—not with the adequacy of the disclosure itself.

All responses to this questionnaire will, of course, remain strictly confidential. Only aggregate level responses will be reported.

If you would like to receive a summary of the research findings, simply check this box. □

Are you a CPA? □ Yes □ No

If not, do you consider yourself knowledgeable about generally accepted auditing standards?

□ Yes □ No
CASE I*

*(This Case Preceded the Issuance of FASB 12 Dealing with Marketable Securities)*

A. In the current balance sheet the company valued its portfolio of marketable equity securities at cost. These securities comprise 11% of the company's total assets and 22% of shareholders' equity. At year end, the market value of this portfolio was approximately 45% of its cost. However, management believes that the depressed market prices are only temporary. Since the company has no plans to dispose of the portfolio, they believe there is no need to value the securities at the currently depressed market prices.

Part 1. Which of the five disclosure alternatives shown below would you choose to communicate the contingency described above?

A. No specific reference to this situation in the annual report.
B. Disclosure only in the unaudited section of the annual report.
C. Disclosure in a footnote to the financial statements, but no further mention.
D. Disclosure in the footnotes, plus a three-paragraph "subject to" audit report in which the middle paragraph provides an explanation of the contingency and its anticipated resolution.
E. A "disclaimer of opinion" by the auditor, along with the auditor's explanation for the disclaimer.

Please circle the letter of your first and second choices:

<table>
<thead>
<tr>
<th>First choice</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Second choice</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
</tr>
</tbody>
</table>

Part 2. Which disclosure alternative do you think the company's annual report would actually show?

Please circle the letter of your choice.

Most likely disclosure actually used

| Most likely disclosure actually used | A | B | C | D | E |

B. In the following year, the market remained depressed. The Company adopted the lower of cost or market method of accounting for the marketable securities. Valuation allowances were set-up to allocate the unrealized losses to the applicable periods. The amount which would eventually be realized from the securities was, of course, still contingent on future events.

Part 1. Which of the five disclosure options listed above would you choose to communicate the contingency in the second year?

Please circle the letter of your first and second choices:

<table>
<thead>
<tr>
<th>First choice</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Second choice</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
</tr>
</tbody>
</table>

Part 2. Which disclosure alternative do you think the company's annual report would actually show?
CASE II—A

The Company has entered into supply contracts with 27 customers which call for delivery of approximately 80 million pounds of uranium over the next twenty years, at an average price of $9.50 per pound, with price escalation based on industrial indices but not keyed to changes in the market price of uranium. The company has approximately 15 million pounds of uranium in inventory or under firm contract, leaving a short fall of approximately 65 million pounds. Recent market price quotations for uranium have been approximately $40 per pound with substantial down payments required. Starting in seven years, the company is also under contract to deliver an additional 9.2 million pounds of uranium if plutonium dilution cannot be perfected. Under current market conditions, if the corporation is required to fulfill all the contracts, the financial impact will be extremely adverse. Under such conditions, the net worth of the company would be virtually eliminated.

Part 1. Which of the five disclosure alternatives shown below would you choose to communicate the contingency described above?

A. No specific reference to this situation in the annual report.
B. Disclosure only in the unaudited section of the annual report.
C. Disclosure in a footnote to the financial statements, but no further mention.
D. Disclosure in the footnotes, plus a three-paragraph “subject to” audit report in which the middle paragraph provides an explanation of the contingency and its anticipated resolution.
E. A “disclaimer of opinion” by the auditor, along with the auditor’s explanation for the disclaimer.

Please circle the letter of your first and second choices:

<table>
<thead>
<tr>
<th>First choice</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Second choice</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
</tr>
</tbody>
</table>

Part 2. Which disclosure alternative do you think the company’s annual report would actually show?

Please circle the letter of your choice.

Most likely disclosure actually used

| A | B | C | D | E |

CASE II—B

During the next year, the company notified its customers that fulfillment of the uranium supply contracts described in Case II—A was
excused under the legal doctrine of commercial impracticability. This notice resulted in 17 law suits being filed against the company by 27 customers alleging failure to fulfill the uranium supply contracts. Two of these law suits allege violations of the anti-trust laws. Under an agreement affirmed by court order, the company will deliver its approximately 15 million pounds of uranium inventory on the basis of the plaintiffs' claimed rights, subject to later final determination of the proper price. This arrangement also provides for establishment of a committee to enter into discussion and negotiations directed toward an amicable resolution of disputes beyond the 15 million pounds.

Three distinct outcomes are possible. The law suits may be settled under the court ordered negotiating arrangements noted above, in which case the costs of settlement to the company could well be substantial. In the meantime, the company will continue vigorously to assert its defense under the Uniform Commercial Code. If the company is not wholly successful and is granted only partial relief from its alleged contractual obligations, the financial impact could be severe. If the company is required to fulfill all the contracts under current market conditions, the financial impact will, of course, be extremely adverse.

Part 1. Which of the five disclosure alternatives shown below would you choose to communicate the contingency described above?
A. No specific reference to this situation in the annual report.
B. Disclosure only in the unaudited section of the annual report.
C. Disclosure in a footnote to the financial statements, but no further mention.
D. Disclosure in the footnotes, plus a three-paragraph "subject to" audit report in which the middle paragraph provides an explanation of the contingency and its anticipated resolution.
E. A "disclaimer of opinion" by the auditor, along with the auditor's explanation for the disclaimer.

Please circle the letter of your first and second choices:

<table>
<thead>
<tr>
<th>First choice</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Second choice</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
</tr>
</tbody>
</table>

Part 2. Which disclosure alternative do you think the company's annual report would actually show?
Please circle the letter of your choice.

| Most likely disclosure actually used | A | B | C | D | E |

CASE III

Inventories applicable to the company's aircraft product line constitute 47% of total assets. The carrying value of this inventory is 28 times
the amount of owners' equity. The company's studies indicate that the carrying value of the aircraft inventories will be recovered and that gross profit will be realized over the remainder of the program.

In the middle of this year, production was in the early, high-cost period with a continuing buildup of unrecovered production start-up costs, although at a declining rate. As of year-end, production costs incurred are less than the sales price of aircraft being delivered. Because production rates are expected to be reduced in the near term, it is expected that current margins between recurring production cost and sales prices may not be maintained. However, over the remaining term of the program, margins are expected to increase. This should result in recovery of the initial planning and tooling costs, plus all unrecovered production costs of previously delivered aircraft, plus a program gross profit.

Recovery of the year-end inventory is dependent on the number of aircraft ultimately sold and on actual selling prices and costs. Through the end of this year, 157 units have been sold or committed under firm orders. The current order backlog is 12 units. The recovery of 2/3 of the inventory carrying amount is dependent on receipt in the future of orders for an additional 150 aircraft. Continued financing will be required until the inventory is liquidated.

Sales significantly under estimates or costs significantly over estimates could result in losses on the program in subsequent periods which would eliminate the net worth of the company. The company's projections of future sales and costs are based on assumptions as to future events which cannot be quantified with precision and which are subject to periodic revision.

Part 1. Which of the five disclosure alternatives shown below would you choose to communicate the contingency desired above?
A. No specific reference to this situation in the annual report.
B. Disclosure only in the unaudited section of the annual report.
C. Disclosure in a footnote to the financial statements, but no further mention.
D. Disclosure in the footnotes, plus a three-paragraph "subject to" audit report in which the middle paragraph provides an explanation of the contingency and its anticipated resolution.
E. A "disclaimer of opinion" by the auditor, along with the auditor's explanation for the disclaimer.

Please circle the letter of your first and second choices:
First choice  A  B  C  D  E
Second choice A  B  C  D  E

Part 2. Which disclosure alternative do you think the company's annual report would actually show?
Please circle the letter of your choice:
Most likely disclosure actually used  A  B  C  D  E
CASE IV—A

Two years ago a major competitor commenced legal action against a wholly-owned subsidiary and two individuals for goods sold and delivered, asserting a claim in the amount of $38,782. The action, initially brought in a state court, has been removed to the U.S. district court.

In January of this year, the original complaint was amended to add several new claims and to add this company, the parent, as a party defendant. As amended, the complaint presently alleges claims for breach of contract, violations of antitrust laws, wrongful interference with business relationships, and trademark infringements. Total damages are claimed which exceed approximately 30% of the total assets and 55% of shareholders' equity. The company filed an answer to the amended complaint and a counter-claim against the plaintiff.

On the basis of all facts known to them, outside counsel believes that each of the claims added by the amended complaint is without merit. Accordingly, in the opinion of the company, the ultimate outcome of the litigation will not have a materially adverse effect on the financial position of the company.

Part 1. Which of the five disclosure alternatives shown below would you choose to communicate the contingency desired above?
   A. No specific reference to this situation in the annual report.
   B. Disclosure only in the unaudited section of the annual report.
   C. Disclosure in a footnote to the financial statements, but no further mention.
   D. Disclosure in the footnotes, plus a three-paragraph "subject to" audit report in which the middle paragraph provides an explanation of the contingency and its anticipated resolution.
   E. A "disclaimer of opinion" by the auditor, along with the auditor's explanation for the disclaimer.

Please circle the letter of your first and second choices:
   First choice  A  B  C  D  E
   Second choice A  B  C  D  E

Part 2. Which disclosure alternative do you think the company's annual report would actually show?

Please circle the letter of your choice.
   Most likely disclosure actually used A  B  C  D  E

CASE IV—B

In the following year a judgment was entered against the company in connection with the litigation described in Case IV—A in amounts totaling approximately 12% of the total assets and 20% of shareholders
equity. In addition to monetary damages, the court enjoined and restrained the company from engaging in certain business practices. This injunction, in the opinion of management, prohibited the subsidiary from engaging in generally accepted business practices. As a consequence of the monetary award and the injunction, the subsidiary terminated all of its business operations as of the end of this year since, in the opinion of the management, it could not be a viable competitor.

Losses relating to closure of the subsidiary are expected to be approximately 1.5 times its total assets. The company has filed an appeal in the United States Court of Appeals from each and every aspect of the judgment adverse to them. Counsel for the defendants state that there are substantial grounds for this appeal. However, they are not in a position to predict what relief, if any, will be granted by the Court of Appeals. No provisions have been made in the financial statements for possible damages.

Part 1. Which of the five disclosure alternatives shown below would you choose to communicate the contingency desired above?
A. No specific reference to this situation in the annual report.
B. Disclosure only in the unaudited section of the annual report.
C. Disclosure in a footnote to the financial statements, but no further mention.
D. Disclosure in the footnotes, plus a three-paragraph "subject to" audit report in which the middle paragraph provides an explanation of the contingency and its anticipated resolution.
E. A "disclaimer of opinion" by the auditor, along with the auditor's explanation for the disclaimer.

Please circle the letter of your first and second choices:

<table>
<thead>
<tr>
<th>First choice</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Second choice</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
</tr>
</tbody>
</table>

Part 2. Which disclosure alternative do you think the company's annual report would actually show?

Please circle the letter of your choice.

Most likely disclosure actually used

A | B | C | D | E

Part 3. If the contingency described in Case IV—B were subsequently resolved unfavorably for the company, litigation might result challenging the adequacy of the auditor's opinion on the financial statements. To what extent do you think the issuance of a "subject to" opinion should affect the auditor's defensive posture in this situation?

Circle the response which most closely reflects your opinion.

How should the issuance of a "Subject to" opinion affect the auditor's defensive posture?

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Significant</td>
<td>Moderate</td>
<td>Adverse</td>
<td>No</td>
<td>Moderate</td>
</tr>
<tr>
<td>Adverse Effect</td>
<td>Adverse Effect</td>
<td>Effect</td>
<td>Favorable</td>
<td>Favorable Effect</td>
</tr>
<tr>
<td>Effect</td>
<td>Effect</td>
<td>Effect</td>
<td>Effect</td>
<td>Effect</td>
</tr>
</tbody>
</table>

147
The company operates as a real estate investment trust and maintains a balanced portfolio of diversified real estate mortgage loans. The current year portfolio has the following composition: construction loans, 35%; intermediate term loans, 30%; long term loans, 18%; development loans, 12%; other loans, 5%. As of the end of the year, approximately 20% of all invested assets were in arrears (i.e., more than 90 days past due). Within three months after the year end, but before the annual report was issued, 25% of the total invested assets were in arrears.

Accrued interest receivable was double the normal level at the end of the fiscal year. Approximately 50% of the current year's accrued interest receivable represents interest on the delinquent accounts. Foreclosure proceedings are currently pending against delinquent loans accounting for about one fourth of accrued interest receivable.

Due to apparent adverse market conditions, the allowance for possible loan losses was increased from .2% to 2.0% of loans at year-end. Last year's income was 4% of total assets. This year's income was 0.5% of total assets.

Part 1. Which of the five disclosure alternatives shown below would you choose to communicate the contingency desired above?

A. No specific reference to this situation in the annual report.
B. Disclosure only in the unaudited section of the annual report.
C. Disclosure in a footnote to the financial statements, but no further mention.
D. Disclosure in the footnotes, plus a three-paragraph "subject to" audit report in which the middle paragraph provides an explanation of the contingency and its anticipated resolution.
E. A "disclaimer of opinion" by the auditor, along with the auditor's explanation for the disclaimer.

Please circle the letter of your first and second choices:

<table>
<thead>
<tr>
<th>First choice</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Second choice</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
</tr>
</tbody>
</table>

Part 2. Which disclosure alternative do you think the company's annual report would actually show?

Please circle the letter of your choice.

Most likely disclosure actually used:

A. A
B. B
C. C
D. D
E. E

Part 3. If the contingency described in Case V were subsequently resolved unfavorably for the company, litigation might result challenging the adequacy of the auditor's opinion on the financial statements. To what extent do you think the issuance of a "subject to" opinion should affect the auditor's defensive posture in this situation?
Circle the response which most closely reflects your opinion.

How should the issuance of a "Subject to" opinion affect the auditor's defensive posture?

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Significant Adverse Effect</td>
<td>Moderate Adverse Effect</td>
<td>No Effect</td>
<td>Moderate Favorable Effect</td>
<td>Significant Favorable Effect</td>
<td></td>
</tr>
</tbody>
</table>

**REFERENCES**


———, "Lending Officers' Attitudes Toward 'Subject to' Audit Opinions," *The Journal of Commercial Bank Lending* (March 1978), pp. 31-45.
Discussant's Response to "'Subject To' Audit Opinions: A Preliminary Investigation of Statement Users' and Statement Issuers' Perceptions"

D. R. CARMICHAEL*

My response covers three aspects of the paper:
1. Analysis of the issue;
2. Design of the questionnaire; and
3. Observations on survey research.
It concludes with some additional observations on reporting practices for uncertainties.

ANALYSIS OF THE ISSUE
Since the conclusions of the Commission on Auditors' Responsibilities on audit reporting requirements for uncertainties apparently, at least in part, stimulated the research, I would like to start by identifying the reasons behind the CAR recommendation. I do not believe Professors Shank and Dillard do them justice.

The CAR identified the following reasons in its report:
1. A responsibility to include information on uncertainties in his report is inconsistent with the auditor's role in expressing an opinion on other aspects of financial statements. (The auditor should not be a reporter or interpreter of financial information.)
2. The "subject to" phrase is ambiguous to users because there is no way for them to tell whether the auditor's intention is only to high-

*The views expressed are those of the author and do not necessarily reflect an official position of the AICPA.
light information more fully disclosed elsewhere or to indicate a deficiency in the financial statements.

3. The "subject to" qualification may cause the user to believe the financial statements will be restated when the uncertainty is resolved, but this will probably not be the case.

4. The standards for uncertainties that require qualification are inherently vague and not susceptible to a desirable degree of uniformity in practice.

5. Since the decision to qualify includes an evaluation of probability, the auditor's reporting decision is essentially a prediction of outcome of some future event. However, the auditor may be in no better position to predict than the average user.

6. Since auditors are known to qualify their opinions for some material uncertainties, the absence of a qualification may lead a user to believe a company faces no significant uncertainties.

7. Business risk and information risk should be distinguished. Evaluating the risks a business faces is a function that should be assumed by the user of financial statements. Evaluating the adequacy of disclosure of such risks is the auditor's responsibility.

8. The present uncertainty-reporting requirement inappropriately diverts the auditor's attention from evaluating the disclosure of uncertainties to high-lighting the existence of some uncertainties.

The research results support the CAR's belief that "subject to" qualifications are misunderstood. However, Shank and Dillard believe their results refute what they incorrectly regard as the CAR's implied belief that "subject to" qualifications "do not serve a valid useful purpose." Nevertheless, the paper recognizes that the CAR "avoids any direct reference to the 'usefulness' of 'subject to' opinions."

I believe Shank and Dillard have underestimated the importance the CAR attached to the incompatibility of the audit reporting requirement on uncertainties with its view of the auditor's role.

Naturally, financial analysts are primarily concerned with the interpretation of financial information. It is their reason for existence. Financial executives are responsible for preparing and presenting financial information. They also are concerned with interpreting financial information in several narrative sections of the annual report, including the president's letter and the financial review as well as the notes to the financial statements.

Financial analysts no doubt regard the auditor as another source of
financial information. Casting the auditor as a reporter or interpreter of information does not bother them a bit. Financial executives may see no reason that auditors should not have responsibilities for presentation not much different than their own. However, to a group such as the CAR charged with defining the role of the independent auditor in society, viewing the auditor as just another source of information or interpretation is objectionable.

Thus, when Shank and Dillard analyze the issue in a manner that results in asking whether the "subject to" qualification is a preferred "disclosure option," they do considerable violence to the CAR's views. They have not, as they seem to believe, merely restated its views in a simpler fashion.

I am not saying that Shank and Dillard should not be permitted to entertain the notion that users find "subject to" qualifications to be useful—the so-called red-flag theory. However, setting up the auditor's report as one of several disclosure options does not contribute much to the neutrality of the research design.

It would have been possible to design research to test one or more of the CAR's statements on the effect on users or to test the red-flag theory directly. For example, one group could have been presented with financial statements with only note disclosure and a control group could have been given identical data with the exception of the qualification in the auditor's report. Each group could have then been asked questions to discern their understanding of the uncertainty.

**Design of the Questionnaire**

Each case (I can not bring myself to call it a caselet) presents the survey subject with a set of facts and asks whether the subject thinks the information would be (or should be) disclosed as follows:

1. No specific reference to this situation in the annual report.
2. Disclosure only in the unaudited section of the annual report.
3. Disclosure in a footnote to the financial statements, but no further mention.
4. Disclosure in the footnotes, plus a three-paragraph "subject to" audit report in which the middle paragraph provides an explanation of the contingency and its anticipated resolution.
5. A "disclaimer of opinion" by the auditor, along with the auditor's explanation for the disclaimer.

The paper states that "The third disclosure alternative above repre-
sents the recommendation of the CAR that major contingencies by
disclosed by management in a separate footnote to the financial
statements, like the accounting policies footnote. . . ."

Shank and Dillard read much more into this disclosure alternative
than I do. Appendix A presents the CAR's recommendation for
comparison.

I believe that a survey subject would assume that alternative three is
the common garden-variety note found in virtually all published
financial statements. On the other hand, the CAR envisioned a note
with the following features:
1. A separate note, similar to the note on accounting policies, with a
standardized position and introduction.
2. A note that emphasizes the relative importance of uncertainties
through arrangement and description.
3. A note that—
   a. Describes the circumstances surrounding the uncertainty;
   b. States management's assumptions about the outcome of the
      uncertainty;
   c. Explains the range of possible outcomes and their potential
effects; and
   d. Includes any other considerations that bear on the probable
      outcome.

Alternative 4 states that the auditor's report will describe the "antici-
pated resolution" of the uncertainty. This is not required by
current standards and usually is not possible. However, this wording
was retained because Shank and Dillard concluded "current practice
does usually include a reference to management's view regarding the
resolution of the contingency." I disagree. Management's view may
appear in a note, but it usually does not appear in the auditor's report.
After all, if the auditor could substantiate management's prediction
he would not need to qualify.

Shank and Dillard offer as evidence an excerpt from the auditor's
report in the 1974 Lockheed annual report. I offer you, as Appendix
B, the auditor's report from the 1975 annual report of Lockheed.

Checking on this matter caused me to wonder how well the descrip-
tion of the uncertainty in each case captured the information that
could be expected in the typical annual report. Thus, I went to the
annual reports of the identifiable companies included in the cases.
You do not have to be a Wall Street whiz to figure out that Case II is
Westinghouse and Case III is Lockheed.
Staying with Lockheed, a look at its 1976 annual report shows what pikers Shank and Dillard were in offering disclosure alternatives to the survey subjects. Paging through Appendix C, which includes several excerpts and the entire financial statement section, you will notice that the uncertainty explained in Case III—recoverability of the Tri Star inventory—is disclosed in

1. An informative paragraph in the president's letter;
2. Over a full page in the financial review;
3. Two paragraphs in management's discussion and analysis of operations;
4. Separate line item coverage in all three basic financial statements; and
5. A two-page note to the financial statements.

In those circumstances, I can imagine that the auditor would feel left out if he could not say something in his report. But consider the user. Would he want to trade all this information for alternative four? How much would the user suffer without the terse, obscurely worded middle paragraph in the auditor's report? What user would care to know anything more about uncertainties by the time he reaches the auditor's report?

Appendix D contains similar information on Westinghouse's coverage of its problems with uranium supply contracts in its 1977 annual report. A user of Westinghouse's annual report would have a better chance of missing the auditor's report altogether than he would of failing to recognize the problems with uranium supply contracts.

Westinghouse had similar disclosures in its 1976 and 1975 annual reports. The problems on uranium supply contracts were first explained to users in the last quarter of 1975. Of course, the trick would have been to disclose the matter in the 1974 annual report. That is the real problem. The user is concerned with receiving an early warning of impending problems.

I cannot imagine that the users of the financial statements of Lockheed and Westinghouse have any need for a "subject to" qualification to make them aware of the existence or significance of the problems those companies face. Advance warning of such problems is another matter. However, all the user gets is a short "me too" comment in the auditor's report after everyone is aware of the problem.

SURVEY RESEARCH
I am biased. I admit it. I believe that most survey research of this type is superficial. It can only measure speech reaction which indicates
hypothesis. What can we learn from the fact that, based on a sketchy explanation of an uncertainty, a survey subject took between one to two minutes* to select one “disclosure option” from a list of matters that are not really alternatives?

I believe there is something drastically wrong with a value system that prefers the kind of information produced by a superficial mail survey to a thoughtful consideration by an experienced analyst.

The survey is replicable. Responses to it can be evaluated statistically. Those qualities cannot simply be dismissed. But I question whether a mail survey such as this one can produce more than trivial knowledge when responses may be influenced more by instrument design than relevant variables. What theories does it test? None. Are important hypotheses tested and do those hypotheses relate to our knowledge of how an annual report is used? In my estimation the answer is “No” to both questions.

The survey results are simply information. We know the preferences of survey subjects among a list of disclosure options. We know the subjects cannot predict very well which option will be used in practice. But we do not know anything more about how users interpret qualified and unqualified audit opinions. Shank and Dillard were apparently interested in this question. At an early point in the paper they say

There is no systematic empirical evidence of which we are aware regarding the extent to which issuers or readers of financial statements do interpret a “subject to” opinion to convey a message which is significantly different from that in an unqualified opinion.

I believe the survey results do little to illuminate that issue.

When research even marginally capable of being labeled systematic and empirical is criticized, the usual response is that it is the lesser of evils. Apparently the results of practically any survey of a group of users is preferred by some to the judgment of a group of auditors about communication with users.

I would not be so quick to dismiss the judgment of auditors on such matters. It is based on experience with

1. Explaining and discussing modifications of audit reports with client executives, outside board members, and the client’s legal counsel;

*The paper notes (p. 126) that the estimated completion time for the questionnaire is twenty-five minutes. There are eighteen individual questions. Thus, each question apparently requires less than 1.4 minutes to read and answer.
2. Responding to questions of stockholders at annual meetings;
3. Responding to inquiries of regulatory agencies; and
4. Responding to questions of bankers and financial analysts in certain circumstances.

Thus, a survey is not essential to gain some understanding of the needs and attitudes of users.

Another interesting source of information on the subject is the comment letters on the exposure draft on contingencies and uncertainties issued by the Auditing Standards Executive Committee. These are available to interested researchers. However, there is no consensus among those who responded. Nevertheless, several commentors agreed with these views expressed by the SEC:

We believe improvements in financial accounting standards concerning uncertainties must be addressed before considering the issue of eliminating “subject to” qualifications. When an adequate level of informative note disclosure has been achieved, it will then be appropriate to determine whether other means of communication to investors can be sufficient to serve the purpose of highlighting uncertainties and thus relieve obstacles to revisions to the auditor’s report.

CONCLUDING REMARKS

I share the SEC’s concern and hope that progress on improving the disclosure of uncertainties is swift because change is required. The difficulty with present practice in reporting on uncertainties is highlighted in the July issue of The Corporate Communications Report referred to in the paper. That report stated, “. . . Qualified opinions clearly are meaningless, perhaps even counter to fair allocation of capital, when two companies with similar legal exposure end up receiving such contradictory treatment at the hands of their auditors.”

The “contradictory treatment” refers to the fact that six major cigarette companies—all defendants in the same anti-trust litigation—received different types of auditor’s reports. Half the companies received qualifications based on the litigation uncertainty and the other half got unqualified opinions. The same public accounting firm gave an unqualified opinion to one client and a “subject to” qualification to another. The existence of a qualification was unrelated to the percentage of sales from tobacco products. The details are in Appendix E.

If we can abandon the notion that the auditor’s report is a disclosure option and concentrate on the critical issue of how best to disclose uncertainties in financial statements, several benefits are
possible. The role of the independent auditor will be clarified. Users will have a better basis for evaluating the importance of uncertainties. And—who knows?—some productive research may even be possible.

APPENDIX A—AN EXCERPT FROM THE "COMMISSION ON AUDITORS' RESPONSIBILITIES REPORT"

Recommended Changes in Financial Accounting Standards

The present requirements for disclosure and presentation of uncertainties should be modified. Users should be better informed about the uncertainties involved in the preparation of financial statements, and the information required to be disclosed should be expanded to improve the ability of users to identify and evaluate significant uncertainties.

A separate note, similar to that on accounting policies, should be required for uncertainties. A standardized position and introduction for the note would contribute to user understanding. The note should indicate clearly that the uncertainties are not recognized in the financial statements because they are not susceptible to estimation but could have a material effect on the financial position and future operations. The note, however, should not be a "boilerplate" description of uncertainties and general business risks. It should emphasize the relative importance of the uncertainties through appropriate arrangement and description.

The uncertainties to be disclosed in the note should generally be those contemplated in FASB Statement No. 5—uncertainties that are the results of events or conditions that have already occurred but whose outcome remains uncertain. The note should include for each material uncertainty information required by Statement of Financial Accounting Standards No. 5, such as a description of the circumstances surrounding the uncertainty, management's assumptions about the outcome of the uncertainty, an explanation of the range of possible outcomes and their potential effects, and any other considerations that bear on the probable outcome of the identifiable future event involved. Such a note should not be limited to uncertainties that might have resulted in a "subject to" qualification.

APPENDIX B—LOCKHEED'S 1975 ANNUAL REPORT—SELECTED PARTS

Auditor's Report

Board of Directors and Shareholders
Lockheed Aircraft Corporation

We have examined the accompanying consolidated balance sheet of Lockheed Aircraft Corporation at December 28, 1975 and the related consolidated
statements of earnings and retained earnings (deficit) and of changes in financial position for the year then ended. Our examination was made in accordance with generally accepted auditing standards, and accordingly included such tests of the accounting records and such other auditing procedures as we considered necessary in the circumstances. We have previously made a similar examination of the consolidated financial statements for the prior year.

As discussed under "Uncertainties" in Note 2, the recovery of TriStar deferred charges aggregating $502.5 million at December 28, 1975 and the ultimate loss from certain TriStar related obligations are dependent on future developments. As discussed in Note 13, a determination cannot be made at this time of the outcome of (1) various disputes and other legal proceedings under certain ship construction contracts, and (2) disclosures of commissions and other payments by the Company and related administrative and legal proceedings.

In our opinion, subject to the effects of such adjustments, if any, as might have been required if the outcome of the matters referred to in the preceding paragraph were known, the statements mentioned above present fairly the consolidated financial position of Lockheed Aircraft Corporation at December 29, 1974 and December 28, 1975 and the consolidated results of operations and changes in financial position for the years then ended, in conformity with generally accepted accounting principles which, except for the changes, with which we concur, in the methods of accounting for TriStar costs (see Note 2) and retirement plan costs (see Note 7), have been applied on a consistent basis during the period.

ARTHUR YOUNG & COMPANY

Los Angeles, California
March 30, 1976

Capital Stock

Listed:
New York Stock Exchange
Pacific Stock Exchange

Transfer Agents:
Chemical Bank
Lockheed Aircraft Corporation

Registrars:
Manufacturers Hanover Trust Company
United California Bank

<table>
<thead>
<tr>
<th>Prices on New York Stock Exchange</th>
</tr>
</thead>
<tbody>
<tr>
<td>1975</td>
</tr>
<tr>
<td>------</td>
</tr>
<tr>
<td>1st quarter</td>
</tr>
<tr>
<td>2nd quarter</td>
</tr>
<tr>
<td>3rd quarter</td>
</tr>
<tr>
<td>4th quarter</td>
</tr>
</tbody>
</table>

158
President's Letter

Through the efforts of our 55,000 employees, Lockheed's achievements in 1976 were notable in many ways. While our annual report normally focuses on our individual products—products that serve the nation and the world—this report pays tribute to Lockheed's greatest resource—its people.

These are the people to whom we owe so much—the people in a wide range of skills and disciplines who make it all happen by designing, developing, producing, marketing, and supporting Lockheed's products.

Their achievements were of major significance in the improving trend of the corporation's financial condition. Operating profits on programs other than the TriStar exceeded $240 million for the second year in a row. We repaid $108 million in debt, entered into new long-term bank credit arrangements, and accomplished a financial restructuring plan which transferred $50 million of bank debt to equity. Total equity increased to $167 million.

The strength of the corporation rests on its technological leadership and on the business that flows to us in recognition of such excellence. The sale of the Canadian CP-140 version of the Orion antisubmarine warfare plane to Canada represented one of our largest-ever international sales. Several multimillion dollar international programs, particularly the rapidly expanding activities of Lockheed Aircraft Service Company, have brought a once small element of the corporation's activity into prominence and significant potential as a contributor to our long-term financial growth. In total, we signed up new business of $3.4 billion, and while funded backlog at year-end was approximately $4.4 billion, recognition of unfunded but otherwise negotiated contractual effort increases the total business backlog to over $6 billion.

Despite substantial financial losses on the TriStar program in 1976, several significant developments brightened long-term TriStar prospects. British Airways' order launching the long-range L-1011-500 achieved our longstanding goal of introducing a 6,000-mile range intercontinental TriStar, broadening the TriStar family and making it much more competitive. Certain technical achievements associated with the current TriStar family hold real promise of additional derivatives, opening even broader markets. Although our projected near-term production rate ranges from six to twelve per year, we did resell several aircraft that had been in storage, and we continue to anticipate higher future production rates extending through the late 1980s.

Our role as a leading contractor to the United States government remains a dominant factor in the corporation's business. Particularly noteworthy, the Trident I missile, our largest defense program, made good progress during 1976 and moved ahead into fully successful flight testing early this year.

Production of the P-3 Orion is projected well into the 1980s. While current production contracts for the S-9A Viking are scheduled for completion early next year, we are continuing Viking derivative development work under Navy contract. Current developments in national defense policy put new emphasis on air mobility, which should favorably impact the future of our airlifter programs. Successful classified programs in space and other fields
also reinforce our confidence that our historic role as a defense contractor will remain sound.

The issue of questionable overseas payments in prior years has brought deep concern to all of us who are a part of this fine corporation. The report of our special review committee is due to be released later this spring, and it is possible that some of our customers may be deferring decisions on potential sales pending its release. We have adopted new policies committing us to rigid standards of business conduct and have taken steps to improve the corporation’s reputation in the world community.

We are sensitive to the fact that for a prolonged period we have been unable to pay dividends to our common shareholders. Our loan agreement as well as our current financial position constrain us from making such payments. We recognize a responsibility to resume dividend payments as soon as we can prudently do so. When that will be, we cannot at this time predict.

On behalf of the board of directors and the Lockheed management, we express our appreciation for the support of our shareholders, employees, banks, customers, and suppliers. With this continued support, Lockheed’s potential is as great and as exciting as it ever has been in the past.

Sincerely,

ROY A. ANDERSON
Vice Chairman of the
Board and Chief Financial
and Administrative Officer

LAWRENCE O. KITCHEN
President and
Chief Operating Officer

March 9, 1977

Financial Review

TriStar Program
(Note 2 to the consolidated financial statements provides an explanation of the status of this program, including a discussion of the uncertainties, the accounting, the inventory, and management’s assessment of the program. For a full understanding of the program’s status, this note should be read in its entirety.)

TriStar deliveries in 1976 totaled 16, and sales including spare parts were $431 million. Comparable figures in 1975 were 25 and $559 million. At year-end, backlog numbered 70 units—24 unfilled firm orders and 46 second-buy orders. The accompanying table details TriStar order status for 1976 and 1975.

During 1976, three second-buy orders were converted to firm orders by Delta Air Lines. Five new firm orders were received. Two of the new orders were from Saudi Arabian Airlines. The other three were for TriStars, pro-
TriStar Program Order Status

<table>
<thead>
<tr>
<th></th>
<th>1976</th>
<th>1975</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Firm Orders</td>
<td>Second Buys</td>
</tr>
<tr>
<td>Beginning of year</td>
<td>52**</td>
<td>49</td>
</tr>
<tr>
<td>Sign-ups</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>Cancellations, etc.</td>
<td>—</td>
<td>(3)</td>
</tr>
<tr>
<td>Transfers from second buys</td>
<td>3</td>
<td>(3)</td>
</tr>
<tr>
<td>Deliveries</td>
<td>(16)</td>
<td>—</td>
</tr>
</tbody>
</table>

Backlog at year-end: 24** 46  32** 49

*Second-buy orders have minimal down payments.
**Including three subject to certain conditions.

Produced for but not accepted by Pacific Southwest Airlines in 1975, which were ordered by the West German inclusive tour carrier LTU (Luftverkehr Unternehmen) under an agreement that calls for us to accept in trade two L-1011s that have been in LTU service. Also, one of our current customers ordered, on a combination lease and buy arrangement, two additional TriStars for which we have assumed lease costs following the bankruptcy of the previous operator in 1975.

Cancellation dates of second-buy orders for 20 TriStars were extended in 1976, and Air Canada canceled second-buy orders for three aircraft.

Of significance was the launching of the Dash 500 long-range version of the TriStar, a potential wide-bodied replacement for aging 707s and DC-8s on long-distance routes. British Airways started the L-1011-500 production program with the conversion of six firm and three second-buy L-1011-1 orders to Dash 500 orders, and added three second-buy orders for the Dash 500.

Not included in year-end TriStar firm orders is an additional conversion of a second-buy to a firm order, made by Delta Air Lines in February of this year.

TriStar operating loss in 1976 was $124.8 million compared with $93.8 million in 1975. An analysis of the operating results is shown in the accompanying table.

Production slowed from an annual rate equivalent to 24 aircraft in 1975 to the equivalent of 10 in 1976, and current market projections indicate a production level of six to 12 aircraft annually in near-term years. As a result of the slowdown in production, $38 million in standby production costs was charged to earnings in 1976, and we expect such costs to remain at approximately the same level in 1977. Standby production costs are associated with the maintenance of a manufacturing capability in excess of the needs of the current level of production.

Further, TriStar results for 1976 were adversely affected by volume-related subcontractor price increases applicable to aircraft previously sold and the interaction of costs with the mix of customer orders (including initial program orders) applicable to aircraft lots currently in production.
Operating Results of TriStar Program
(Dollars in millions)

<table>
<thead>
<tr>
<th></th>
<th>December 26, 1976</th>
<th>December 28, 1975</th>
</tr>
</thead>
<tbody>
<tr>
<td>Net sales</td>
<td>$431.0</td>
<td>$559.3</td>
</tr>
<tr>
<td>Cost of sales</td>
<td>426.3</td>
<td>558.6</td>
</tr>
<tr>
<td>Standby production costs</td>
<td>38.0</td>
<td>17.1</td>
</tr>
<tr>
<td>Amortization of deferred charge</td>
<td>50.0</td>
<td>12.5</td>
</tr>
<tr>
<td>Development costs</td>
<td>5.9</td>
<td>4.9</td>
</tr>
<tr>
<td>Costs related to resale obligations</td>
<td>—</td>
<td>10.9</td>
</tr>
<tr>
<td>General and administrative expense</td>
<td>33.6</td>
<td>49.1</td>
</tr>
<tr>
<td></td>
<td>555.8</td>
<td>653.1</td>
</tr>
</tbody>
</table>

Operating loss $124.8 $98.8

Commencing with the fourth quarter of 1975, deferred initial planning and tooling and unrecovered production start-up costs are being amortized in equal quarterly amounts through 1985, and cost of sales is being charged with the actual production costs of aircraft and spares delivered. Deferred costs amounted to $453.3 million at year-end 1976. Recovery of these deferred costs is dependent on the number of aircraft ultimately sold, continuity and rate of production, and actual selling prices and costs.

We believe there is a potential market for more than 800 TriStar aircraft, including the 138 already delivered through year-end. Production and deliveries are expected to extend into the late 1980s, and gross profit on future deliveries of aircraft is expected to exceed deferred costs. Should future assessments indicate that unamortized deferred costs cannot be recovered through future gross profits, the portion so determined would be charged immediately to earnings.

There are many uncertainties surrounding the TriStar program, including those involving projections of sales and costs and the underlying assumptions as to future events. Consequently, estimates and assessments regarding the TriStar program are subject to periodic revision.

Financial Restructuring and Debt

Completion in 1975 of the first phase of a financial restructuring program with our lending banks and the government’s Emergency Loan Guarantee Board resulted in an extension of our 1971 credit agreement and a reduction of interest to 4% on $400 million of nonguaranteed bank loans.

The remaining two phases of the original plan were replaced by a new plan, placed in effect on October 27, 1976, after approval by shareholders and debenture holders.

This plan provided for the conversion of $50 million of nonguaranteed loans into preferred stock and placed financing of the remaining $350 million of nonguaranteed bank debt on a long-term basis, extending through March 1981. It also provided for the issuance to our banks of 10-year warrants to
purchase 1.25 million shares of Common Stock at $7 per share and 500,000 shares at $10 per share. This was in addition to 10-year warrants to purchase 1,750,000 shares of Common Stock at $7 per share issued in connection with the first phase of the financial restructuring plan completed in 1975.

In addition to reduction of nonguaranteed loans through this new recapitalization plan, we paid back during the year $95 million of government-guaranteed bank loans.

Total bank debt at December 26, 1976 was $450 million, consisting of $100 million guaranteed by the U.S. government and $350 million of long-term nonguaranteed loans. An additional repayment of $20 million on February 18, 1977 reduced borrowings under the guarantee to $80 million. Fees paid or accrued to the government under the guarantee since inception through year-end totaled $24.3 million.

Other than bank debt of $450 million, long-term debt—excluding the current portion but including $125 million of convertible subordinated debentures—amounted to $211 million at year-end.

Additional information on long-term debt is shown in Note 8 to the consolidated financial statements.

Management's Discussion and Analysis of Operations
Sales and program profit (loss) by class of similar products are shown in the following summaries. (The years 1975 and 1976 reflect the changes in accounting described in Note 2 to the consolidated financial statements and also reflect the change in accounting for retirement plan costs adopted in 1975, the income effect of which is not practical to compute.)

Sales of TriStars increased through 1974 as a result of deliveries (17 in 1972, 39 in 1973, and 41 in 1974) and higher average unit prices due to contract escalation provisions. TriStar sales decreased progressively in 1975 and 1976 as a result of fewer deliveries in each succeeding year (25 in 1975 and 16 in 1976).

TriStar losses for 1975 were greater than in 1974 primarily as a result of the 1975 accounting change for the TriStar program explained in Note 2 to the consolidated financial statements. The greater TriStar loss in 1976 reflects the slowdown in production, the effect of volume-related subcontractor price increases applicable to aircraft previously sold, and the interaction of costs with the mix of customer orders (including initial program orders) applicable to aircraft lots currently in production.

Development costs have declined since completion of the original design and testing and at present consist of the cost of design and testing of improvements and the Dash 500 model.

General and administrative costs allocated to the TriStar program, which have been charged against earnings as incurred, decreased by $14 million in 1976 compared with 1975 and by $9 million in 1975 as compared with 1974, because of decreased production hours (the basis for allocation of such costs).

Sales of all other aircraft programs and related services decreased slightly in 1976, primarily because of lower sales on the S-3A Viking aircraft program. Partly offsetting this decrease were increased sales on the C-130 Hercules aircraft program. The improved profit rates on the C-130 aircraft program
### Sales (Dollars in millions)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Aircraft and related</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>services:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TriStar</td>
<td>$431</td>
<td>$559</td>
<td>$811</td>
<td>$730</td>
<td>$302</td>
</tr>
<tr>
<td>All other</td>
<td>1,442</td>
<td>1,458</td>
<td>1,225</td>
<td>1,007</td>
<td>1,170</td>
</tr>
<tr>
<td>Missiles, space,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>propulsion, and</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>electronics</td>
<td>1,210</td>
<td>1,263</td>
<td>1,153</td>
<td>967</td>
<td>905</td>
</tr>
<tr>
<td>Shipbuilding and</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>other</td>
<td>120</td>
<td>107</td>
<td>60</td>
<td>53</td>
<td>96</td>
</tr>
<tr>
<td><strong>Total sales</strong></td>
<td>$3,203</td>
<td>$3,387</td>
<td>$3,279</td>
<td>$2,757</td>
<td>$2,473</td>
</tr>
</tbody>
</table>

### Program Profit (Loss)* (Dollars in millions)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Aircraft and related</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>services:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TriStar</td>
<td>$(125)</td>
<td>$(94)</td>
<td>$(49)</td>
<td>$(70)</td>
<td>$(124)</td>
</tr>
<tr>
<td>All other</td>
<td>202</td>
<td>192</td>
<td>128</td>
<td>100</td>
<td>88</td>
</tr>
<tr>
<td>Missiles, space,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>propulsion, and</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>electronics</td>
<td>62</td>
<td>71</td>
<td>65</td>
<td>67</td>
<td>63</td>
</tr>
<tr>
<td>Shipbuilding and</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>other</td>
<td>(4)</td>
<td>(22)</td>
<td>(17)</td>
<td>(15)</td>
<td>(1)</td>
</tr>
<tr>
<td>**Total program</td>
<td>$135</td>
<td>$147</td>
<td>$127</td>
<td>$82</td>
<td>$26</td>
</tr>
<tr>
<td>profits</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Before other income, interest, taxes, and extraordinary gain.

and on foreign technical assistance contracts accounted for the profit increase in 1976 over 1975. However, there were lower profits on the S-5A Viking as a result of decreased sales and lower profit rates.

Sales and profits in 1975 were greater than in 1974 as a result of increased deliveries of Hercules cargo aircraft (71 in 1975 and 46 in 1974) and increased sales and profits on the S-3A Viking and on foreign contracts for technical assistance.

The amount of profits on missiles, space, propulsion, and electronics work has remained relatively constant. Although sales increased in each of the last four years prior to 1976, there was a change in mix of programs and their related profit rates. Sales and program profits in 1976 were somewhat lower than for the 1975 period. The decreased sales and profits reflect the gradual phasing out of the Poseidon missile program, and a lower level of activity on the current Trident missile program.

The increased sales and reduced loss in the shipbuilding and other category in 1976 over 1975 reflect increased construction effort on two submarine tenders for the U.S. Navy (on a cost-type contract) offset, in part, by no icebreaker deliveries in 1976 and lower sales on other ocean systems work.
Increased sales in 1975 over 1974 in the shipbuilding and other category were attributable to the delivery of the first of two icebreakers to the U.S. Coast Guard. The company reported new ship construction profits of $9 million in 1976, compared with losses of $10.9 million in 1975, $16.1 million in 1974, $14.0 million in 1973, and $.9 million in 1972. These losses are applicable to a bulk sugar carrier delivered in 1973, a ferry for the State of Alaska delivered in 1974, and the two icebreakers for the U.S. Coast Guard noted above, the second of which was delivered in January 1977. These losses resulted from continuing technical, production, and schedule problems and higher than anticipated costs. There are no provisions for additional losses on any new ship construction. The company is involved in a dispute with the U.S. Coast Guard regarding prices to be established for the second icebreaker as described in Note 13 to the consolidated financial statements.

As a result of losses on government programs in 1969 and 1970 and the substantial development and initial tooling and production costs on the TriStar program, long-term debt increased through 1973 and most of 1974. These borrowings have resulted in substantial interest expense that, with higher interest rates in 1974, reached a peak of $103 million in that year. The substantial decrease in interest expense in 1975 and 1976 resulted from decreased borrowings, the generally lower interest rates, and the reduction in interest rate, from prime plus 1% to 4%, on $100 million of long-term debt as explained in Note 8 to the consolidated financial statements.

NOTES TO CONSOLIDATED FINANCIAL STATEMENTS

For years ended December 26, 1976, and December 28, 1975

Note 1: Summary of Significant Accounting Policies
The consolidated financial statements include the accounts of all wholly owned subsidiaries, and material intercompany transactions and accounts have been eliminated. The consolidated financial statements are based on the following accounting policies and practices:

a. Inventories:
   — Stated at the lower of cost or estimated net realizable value;
   — For materials and spare parts, cost represents average cost;
   — Work in process includes direct costs (including commissions on foreign government contracts) and allocable operating overhead;
   — General and administrative expenses are allocated to contracts or programs and:
     — Included in work in process for U.S. and foreign governments fixed price contracts;
     — Excluded from work in process and charged to earnings as incurred for commercial contracts and programs;
     — Sales, both military and commercial, are principally under long-term contracts or programs. Accordingly, the Company's operating cycles are more than one year and inventories related to such contracts and programs are included in current assets.
### CONSOLIDATED BALANCE SHEET

<table>
<thead>
<tr>
<th>Assets</th>
<th>December 26, 1976</th>
<th>December 28, 1975</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Current assets:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cash and equivalents (Note 3)</td>
<td>$140.8</td>
<td>$52.7</td>
</tr>
<tr>
<td>Restricted cash (Note 3)</td>
<td>36.7</td>
<td>5.6</td>
</tr>
<tr>
<td>Accounts receivable (Note 4)</td>
<td>195.8</td>
<td>209.8</td>
</tr>
<tr>
<td>Inventories (including TriStar inventories of $166.0 in 1976 and $204.3 in 1975) (Notes 2, 5, and 8)</td>
<td>405.4</td>
<td>387.5</td>
</tr>
<tr>
<td>Future tax benefits (Note 6)</td>
<td></td>
<td>43.0</td>
</tr>
<tr>
<td>Deferred tax charges (Note 6)</td>
<td>15.1</td>
<td>18.7</td>
</tr>
<tr>
<td>Prepaid expenses</td>
<td>45.0</td>
<td>49.1</td>
</tr>
<tr>
<td><strong>Total current assets</strong></td>
<td>838.8</td>
<td>766.4</td>
</tr>
<tr>
<td><strong>Property, plant, and equipment, at cost (Note 8):</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Land</td>
<td>29.9</td>
<td>29.6</td>
</tr>
<tr>
<td>Buildings, structures, and leasehold improvements</td>
<td>289.5</td>
<td>288.4</td>
</tr>
<tr>
<td>Machinery and equipment</td>
<td>422.4</td>
<td>414.6</td>
</tr>
<tr>
<td><strong>Less accumulated depreciation and amortization</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Net property, plant, and equipment</td>
<td>741.8</td>
<td>732.6</td>
</tr>
<tr>
<td><strong>Future tax benefit, long-term portion (Note 6)</strong></td>
<td>494.5</td>
<td>476.6</td>
</tr>
<tr>
<td>TriStar initial planning and tooling and unrecovered production start-up costs (Note 2)</td>
<td>247.3</td>
<td>256.0</td>
</tr>
<tr>
<td>Other noncurrent assets (net of allowance of $4.2 for doubtful notes receivable) (Note 8)</td>
<td>453.3</td>
<td>502.5</td>
</tr>
<tr>
<td><strong>Noncurrent assets and deferred charges:</strong></td>
<td>1,585.9</td>
<td>1,573.4</td>
</tr>
<tr>
<td>Liabilities and Shareholders' Equity</td>
<td>December 26, 1976</td>
<td>December 28, 1975</td>
</tr>
<tr>
<td>----------------------------------------------------------</td>
<td>-------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td><strong>Current liabilities:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accounts payable</td>
<td>$195.7</td>
<td>$220.4</td>
</tr>
<tr>
<td>Salaries and wages</td>
<td>111.7</td>
<td>102.2</td>
</tr>
<tr>
<td>Income taxes (Note 6)</td>
<td>9.8</td>
<td>7.1</td>
</tr>
<tr>
<td>Other taxes</td>
<td>36.2</td>
<td>32.1</td>
</tr>
<tr>
<td>Customers' advances in excess of related costs</td>
<td>208.5</td>
<td>111.4</td>
</tr>
<tr>
<td>Retirement plan contribution (Note 7)</td>
<td>63.1</td>
<td>62.8</td>
</tr>
<tr>
<td>Other liabilities</td>
<td>116.8</td>
<td>115.0</td>
</tr>
<tr>
<td>Current portion of long-term debt (Note 8)</td>
<td>111.8</td>
<td>18.4</td>
</tr>
<tr>
<td>Total current liabilities</td>
<td>853.6</td>
<td>669.4</td>
</tr>
<tr>
<td><strong>Deferred income tax—long term</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Note 6)</td>
<td>4.4</td>
<td>15.7</td>
</tr>
<tr>
<td>Long-term senior debt (Note 8)</td>
<td>436.2</td>
<td>688.0</td>
</tr>
<tr>
<td>4½% Convertible Subordinated Debentures (Note 8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Commitments and contingencies (Notes 2, 7, 11, 12, and 13)</td>
<td>125.0</td>
<td>125.0</td>
</tr>
<tr>
<td><strong>Shareholders' equity</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Notes 2, 8, 9, and 13):</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$9.50 Senior Preferred Stock, $1 par value</td>
<td></td>
<td>47.4</td>
</tr>
<tr>
<td>Common Stock, $1 par value</td>
<td>11.4</td>
<td>11.4</td>
</tr>
<tr>
<td>Additional capital</td>
<td>88.1</td>
<td>82.5</td>
</tr>
<tr>
<td>Retained earnings (deficit)</td>
<td>19.8</td>
<td>(18.6)</td>
</tr>
<tr>
<td>Total shareholders' equity</td>
<td>166.7</td>
<td>75.3</td>
</tr>
<tr>
<td><strong>Total shareholders' equity</strong></td>
<td>$1,585.9</td>
<td>$1,573.4</td>
</tr>
</tbody>
</table>
## CONSOLIDATED STATEMENT OF EARNINGS

<table>
<thead>
<tr>
<th></th>
<th>December 26, 1976</th>
<th>December 28, 1975</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sales</td>
<td>$3,202.7</td>
<td>$3,387.2</td>
</tr>
<tr>
<td>Costs and expenses</td>
<td>3,068.2</td>
<td>3,239.9</td>
</tr>
<tr>
<td>Program profit</td>
<td>134.5</td>
<td>147.7</td>
</tr>
<tr>
<td>Interest and other income</td>
<td>13.2</td>
<td>10.4</td>
</tr>
<tr>
<td></td>
<td>147.7</td>
<td>157.7</td>
</tr>
<tr>
<td>Interest expense ($15.2 million savings in 1976 and $13.4 million in 1975 as a result of reduction in interest rate, Note 8)</td>
<td>54.3</td>
<td>67.5</td>
</tr>
<tr>
<td>Earnings before income taxes</td>
<td>93.4</td>
<td>90.2</td>
</tr>
<tr>
<td>Provision for income taxes (Note 6)</td>
<td>54.7</td>
<td>44.9</td>
</tr>
<tr>
<td>Net earnings for the year (Notes 2 and 13)</td>
<td>38.7</td>
<td>45.3</td>
</tr>
<tr>
<td>Preferred stock—dividend requirement and provision for redemption value</td>
<td>(1.1)</td>
<td></td>
</tr>
<tr>
<td>Net earnings applicable to Common Stock</td>
<td>$37.6</td>
<td>$45.3</td>
</tr>
<tr>
<td>Earnings per share of Common Stock (Note 10):</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary</td>
<td>$3.10</td>
<td>$3.86</td>
</tr>
<tr>
<td>Fully diluted</td>
<td>$2.80</td>
<td>$3.49</td>
</tr>
</tbody>
</table>

†TriStar and other program profit (loss):

<table>
<thead>
<tr>
<th></th>
<th>1976</th>
<th>1975</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sales</td>
<td>TriStar $431</td>
<td>Other $2,772</td>
</tr>
<tr>
<td>Costs and expenses</td>
<td>556</td>
<td>2,512</td>
</tr>
<tr>
<td>Program profit (loss)</td>
<td>(125)</td>
<td>250</td>
</tr>
<tr>
<td></td>
<td>(94)</td>
<td>241</td>
</tr>
</tbody>
</table>

See Note 2 for accounting for the TriStar Program.
### CONSOLIDATED STATEMENT OF SHAREHOLDERS' EQUITY

<table>
<thead>
<tr>
<th></th>
<th>Dollars in millions</th>
<th>$9.50 Senior Preferred Stock</th>
<th>Common Stock</th>
<th>Additional Capital</th>
<th>Retained Earnings (Deficit)</th>
</tr>
</thead>
<tbody>
<tr>
<td>At December 29, 1974</td>
<td></td>
<td></td>
<td></td>
<td>$11.4</td>
<td>$79.0</td>
</tr>
<tr>
<td>Fair value of warrants issued</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Net earnings</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>At December 28, 1975</td>
<td></td>
<td></td>
<td>$11.4</td>
<td>82.5</td>
<td>(18.6)</td>
</tr>
<tr>
<td>Issuance of $9.50 Senior Preferred Stock</td>
<td></td>
<td>$47.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fair value of warrants issued</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Provision for preferred stock redemption value</td>
<td></td>
<td></td>
<td></td>
<td>0.3</td>
<td></td>
</tr>
<tr>
<td>Net earnings</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>At December 26, 1976</td>
<td>(Notes 2, 8, 9, and 13)</td>
<td>$47.4</td>
<td>$11.4</td>
<td>$88.1</td>
<td>$19.8</td>
</tr>
</tbody>
</table>


# CONSOLIDATED STATEMENT OF CHANGES IN FINANCIAL POSITION

## Source of working capital:

<table>
<thead>
<tr>
<th>Description</th>
<th>Dollars in millions</th>
<th>Year Ended</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operations:</td>
<td></td>
<td>December 26, 1976</td>
</tr>
<tr>
<td>Net earnings for the year</td>
<td>$ 38.7</td>
<td>$ 45.3</td>
</tr>
<tr>
<td>Add charges against earnings not involving working capital:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depreciation and amortization of plant and equipment</td>
<td>39.5</td>
<td>39.8</td>
</tr>
<tr>
<td>Amortization of TriStar initial planning and tooling and recovered production start-up costs</td>
<td>50.0</td>
<td>47.3</td>
</tr>
<tr>
<td>Other</td>
<td>(8.5)</td>
<td>8.0</td>
</tr>
<tr>
<td></td>
<td><strong>119.7</strong></td>
<td><strong>140.4</strong></td>
</tr>
<tr>
<td>Issuance of preferred stock</td>
<td>47.1</td>
<td></td>
</tr>
<tr>
<td>Fair value of warrants issued</td>
<td>5.6</td>
<td>3.5</td>
</tr>
<tr>
<td>Current portion of future tax benefit</td>
<td>9.1</td>
<td>55.5</td>
</tr>
<tr>
<td></td>
<td><strong>181.5</strong></td>
<td><strong>199.4</strong></td>
</tr>
<tr>
<td>Additions to property, plant and equipment</td>
<td>33.0</td>
<td>39.4</td>
</tr>
<tr>
<td>Reduction of long-term debt</td>
<td>251.8</td>
<td>18.7</td>
</tr>
<tr>
<td>Other</td>
<td>8.5</td>
<td>5.8</td>
</tr>
<tr>
<td></td>
<td><strong>293.3</strong></td>
<td><strong>63.9</strong></td>
</tr>
</tbody>
</table>

## Application of working capital:

<table>
<thead>
<tr>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Additions to property, plant and equipment</td>
</tr>
<tr>
<td>Reduction of long-term debt</td>
</tr>
<tr>
<td>Other</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Increase (decrease) in working capital</td>
</tr>
<tr>
<td>--------------------------------------</td>
</tr>
<tr>
<td>Cash</td>
</tr>
<tr>
<td>Accounts receivable</td>
</tr>
<tr>
<td>Inventories</td>
</tr>
<tr>
<td>Deferred tax charges and prepaid expenses</td>
</tr>
<tr>
<td>Total decrease in current assets</td>
</tr>
<tr>
<td>Increases (decrease) in current liabilities:</td>
</tr>
<tr>
<td>Accounts payable</td>
</tr>
<tr>
<td>Customers' advances in excess of related costs</td>
</tr>
<tr>
<td>Accrued expenses and other liabilities</td>
</tr>
<tr>
<td>Current portion of long-term debt</td>
</tr>
<tr>
<td>Total increase (decrease) in current liabilities</td>
</tr>
</tbody>
</table>
b. Depreciation:
   — Principal method used to depreciate plant and equipment is the double
decreasing-balance method during the first half of the estimated useful life and
straight-line thereafter.

c. Research and Development Costs:
   — Research and development costs contractually covered are included in
   inventory;
   — Development costs directly related to specific production programs are
   charged to earnings as incurred;
   — Independent research and development costs are allocated to all contracts
   and programs as part of general and administrative expenses.

d. Sales and Earnings:
   — Under cost reimbursement type contracts, sales are recorded as costs are
   incurred and include applicable fees in the proportion that costs incurred
   bear to total estimated costs;
   — Sales under commercial aircraft programs are recorded at contract sales
   prices as aircraft are delivered, and cost of sales is charged with the actual
   production costs of the delivered aircraft. See Note 2 for a description of
   accounting for the TriStar;
   — Under all other contracts, sales are recorded on deliveries or completion of
   contractually specified tasks, as applicable, and the estimated profit on each
   contract is taken into earnings in proportion to recorded sales;
   — The Company's contracts involve numerous price adjustments for specifica-
   tion changes, performance incentives or penalties, and other matters. The
   amount of many such adjustments is estimated prior to final determination.
   Sales and operating profits are recorded on the basis of such estimates. When
   such matters are finally determined, any resulting adjustments are reflected in
   earnings in the year of determination. Performance profit incentives or
   penalties provided for in contracts are taken into account in estimating sales
   and profit rates to be recorded only when there is sufficient information to
   assess anticipated contract performance;
   — Any anticipated losses on contracts or programs are charged to earnings
   when identified.

e. Income Taxes:
   — General and administrative expenses are taken as a tax deduction in the
   year incurred;
   — Tooling, production, and other recurring costs (including costs to com-
   plete) applicable to commercial aircraft are amortized over current year's
   deliveries and remaining production required for firm orders for tax pur-
   poses;
   — Fixed price shipbuilding contracts are accounted for on a completed con-
   tract basis for tax purposes;
   — Investment tax credits are recognized on the flow-through method.

Note 2: L-1011 TriStar Program

Uncertainties. Management believes there is a continuing market for over
500 aircraft (including 188 already delivered) of the basic TriStar model and
proposed derivatives and improvements (which should not involve signifi-

172
cant development costs), with production and deliveries now expected to extend into the late 1980s. The Company's projections of sales and costs are based on many underlying assumptions as to future events, including those concerning the U.S. and world economies, aircraft and other prices, cost performance, production rates, labor performance improvements, inflation, competition, and foreign exchange rates. Projected sales and manufacturing costs take into account expected price level increases. All of these factors are subject to variations and many of them are beyond the Company's control. Consequently, these factors cannot be quantified with precision, and estimates are subject to periodic revision.

Interruption of the TriStar production line for any appreciable period of time could make it economically impracticable to resume production and, in turn, could severely affect the Company's financial condition.

TriStar initial planning and tooling costs and unrecovered production start-up costs are being amortized ratably through 1985 commencing with the fourth quarter of 1975 (see 1975 Accounting Change below). Amortization amounted to $12.5 million in 1975 and $50 million in 1976. Although management believes that gross profit on future TriStar deliveries will be sufficient to recover the remaining unrecovered start-up production and initial planning and tooling costs at December 26, 1976, recovery of such costs is dependent on the number of aircraft ultimately sold (firm and second-buy orders on hand, plus additional orders, less cancellations), continuity and rate of production, and actual selling prices and costs. If future assessments indicate that any such unamortized costs would not be recoverable, the Company would be required to charge to earnings immediately any such costs determined to be nonrecoverable.

Additional uncertainties arise from disclosures of commissions and other payments by the Company referred to in Note 13.

1975 ACCOUNTING CHANGE. In recognition of increased uncertainties, the Company decided during the third quarter of 1975 to recognize no gross profit in earnings on TriStars and spare parts delivered during the first three quarters of 1975. The sales prices thereof exceeded their production costs during that period by $34.8 million, which served to reduce the carrying amount of initial planning and tooling costs and unrecovered production start-up costs of delivered aircraft. Commencing with the fourth quarter of 1975, the Company changed its method of accounting for the TriStar to amortize such costs to earnings ratably through 1985, so long as studies continue to indicate they are recoverable.

In conjunction with this change, management also changed its accounting for TriStar production costs to charge cost of sales commencing with the fourth quarter of 1975 with the actual production costs of aircraft and spare parts delivered; costs remaining in inventories represent the production cost of aircraft in inventory or in process of manufacture.

The effects of the accounting changes made in the fourth quarter of 1975 were to increase the TriStar operating loss by $10.2 million and decrease net earnings by $7.8 million ($0.66 per share) in the quarter and year ended December 28, 1975.
RECENT DEVELOPMENTS. During 1976 new firm orders were received for eight TriStars, three of which were conversions of second-buy orders. In addition, new second-buy orders for three TriStars were received and second-buy orders for three other TriStars were canceled. In 1976, 16 TriStars were delivered as compared to 25 TriStars in 1975. As anticipated, production rates decreased in 1976 and are expected to remain in the near term at a rate of 6-12 aircraft annually.

In 1976 British Airways converted firm orders for six airplanes and second-buy orders for three airplanes from the basic TriStar to the new Dash 500 extended-range TriStar and added three second-buy orders for Dash 500 TriStars. (These are the three second-buy orders mentioned in the paragraph above.) The first extended range TriStars are scheduled for delivery in 1979. Development costs related to the extended range TriStar are being charged against earnings as incurred. Initial planning and tooling costs related to this aircraft (approximately $8 million in 1976) have been deferred and will be amortized commencing with the delivery of the first extended range TriStar.

In 1975 a customer notified the Company that it was unable to accept delivery of three substantially completed TriStars which were on firm order. In January 1977 the Company executed a sales contract with a charter airline for the purchase of these three aircraft. The three TriStars are scheduled for delivery in 1977. The Company will receive two used TriStars in partial payment for the three new TriStars.

The Company became obligated in 1975 to locate purchasers or lessees for two TriStars originally delivered in 1976. In connection with this matter, the Company charged fourth quarter 1975 earnings with $10.9 million. In December 1976 the Company signed a contract with a TriStar customer for the sale of one aircraft to be delivered in 1977 and the lease of the second aircraft with an option to purchase in 1978.

TRISTAR OPERATING RESULTS. The following information summarizes TriStar operating results (in millions of dollars):

<table>
<thead>
<tr>
<th></th>
<th>Year Ended</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>December</td>
<td>December</td>
</tr>
<tr>
<td></td>
<td>26, 1976</td>
<td>28, 1975</td>
</tr>
<tr>
<td>Net Sales</td>
<td>$431.0</td>
<td>$559.3</td>
</tr>
<tr>
<td>Cost of sales</td>
<td>426.8</td>
<td>558.6</td>
</tr>
<tr>
<td>Standby production costs</td>
<td>38.0</td>
<td>17.1</td>
</tr>
<tr>
<td>Amortization of deferred charges</td>
<td>50.0</td>
<td>12.5</td>
</tr>
<tr>
<td>Development costs</td>
<td>5.9</td>
<td>4.9</td>
</tr>
<tr>
<td>Costs related to resale obligations</td>
<td></td>
<td>10.9</td>
</tr>
<tr>
<td>General and administrative expenses</td>
<td>35.6</td>
<td>49.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>555.8</td>
</tr>
<tr>
<td>Operating loss</td>
<td>$(124.8)</td>
<td>$(93.8)</td>
</tr>
</tbody>
</table>

Fourth quarter 1976 TriStar results were affected by volume-related subcontractor price increases applicable to aircraft previously sold and the interaction of costs with the mix of customer orders (including initial program orders) applicable to aircraft lots currently in production.
Because of the TriStar production slowdown commencing in 1975, $17.1 million of related standby production costs were charged against earnings in the fourth quarter of 1975, and $38.0 million in the year 1976. A substantial portion of such costs is associated with the maintenance by the Company of a manufacturing capability in excess of the needs of the current level of production and is not considered to add value to units in process.

The Company expects TriStar operating losses of comparable amounts in near-term years, after deducting related standby production costs, development costs, general and administrative expenses, and amortization of deferred charges.

**Inventories and Orders.** Through December 26, 1976, a total of 138 TriStar commercial jet transports had been delivered. At December 26, 1976, there remained 24 unfilled firm orders and 46 second-buy orders. Firm orders for three aircraft are conditioned upon receipt by the buyer of foreign government approval and financing. Second-buy orders have minimal down payments that are retained by the Company if the order is canceled by the buyer.

TriStar inventories were as follows (in millions of dollars):

<table>
<thead>
<tr>
<th></th>
<th>December 26, 1976</th>
<th>December 26, 1975</th>
</tr>
</thead>
<tbody>
<tr>
<td>Completed and substantially complete aircraft</td>
<td>$55</td>
<td>$115</td>
</tr>
<tr>
<td>Work in process</td>
<td>168</td>
<td>251</td>
</tr>
<tr>
<td>Materials and spare parts</td>
<td>57</td>
<td>55</td>
</tr>
<tr>
<td>Advances to subcontractors</td>
<td>26</td>
<td>34</td>
</tr>
<tr>
<td>Gross inventories</td>
<td>306</td>
<td>455</td>
</tr>
<tr>
<td>Less customer advances</td>
<td>140</td>
<td>251</td>
</tr>
<tr>
<td>Net inventories</td>
<td>$166</td>
<td>$204</td>
</tr>
</tbody>
</table>

Customers' advances at December 26, 1976, include $21.9 million of interest-bearing prepayments from an airline customer. These prepayments will be liquidated against deliveries scheduled through 1978. In addition, advances aggregating $16.5 million are secured by specific equipment included in inventories.

**Note 13: Contingencies**

In 1975 the Armed Services Board of Contract Appeals ruled in favor of the Company with respect to the $62 million settlement in 1971 of the Company's claim of $150 million on certain ship construction contracts. In December 1974 the Company was advised by the Navy that material relating to the ship claims had been referred to the Department of Justice for investigation. This investigation is continuing and has been extended to other ship construction contracts. The Department of Defense has advised that it will not implement the decision of the Armed Services Board of Contract Appeals until it has been advised by the Department of Justice that it is proper to do so. The $15 million uncollected balance of the settlement is recorded by the Company as an account receivable.

The Contract Appeals Board of the Department of Transportation has not
yet issued a decision on a dispute with the Coast Guard on the price to be paid to the Company for construction of the second Polar-class icebreaker. The Company believes it is entitled to a price which exceeds the amount which has unilaterally been set by the Coast Guard. Pending the final determination of price, the inventory under the contract is recorded at an amount which the Company believes would be the minimum recovery under the terms of the original letter contract.

A class action was filed in February 1976 in the United States District Court for the Southern District of Texas and subsequently transferred to the United States District Court for the Central District of California by two shareholders of the Company, and another class action was filed in February 1976 in the United States District Court for the Central District of California by another shareholder of the Company. The class actions allege, among other things, violation of the Federal securities laws as to proxy materials and financial and other reports because of omission of material facts regarding alleged corrupt business practices for which plaintiffs seek injunctive relief, but not damages. In both actions the plaintiffs also sued derivatively on behalf of the Company against certain current and former directors and against the Company as a nominal defendant. In these derivative counts, recovery of damages, restitution, interest, and punitive damages is sought.

Proceedings in both cases have been stayed until May 15, 1977. Counsel are unable at this time to express any opinion as to the probable ultimate liability of the Company on the class action counts in either case, but a substantial part of the relief sought in one of the actions is already included in a consent judgment entered in a suit brought against the Company by the Securities and Exchange Commission.

The appointment of a Special Review Committee of nonmanagement members of the Company's Board of Directors to conduct an investigation of foreign payments and other matters specified in the consent judgment was reported in the Company's 1975 annual report to shareholders. The Committee has received an extension to March 31, 1977, to prepare and submit its report to the Board of Directors. The report is to be filed with the United States District Court for the District of Columbia and the Securities and Exchange Commission within 30 days after it has been submitted to the Board, or such further time as is approved by the Commission, and a statement based on the report is to be sent to the Company's shareholders.

The Company is furnishing documents to the Federal Trade Commission and to a Federal grand jury, both of which are investigating foreign payments and related activities by various companies. Employees and former employees of the Company have been subpoenaed to testify before the grand jury.

These inquiries and investigations and the report and statement of the Special Review Committee could result in the identification of additional foreign payments or further disclosure of identifying details of foreign payments and practices. Such disclosures could significantly impair the Company's ability to obtain future orders, including orders for TriStar aircraft. (See Note 2 with respect to the TriStar program.)

The Internal Revenue Service, through its audit and intelligence divisions, is currently examining the Company's Federal income tax returns for 1973 and 1974 and is reexamining the Company's Federal income tax returns for 1970 through 1972, particularly with respect to foreign payments. Although
the Company's income tax losses for the years 1970 through 1978 effectively preclude, in the Company's judgment, the assertion of any income tax deficiencies for those years, the possibility of the assertion of penalties is recognized even though the Company believes such assertion would not be justified. In the Company's income tax returns for 1974 and 1975 foreign payments which could be viewed as questionable were treated as unallowable for income tax purposes.

In May 1976 an action was filed against Lockheed Electronics Company, Inc. (LEC), a wholly owned subsidiary of the Company, and two other defendants in the United States District Court for the Northern District of Illinois alleging $1.8 million damages in each of three antitrust counts in which LEC is named and seeking treble damages, and seeking $1.8 million compensatory damages and $3 million punitive damages in another count in which LEC is named. LEC and its counsel are in the initial stages of discovery in this case.

In June 1976 an action was filed against the Company in the Los Angeles County Superior Court seeking injunctive relief and damages of $9.9 billion on various claims, as well as other general and compensatory damages and punitive damages of $101 million based on allegations of, among other things, breach of a nondisclosure agreement and of other agreements relating to a proprietary process, and fraud and misrepresentation. Counsel for the Company believe that the damages claimed are grossly excessive in light of the allegations of the complaint and the facts known to them, but are otherwise unable to express any opinion as to probable ultimate liability.

The Company is also involved in other litigation and administrative proceedings. While the Company believes that such other claims against the Company, including various civil rights actions and previously reported actions arising from the crash in 1973 of a Government-owned C-5 aircraft, are either without merit or if successful would be covered by insurance or by government contract or are not of material significance, and while the Company has received opinions from counsel concerning various litigation and other proceedings, including those described above, the outcome of litigation and administrative proceedings is always uncertain, and accordingly, no assurance can be given that the Company will not be adversely affected by the litigation and other proceedings in which it is now engaged.

At December 26, 1976, the Company was contingently liable as guarantor of customers' notes aggregating $6.2 million.

Auditor's Report

Board of Directors and Shareholders
Lockheed Aircraft Corporation

We have examined the accompanying consolidated balance sheet of Lockheed Aircraft Corporation at December 26, 1976 and the related consolidated statements of earnings, shareholders' equity and changes in financial position for the year then ended. Our examination was made in accordance with generally accepted auditing standards, and accordingly included such tests of the accounting records and such other auditing procedures as we considered
necessary in the circumstances. We have previously made a similar examination of the consolidated financial statements for the prior year.

As discussed under "Uncertainties" in Note 2, the recovery of TriStar deferred charges aggregating $453.3 million at December 26, 1976 is dependent on future developments. As discussed in Note 13, a determination cannot be made at this time of the outcome of (1) various disputes and other legal proceedings under certain ship construction contracts, and (2) disclosures of commissions and other payments by the Company and related administrative and legal proceedings. In addition to the matters mentioned above, our report on the prior year was also qualified with respect to uncertainties involving certain TriStar-related obligations. These uncertainties have been substantially diminished, and accordingly our qualification with respect thereto is removed.

In our opinion, subject to the effects of such adjustments, if any, as might have been required if the outcome of the matters referred to in the preceding paragraph were known, the statements mentioned above present fairly the consolidated financial position of Lockheed Aircraft Corporation at December 28, 1975 and December 26, 1976 and the consolidated results of operations and changes in financial position for the years then ended, in conformity with generally accepted accounting principles which, except for the change in 1975, with which we concur, in the method of accounting for TriStar costs (see Note 2) have been applied on a consistent basis during the period.

Los Angeles, California
March 9, 1977

ARTHUR YOUNG & COMPANY

COMMON STOCK

Stock Symbol: LK

Listed:
New York Stock Exchange
Pacific Stock Exchange

Transfer Agent:
Chemical Bank
Lockheed Aircraft Corporation

Registrars:
Manufacturers Hanover Trust Company
United California Bank

<table>
<thead>
<tr>
<th>Prices on New York Stock Exchange</th>
</tr>
</thead>
<tbody>
<tr>
<td>1976</td>
</tr>
<tr>
<td>------</td>
</tr>
<tr>
<td>1st quarter</td>
</tr>
<tr>
<td>2nd quarter</td>
</tr>
<tr>
<td>3rd quarter</td>
</tr>
<tr>
<td>4th quarter</td>
</tr>
</tbody>
</table>
Chairman's Letter

Fellow Stockholders:

1977 was another year of substantial progress for Westinghouse.

Earnings again were at a record high. The overall financial position of the Corporation was further strengthened. Each of our three operating companies, as well as Westinghouse Broadcasting, recorded operating profits higher than those of the previous year.

The Industry Products Company once again was the biggest profit contributor. Its record year is a reflection of the confidence Westinghouse has earned among a broad range of customers in such important basic industries as metals, paper, chemicals and mining. Within Industry Products, each of the product groups improved on its 1976 performance, further evidence of the overall strength and balance of this important part of Westinghouse.

The Public Systems Company also had its highest operating profits ever. The company has made great progress in streamlining and strengthening its various business elements.

The Power Systems Company's improved sales were achieved despite the continuing soft market for power generation equipment and the absence of any domestic nuclear plant orders. Sales for transmission and distribution equipment remained high. The power generation group offset a slump in turbine-generator orders by concentrating on the service aspect of its business.

Sales by the nuclear energy systems group were above 1976, with service and fuel fabrication business both strong. Utility customers are confronted with a number of uncertainties as to the future of nuclear power. Many of these are related to questions that must be resolved by government decisions. Among them are nuclear export regulations, environmental controls and the future of the fast breeder reactor program. A review of some of these uncertainties can be found in this report on page 46 in the Management Discussion section.

I would like to emphasize that we still believe that the United States will have to become much more heavily dependent on coal and nuclear power to meet its future energy and economic needs.

Because we do believe nuclear power over the longer term will have an increasingly important role in meeting our energy needs, Westinghouse intends to maintain its technological leadership in this field.

The past year brought us closer to a resolution of the problems involving our uranium supply contracts, over which a number of utility customers have filed lawsuits. As previously reported, three suits involving utilities in Pennsylvania, Ohio, Alabama and Texas were settled out of court last year. We are continuing efforts to negotiate settlements of the other suits. In a Federal Court in Richmond, Virginia, Westinghouse is vigorously defending the suits brought by the remaining utility plaintiffs. Resolving the uranium problem remains one of our primary objectives. A detailed discussion of uranium matters can be found on pages 40 and 41 of this report.

Let me talk about people for a minute.
Two distinguished members of the Board of Directors stepped down in 1977. John W. Simpson, former president of the Power Systems Company and a 40-year employee, had served as a director since 1971, and Max Nokin, of Brussels, Belgium, a director since 1972.

One of the most heartening aspects of my role as chairman has been to witness the enthusiasm with which employs throughout Westinghouse have responded to efforts to strengthen the Corporation in every way possible—by manufacturing more reliable products, by improving our financial position, by making sure we are meeting the needs of our customers, by seeing to it that we live up to our obligations as a corporate citizen.

We have not yet reached every goal we have set for ourselves. We want to do an even better job of getting to know our customers and their problems. We want to achieve higher profit margins. We want to make more progress in our affirmative action programs toward equal employment opportunity.

Still, looking back at the substantial progress we have made over the past three years, I point with pride to the men and women in hundreds of plants, offices and field locations around the world who have contributed to this achievement. This report contains photographs and captions describing just a few of these 140,000 Westinghouse employees worldwide. They all constitute an asset that gives me confidence in the future of Westinghouse.

R. E. Kirby
January 27, 1978

Note 17

Uranium Litigation

The Corporation is presently defending 14 lawsuits by public utility customers alleging breach of uranium supply contracts. These suits followed a notification by the Corporation to its customers in September 1975 that performance was excused under the legal doctrine of commercial impracticability. Three of these lawsuits are in Sweden. A consolidated trial of eight domestic suits is now in progress in the United States District Court for the Eastern District of Virginia. Three other domestic suits have been transferred to such Court for later trial. Two of the suits alleged violations of the antitrust laws. In two cases there are questions as to the existence of any contract to deliver uranium and, in one case, there is a dispute concerning the obligation to dispose of spent fuel. It is expected that the eventual outcome of the spent fuel issue may be impacted by the governmental policies adopted with respect to the reprocessing and storage of spent fuel.

The parties to the domestic suits, pursuant to order of the Court, have submitted settlement proposals to a Court-appointed Special Master, and the Corporation is continuing its efforts to resolve amicably these and the Swedish suits.

The approximate total amount of asserted commitment of uranium in the litigation referred to above is 69 to 85 million pounds with delivery scheduled under the alleged contracts over approximately the next twenty years. The average price under the alleged contracts is approximately $11.00 per pound, assuming application of contract escalation provisions to date, with price adjustment based on industrial indices and not based on changes in the market price of uranium. Recent spot market price quotations have been in
Westinghouse Consolidated Statements of Income and Retained Earnings

Income Statement

(Amounts in thousands) Year Ended December 31

<table>
<thead>
<tr>
<th></th>
<th>1977</th>
<th>1976</th>
</tr>
</thead>
<tbody>
<tr>
<td>Income:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sales</td>
<td>$6,137,661</td>
<td>$6,145,152</td>
</tr>
<tr>
<td>Equity in income from non-consolidated subsidiaries and affiliated companies</td>
<td>34,001</td>
<td>21,050</td>
</tr>
<tr>
<td>Other income</td>
<td>128,019</td>
<td>88,089</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>6,299,681</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6,254,291</td>
</tr>
<tr>
<td>Costs and Expenses:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cost of sales</td>
<td>4,767,338</td>
<td>4,786,186</td>
</tr>
<tr>
<td>Distribution, administration and general</td>
<td>921,392</td>
<td>914,903</td>
</tr>
<tr>
<td>Depreciation</td>
<td>137,917</td>
<td>139,650</td>
</tr>
<tr>
<td>Interest</td>
<td>46,107</td>
<td>52,347</td>
</tr>
<tr>
<td>Income taxes (Note 5)</td>
<td>150,989</td>
<td>130,892</td>
</tr>
<tr>
<td>Minority interest in net income</td>
<td>4,609</td>
<td>7,096</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>6,028,352</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6,031,074</td>
</tr>
<tr>
<td>Income before extraordinary loss</td>
<td>271,329</td>
<td>223,217</td>
</tr>
<tr>
<td>Extraordinary loss—1977 uranium contract litigation settlements, net of income taxes of $18,970 (Note 17)</td>
<td>(20,550)</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Net income</td>
<td>$250,779</td>
<td>$223,217</td>
</tr>
<tr>
<td>Earnings per common share:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income before extraordinary loss</td>
<td>$9.16</td>
<td>$2.54</td>
</tr>
<tr>
<td>Extraordinary loss, net of income taxes</td>
<td>(.24)</td>
<td>—</td>
</tr>
<tr>
<td>Net income per common share</td>
<td>$2.86</td>
<td>$2.54</td>
</tr>
<tr>
<td>Average common shares outstanding</td>
<td>87,329</td>
<td>87,492</td>
</tr>
</tbody>
</table>

Retained Earnings

(Amounts in thousands) Year Ended December 31

<table>
<thead>
<tr>
<th></th>
<th>1977</th>
<th>1976</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retained earnings at beginning of year</td>
<td>$1,380,083</td>
<td>$1,242,341</td>
</tr>
<tr>
<td>Plus: Net income</td>
<td>250,779</td>
<td>223,217</td>
</tr>
<tr>
<td>Less: Dividends paid on preferred stock</td>
<td>612</td>
<td>631</td>
</tr>
<tr>
<td>Dividends paid on common stock</td>
<td>84,664</td>
<td>84,894</td>
</tr>
<tr>
<td>Retained earnings at end of year</td>
<td>$1,545,536</td>
<td>$1,380,033</td>
</tr>
</tbody>
</table>

181
$40 to $45 per pound range. Industrial marketing information currently indicates that purchases of uranium for long-term delivery may require large down payments and specify that the price will be the world market price at the time of delivery or a stated minimum, whichever is the higher at the time of delivery.

Based upon the Corporation's estimates as of September 1975, with certain modifications to reflect changed circumstances as to the asserted commitment, approximately 69 million pounds would be involved in the event of an unfavorable outcome in all the litigation. Included for this purpose is the uranium at issue in the two suits in which Westinghouse is contending that it has no contractual obligation to deliver any uranium. This estimate is based on a number of assumptions which materially affect the quantities involved in the asserted commitments. Among these assumptions are the following: (a) that the Corporation's current position in the litigation with respect to the scope and duration of each of the asserted commitments would prevail, (b) that future United States government policy regarding tails assay requirements for the operation of government-owned uranium enrichment facilities, which has a significant impact on the quantities of uranium involved in several of the asserted commitments, will be in accord with the Corporation's current expectations and (c) that uranium requirements will be reduced in the future by the commercial availability of plutonium recycle.

Several of the plaintiff utilities have claimed, in their original and amended complaints and in other statements, that the Corporation's asserted commitments to them involve substantially greater quantities of uranium than the Corporation has estimated. By the Corporation's current estimate of the plaintiffs' claims of the quantities involved in the asserted commitments, about 85 million pounds of uranium would be involved in the event of an unfavorable outcome of all the litigation.

Under an arrangement confirmed by court order in February 1976 in the consolidated action, the Corporation became obligated to deliver, subject to a later determination of proper prices, the uranium it had in inventory or under contract to the extent received by the Corporation, then estimated to be approximately 15 million pounds. Approximately 46 per cent of the 15 million pounds has since been delivered to certain plaintiff utilities under the court-ordered allocation plan. In connection with the settlements described below, the Corporation has released its right to adjustment of quantities or prices respecting 722,000 pounds.

There has been a dispute concerning the meaning of the court order with respect to the extent of the delivery obligations of the Corporation. In October 1976 a uranium supplier filed suit asking for a declaratory judgment that it is not required to perform its contract for the delivery of 450,000 pounds of uranium. Two other uranium suppliers under contract for about seven million pounds have suggested they may be entitled to price relief. One has completed delivery under protest; the other has continued delivery to date under protest. The Corporation is pursing its contractual rights against these suppliers and believes it is in compliance with the court order.

Certain of the contracts in litigation are also included in a group of fuel fabrication contracts that may require the Corporation to supply, starting about 1982, an additional 3.7 million pounds of uranium for use as diluent in
making plutonium fuel. As to diluent uranium required by these certain contracts, the Corporation has not asserted the doctrine of commercial impracticability. If the plutonium fabrication option in each of those certain contracts should not be exercised for some reason (e.g., government prohibition), the Corporation may be obligated under those certain contracts to supply sufficient uranium to provide equivalent energy, which depending on several variables could require an estimated nine million pounds. (There would be no such additional requirement under the other contracts in the group.) There continues to be uncertainty as to the need for this additional uranium, although in the Corporation's estimate of the 85 million pounds asserted commitments of the plaintiffs, no plutonium recycle was assumed. The absence of plutonium recycle could also limit the alternatives available to the Corporation to discharge any spent fuel obligations it may have to two of its utility customers.

In 1977 the Corporation settled a uranium supply contract suit in a Pennsylvania court involving Duquesne Light Company and two other utilities. The settlement agreement provides that plaintiffs will be entitled to additional benefits if the Corporation settles with any other plaintiff on a more favorable basis. The Corporation also in 1977 settled the Alabama Power Company uranium supply contract suit which had been a part of the consolidated trial.

In December 1977 the Corporation settled the Texas Utilities Services Inc. uranium supply contract suit, which also was a part of the consolidated trial and involved three utilities. Pursuant to the settlement agreement, the Corporation will make a cash payment, will supply certain goods and services at no charge and certain uranium under a deferred payment arrangement and will transfer certain rights to mining properties, production and technology. The settlements with Alabama Power Company and Texas Utilities Services Inc. are not expected to result in payment by the Corporation of additional benefits under the settlement agreement reached with three utilities in the Pennsylvania state court action referred to above.

The amount of uranium involved in the three lawsuits settled to date approximates 4.4 million pounds, including the uranium originally involved in the court-ordered allocation plan which the Corporation has delivered or will deliver at the original contract price.

These settlements, representing a total of about five per cent of all the uranium originally estimated to be in dispute, resulted in 1977 in an extraordinary loss, net of income taxes, of approximately $20.5 million. In the opinion of management and counsel, because of the particular circumstances of these settled cases and of the remaining cases, neither the costs of settlements made to date nor any other basis provides a reliable means for arriving at a reasonable estimate of the aggregate loss that will ultimately be incurred by the Corporation in respect of the remaining lawsuits.

Regardless of the differences among the cases and the other uncertainties, however, the ultimate disposition of the remaining uranium lawsuits will in all likelihood be very costly to the Corporation, and the aggregate loss could be substantially more, or could be less costly to Westinghouse, proportionately, than the loss represented by the cases settled to date. Failing settlement, if the Corporation is not wholly successful and is granted only partial relief
from its alleged contractual obligations, the effect on the Corporation's financial condition will also be very substantial. If the Corporation on the other hand were required to fulfill all the remaining contracts under current market conditions, the financial impact would, of course, be extremely adverse.

The Corporation is also defending against several purported shareholder class actions alleging securities law violations for failure to make proper disclosure of, among other things, the uranium situation. In January 1978 the court decertified a class in one action which had been proceeding as a class action and ordered that another action concerning the uranium situation only proceed as a class action. The court has denied class certification with respect to the remaining actions. All allegations of wrongdoing have been denied. These actions have been consolidated for pretrial purposes in the United States District Court for the Eastern District of New York.

Due to the uncertainties pertaining to the remaining suits described above, eventual outcome cannot be predicted and financial effect cannot reasonably be estimated. Because the minimum loss cannot be reasonably estimated, no liability has been recorded in the consolidated financial statements.

**Note 18**

**Uranium Cartel Antitrust Suit**

In October 1976 the Corporation filed an antitrust suit in the United States District Court for the Northern District of Illinois, Eastern Division, against 29 domestic and foreign uranium producers. The suit seeks treble damages from the defendants as well as injunctive relief from illegal combinations and conspiracies entered into by defendants to restrain both the interstate and foreign commerce of the United States in uranium. To date, five of the defendants have filed antitrust counterclaims and it is possible that others may also file counterclaims. Direct discovery has been stayed pending the court's decision on a motion by certain defendants to disqualify the Corporation's counsel. Although the outcome of litigation is always uncertain, if the Corporation is successful in this suit, the judgment could, depending on size and collectibility, have a favorable financial effect. No accounting recognition has been given to this suit in the consolidated financial statements.

**Note 19**

**Other Litigation**

Eight pending cases by utilities claim damages totaling approximately $2.5 billion in connection with outages of turbine generators sold by the Corporation. Most of the damages claimed are consequential or punitive. To date, the Corporation's contractual disclaimers against consequential damages have been sustained in all resolved cases.

With respect to the suit brought in November 1974, by the San Francisco Bay Area Rapid Transit District (BART) against the Corporation and other defendants for damages in connection with equipment Westinghouse supplied to BART, the parties have reached a tentative agreement in principle in settlement of this suit. Under the terms of the settlement, which is subject to agreement by all parties on final terms and conditions, the Corporation will make a cash payment and release certain claims it has against BART.
Management believes the cost of effecting the settlement approximates the estimated costs and expenses of conducting the litigation to a conclusion. The cost of the settlement has been charged to earnings. The Corporation also reached a tentative agreement with a co-defendant regarding a resolution of certain contractual differences involved in the BART project.

The Corporation, Mitsubishi Electric Corporation and Mitsubishi Heavy Industries, Ltd. are defendants in a civil antitrust action brought by the United States Department of Justice asserting that certain agreements under which the Corporation grants technical assistance and foreign patent rights to the other defendants constitute a conspiracy to restrain trade between the United States and Japan in violation of Section 1 of the Sherman Act. The Corporation denies any violation of law. This case does not involve any claim for money damages, but seeks to require the Corporation to change its practices with respect to the licensing of foreign companies. Trial of the case commenced in May 1976 and, following the conclusion of the government's case in chief, all defendants moved to dismiss. Those motions are pending.

The Corporation is also defending against three shareholder actions alleging securities law violations for failure to make proper disclosure of the sale of the Corporation's major appliance business. All allegations of wrongdoing have been denied. These lawsuits were originally initiated as purported shareholder class actions. The trial court refused to certify plaintiffs' proposed classes, and the appellate court dismissed, as premature, the plaintiffs' appeal from the trial court's decision. Trial of these cases, as individual shareholder actions, commenced in January 1978.

In January 1975 a lawsuit, the subject matter for which the Corporation has assumed liability, arose out of a subsidiary's earlier acquisition of a recreation land development business. As filed, the suit claimed an unspecified amount of damages.

In addition to pending litigation, disputes have arisen between the Sacramento Municipal Utility District (SMUD) and the Corporation concerning responsibility for losses related to turbine and generator outages at SMUD's Rancho Seco nuclear plant located in Sacramento County, California. Also, SMUD has notified the Corporation that it has rejected a claim against it by Pacific Gas & Electric Co. (PG&E) for alleged damages flowing from an agreement for the purchase and sale of electricity into which PG&E was allegedly induced by alleged misrepresentations by SMUD. SMUD has notified the Corporation of its intention to tender to the Corporation the defense of any suit brought against SMUD by PG&E for losses due to those outages. PG&E has recently indicated that it may present a direct claim against the Corporation. In October 1977 the Corporation filed a lawsuit against SMUD seeking damages arising out of repair work performed by the Corporation on a SMUD turbine-generator.

In September 1977 Nordostschweizerische Kraftwerke AG (NOK), a Swiss utility, filed a Declaratory Judgment proceeding in Switzerland claiming damages for alleged deficiencies with respect to steam generators supplied by the Corporation as part of the Beznau II nuclear plant. The Corporation and NOK have commenced settlement negotiations.

The Korean Educational Development Institute is claiming a material amount of damages for an alleged breach of contract arising out of the sale of
a telecommunications balloon and mooring system by a subsidiary of the Corporation. Management believes that the claim will be tried, if at all, by means of arbitration.

The Corporation has settled the antitrust suit brought by four subsidiaries of American Electric Power Company by payment of a portion of the legal fees incurred by the plaintiffs in connection with the lawsuit. The Corporation had previously agreed with 126 other utilities to a suspension of the running of the statute of limitations with respect to any alleged violation of the antitrust laws arising out of their purchases of turbine-generators manufactured by the Corporation. In November 1977 the Corporation gave the notice necessary under those agreements to require the other utilities to sue or be barred. Accordingly, under the agreements, including extensions thereof, these utilities may file suit prior to April 1, 1978. Should any utility file suit within the prescribed period, other utilities will receive an automatic extension of an additional 30 days.

Management believes the Corporation has meritorious defenses to the suits and claims mentioned in this note and that the ultimate outcome thereof will not result in a material adverse effect on operations and financial position.

Note 20
Guarantees
At December 31, 1977 the Corporation was guarantor of customers' notes sold to banks and other liabilities aggregating $128 million.

Note 21
Management Incentive Compensation
The Board of Directors voted to award $10,294,900 to 1,130 employees under the management incentive program for 1977.

Note 22
Replacement Costs and Quarterly Financial Data
Unaudited information regarding asset replacement costs and quarterly financial data appears on page 47.

Report of Independent Accountants
To the Board of Directors and Stockholders
of Westinghouse Electric Corporation

We have examined the consolidated financial statements of Westinghouse Electric Corporation and its subsidiaries appearing on pages 28 through 43. Our examination of these statements were made in accordance with generally accepted auditing standards and accordingly included such tests of the accounting records and such other auditing procedures as we considered necessary in the circumstances.

Note 17 to the consolidated financial statements discusses litigation involving uranium supply contracts with customers, uncertainties regarding the plutonium fabrication option and disposal of spent uranium fuel, and certain purported class actions by shareholders. Notwithstanding the uncertainties which preclude the estimation of a minimum loss and recording of a liability, management and counsel believe that the
ultimate resolution of the uranium litigation will, in all likelihood, be very costly to the Corporation.

In our opinion, subject to the effects of the adjustments as would be required by the resolution of the matters discussed in the preceding paragraph, the financial statements referred to above present fairly the financial position of Westinghouse Electric Corporation and its subsidiaries at December 31, 1977 and 1976, the results of their operations and the changes in financial position for the years then ended, in conformity with generally accepted accounting principles consistently applied.

Price Waterhouse & Co.
600 Grant Street
Pittsburgh, Pennsylvania 15219
January 24, 1978

APPENDIX E—AN EXCERPT FROM "THE CORPORATE COMMUNICATIONS REPORT," VOLUME 3, NUMBER 1, JULY 1977

Funny Stuff in the Tobacco Industry

The six major U.S. cigarette companies all are defendants in two class action anti-trust lawsuits by tobacco growers seeking $379 million and a third class action suit seeking unspecified damages. Yet, the auditors have reacted in varied ways in deciding whether these lawsuits mandate a qualified audit report. Coopers & Lybrand, for instance, has issued a qualified opinion to one client on the basis of the lawsuits but given a clean certificate to another, even though the latter company actually draws a higher proportion of its sales from tobacco products.

<table>
<thead>
<tr>
<th>Company</th>
<th>Percentage of 1976 sales from tobacco products</th>
<th>Auditor</th>
<th>Auditor's opinion qualified in 1975?</th>
</tr>
</thead>
<tbody>
<tr>
<td>American Brands</td>
<td>65</td>
<td>Coopers &amp; Lybrand</td>
<td>Yes</td>
</tr>
<tr>
<td>Liggett Group</td>
<td>48</td>
<td>Haskins &amp; Sells</td>
<td>No</td>
</tr>
<tr>
<td>Loews Corp.</td>
<td>25</td>
<td>Touche Ross &amp; Co.</td>
<td>No</td>
</tr>
<tr>
<td>Philip Morris</td>
<td>70</td>
<td>Coopers &amp; Lybrand</td>
<td>No</td>
</tr>
<tr>
<td>R. J. Reynolds</td>
<td>65</td>
<td>Ernst &amp; Ernst</td>
<td>Yes</td>
</tr>
<tr>
<td>Universal Leaf Tobacco</td>
<td>90</td>
<td>Arthur Young &amp; Co.</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Seeming inconsistencies of this sort have occurred in other industries—the armored carrier business being a notable example. In 1975, the three major companies—Baker Industries (Wells Fargo), Brink's and Purolator—all stated in their footnotes that a federal grand jury in Atlanta was investigating the industry for possible anti-trust violations, but only one of the three companies—Brink's—received a qualified opinion. Yet, each company has the same auditor, Peat, Marwick, Mitchell & Co. In 1976, Peat Marwick changed course and qualified its opinion on all three companies. The auditor's report of a fourth company in the armored truck business—Pittston, whose position is held through 85 per cent ownership of Brink's—was not qualified in either year on the basis of the anti-trust investigation (but was qualified for other reasons). Again, Peat Marwick is the auditor.
Some accountants cite varying lawyers' opinions for inconsistencies of this sort. While the lawyers for one company might be willing to go on record that they believe a lawsuit to be without merit, the lawyers for another might refuse to do so even when the same lawsuit is involved—meaning that the former company could receive a clean certificate and the latter a qualified certificate from the very same accounting firm. And, of course, it is possible that the potential impact of a specific suit falls differently from one company to the next, although this appears to be less a factor than the varying ways of varying lawyers.

Regardless of where the blame lies, qualified opinions clearly are meaningless, perhaps even counter to fair allocation of capital, when two companies with similar legal exposure end up receiving such contradictory treatment at the hands of their auditors.
Discussant's Response to "Subject To' Audit Opinions: A Preliminary Investigation of Statement Users' and Statement Issuers' Perceptions"

ROBERT LIBBY*

The key question of interest to Shank and Dillard, as indicated on page two of their paper, is, "Do readers interpret a "subject to" opinion to convey a message which is significantly different from that in an unqualified opinion?" The need for empirical research into this question as recognized by the authors, myself, and others has been enhanced by the willingness of the Commission on Auditor's Responsibilities to decide the question based on armchair analysis. I approve of the SEC's action to delay AudSEC's decision to delete the "subject to" opinion because I believe that it was premature. So, I applaud Shank and Dillard (and their coauthor, Murdock, on other projects dealing with the same topic) in their attempt to test some of the Commission's hypotheses before they were acted upon by a regulatory body and the proverbial barn door was closed.

However, I believe that there are a number of flaws in the design of the project which affect the ways in which the results may be interpreted. The remainder of the discussion will be organized as follows. First, the way in which the authors attempted to answer each of their five research questions will be evaluated. Second, an interpretation of the results will be made, based on the above evaluation. And third, some miscellaneous methodological comments will be made.

*This paper was completed while Mr. Libby was Price Waterhouse & Co. Accounting Research Fellow at the University of Chicago.
RESEARCH ISSUES

Perceived Usefulness

The study first addresses whether the respondents to the survey “consider a ‘subject to’ opinion to provide useful information (p. 189).” This question is answered, according to the authors, by asking the subjects to indicate the “best” method of disclosure for a number of contingency cases. The authors assume (p. 187) that, if the respondents indicate that a “subject to” is preferred, they also believe that there is a useful distinction between it and the clean opinion. There are two major problems with drawing this inference in this fashion. First, if one looks closely at the questionnaire, while the instructions were consistent with the author’s description in the body of the questionnaire, the subjects were not asked to indicate the “best” method of disclosure, but the one they would choose to communicate the contingency. While this inconsistency may or may not have confused the respondents, I suggest that many would take both questions to be asking them to apply current GAAS. The authors indicated that they share my concern by making a change in the cover letter (sent only with the second requests for responses to a group of CPAs who also received the questionnaire) which specifically stated that this was not the question being asked. A measure of the respondents’ interpretations of GAAS is in no way related to perceived usefulness which is the question of interest.

Second, let us assume—as the authors suggest—that the participants did indicate their preferred disclosure choice without consideration of GAAS. It is still far from clear that a preference for a “subject to” disclosure implies a belief that the auditor’s choice (between footnote disclosure alone and the addition of the “subject to” qualification to the footnote) would affect decisions. It is important to note that the literature is replete with surveys showing disclosure preferences which are followed by experimental studies indicating that the disclosures have no effect on decisions. It is difficult for me to understand why, in place of the need for this grand assumption, the authors did not take a more straightforward approach such as asking the question of interest, “Would your decisions be different if the method of disclosure were different?” The validity of this approach is also open to question, but would appear superior to that used in this study.
Predictive Accuracy

The second issue addressed in the study is whether the respondents understand "the way CPAs use 'subject to' opinions (p. 128)." This question is approached based on the assumption that an inability to predict the type of audit opinion issued based on the information normally presented in a footnote suggests a lack of understanding of the message implied by different reports. Such an assumption presents a number of problems which make the interpretation of the results unclear. First, it is possible for a reader to understand the message the auditor intends to transmit when he issues alternative reports without being able to predict what report will be issued given a set of circumstances. Consider the following analogy. A pathologist has analyzed a tumor biopsy and decided that there is a 95 percent chance that the patient will die within twelve months. Can the patient understand the message implied by the statement, "There is a 95 percent chance you will die within twelve months," without understanding the specific attributes of the biopsy which lead the pathologist to his judgment? I suggest that the answer is "Yes." In a sense, the logic of the Shank and Dillard analysis is backwards. A more relevant comparison might be the congruence between the message the auditor intends and the message the user believes the auditor intends.

A second problem exists in the authors' conceptualization of the process which generates audit reports. For any one case, there is no single appropriate disclosure level in accordance with GAAS. GAAS requires auditors to use their professional judgment in making this determination. Any disagreement between different auditors' judgments or within an auditor's judgments over time will result in a probabilistic process or a distribution of "appropriate disclosures" for every case. Such inconsistencies in application of both accounting and auditing standards have been documented time and time again in the research literature. If a single set of circumstances can result in a number of different disclosures, this places a limit on the maximum predictability obtainable in the survey task. Without knowing this limit, statements about the quality of the respondents' predictions or inferences concerning levels of confusion are not possible. Some of you may take evidence of inconsistencies in application alone as sufficient to suggest that we eliminate the "subject to" report. Let me warn you against this argument on two grounds. First, information which is probabilistically related to some event can be useful. And,
second, if we were to do away with all accounting disclosures which were applied inconsistently in practice, we would report only on cash.

It should be emphasized that not only do inconsistent auditor’s judgments set a limit on prediction accuracy, but also (as admitted by the authors) only the very limited information contained in the footnote was available as a basis for the subjects’ predictions. The combination of these two factors would lead one to expect even the most knowledgeable auditor to be only mildly accurate in his predictions. As a final note on this point, if users could predict audit reports from footnotes, by definition, the information in the “subject to” report would be useless and we would have a rationale for eliminating it.

Litigation vs. Asset Realization and Going-Concern
vs. Less Cataclysmic Proportions
The third issue addressed deals with differences in disclosure preferences for different types of contingencies. The authors state (p. 124) that, in effect, they are testing a 2x2x2 repeated-measures ANOVA design. Consideration of this view illustrates some confounding of variables in these tests. For such tests to be internally valid, the cases must differ only on the variable of interest (e.g., litigation vs. asset realization). As the cases are not comparable on other attributes (e.g., size of possible loss), it is impossible to separate the impact of the variables of interest from these other confounding variables.

Interpretation of the Remaining Results
The above difficulties question the interpretability of the results relating to three of the research issues. Two other issues are relatively unaffected. The first is the effect of the “subject to” on the auditors’ defensive posture. These results are particularly interesting. The Commission decided that the “subject to” would not provide the auditor with any protection against legal action. Their conclusion was based on analysis of two cases which—some have suggested—are only marginally relevant to the question. As financial analysts and financial executives are potential litigants in such situations, their opinions are of interest. The finding of a total contradiction of the Commission’s conclusion should be considered carefully by policy makers.

The second finding of interest is the difference in attitudes between the executives and analysts. As the authors point out, it is not sur-
prising that the analysts are more willing to shift to the auditor and management the responsibility for estimating the outcomes of uncertainties. The authors' second explanation is of interest to those who study the psychology of human judgment. Without realizing it, they suggest that use of the "availability" heuristic would also explain the results. When using this simplified decision rule, or heuristic, people estimate the frequency of an event based upon their ability to recall similar events that they have experienced. As more memorable experiences will be more easily recalled, they overestimate their frequency of occurrence. Given the differing loss functions faced by the analysts and executives we would expect differing events to be the most memorable to the two groups. It seems reasonable to assume that the analyst will remember cases which were widely reported in the press in which audit disclosures have been reported as "sub-standard." Virtually all such cases chastise the auditor for not making additional disclosures. The financial executive, on the other hand, is more likely to remember cases in which he was unable to convince his own auditor that additional disclosure or some write down of an asset was unnecessary. This would result in analysts expecting less disclosure and executives expecting more. The results are consistent with these expectations.

**METHODOLOGICAL PROBLEMS**
Some additional methodological problems are worthy of a short note. First, the difference between the first question asked in the instructions and that asked in the body of the questionnaire may cause confusion and is undesirable. Second, it is normally not considered to be good practice to change instructions in a cover letter sent to only some of the participants (those receiving the third mailing). Third, the "footnote only" disclosure alternative presented to the participants is misleading because the subjects were not told that the footnote would contain the same information disclosed under the "subject to" alternative, only without the auditor's choice of the "subject to" vs. "clean" opinion being offered. Fourth, if subjects have experienced frequencies of the disclosure alternatives which are different from the frequencies represented by the cases, it should result in a decrease in their prediction accuracy due to incorrect prior probability estimates. Fifth and last, it is not clear from the paper why financial executives' judgments are relevant to the questions concerning the usefulness of the disclosures to financial statement readers.
CONCLUDING NOTE

Finally, the following observation is of general relevance to those who commission, conduct, or read the results of surveys. This paper, like most surveys, is at its worst when it attempts to draw inferences concerning the effects of disclosures on decisions. In general, both laboratory and field experimental methods and ex post analyses are more appropriate for this task. On the other hand, the paper, like most surveys, is at its best when it is soliciting expert predictions of the future behavior of third parties (the courts in this case) or estimates of past behavior where archival evidence of past behavior is not sufficient to allow reliable inferences to be made. While this point has been made previously at conferences here at Illinois, it needs to be made again as its significance has yet to be reflected in the accounting literature.
Session

FIVE
Thoughts on Practical Auditing Research

RAY J. GROVES

This evening my remarks are intended to express a sincere concern from one professional to another—and in particular to those of you who are professional educators and accountants with a strong interest in auditing research. My own experience is that of a professional practicing accountant. All of my working career since graduating from The Ohio State University in 1957 has been with Ernst & Ernst working up through the ranks as a staff assistant, senior, supervisor, manager, and partner.

If there has been a major emphasis in my experience, it has been on the technical side of practice. In our organization, the chairman of our accounting and auditing standards committee has final responsibility with our managing partner, for the application of generally accepted accounting principles and generally accepted auditing standards within the entire firm. Having served in that capacity until three years ago, I came into practice administration fairly recently with a strong practical background in meeting and resolving the kinds of technical issues that practicing accountants face every day.

WE HAVE MUCH IN COMMON
As professionals, we have much in common. Each of us has a strong motivation to succeed in our chosen activities. Each of us would include in our definition of success something more than monetary income—some contribution to society, whether that be the intro-
duction of new methods, the reemphasis of old ones, or the effective performance of a service which adds to the well being of our society.

We all appreciate and support the freedom and opportunity of our market economy. We all have a keen interest in accounting and auditing in the long run as well as in terms of our own personal success. I trust that you take as much pride in being an accountant as I do, and that we both see it as an essential part of our complex economic system.

My purpose in making these introductory remarks is first to let you know the experience and philosophical base from which I comment, and second, to point out why you and I might see some things differently, because there are some differences between us.

As described, my background is audit practice. This background represents a world of complex problems requiring immediate resolutions—resolutions that unavoidably impact the personal fortunes and prestige of those involved. It is a world of deadlines, time budgets, control of costs and expenses, and responsibility for fee collection. It is a world in which failure to understand the latest regulations or to provide adequate supervision of assistants, or to note the importance of trends in a client's business or weaknesses in its controls, can lead to the exorbitant burdens of litigation for both an individual practitioner and his firm.

I have seen close up what happens to a decent, competent, conscientious audit professional when his work is called into question in an SEC hearing or in a court of law. Even when his work passes every test, the experience remains a traumatic one.

Audit practice calls for hard decisions, often in situations where right and wrong are not all that clear and where time for meditation is limited, if it exists at all. And then those decisions, made in view of specific facts and conditions, are a matter of record. Then, at a later date if any one of a number of events far beyond the auditor's control should happen, the decisions may be called up for scrutiny when the facts are obscured and the conditions have changed.

**SHARED INTEREST IN AUDITING RESEARCH**

My understanding is that part of the reason you have become educators is your keen interest in research. The subject of this conference is auditing research, and I am pleased to see auditing receive such serious attention. As I pointed out at our recent annual partners' meeting, the development of accounting principles over the past two decades has received so much attention that the development of
auditing has been short changed. We are very hopeful that over the next few years we can all give auditing the attention it warrants. Thus, your interest in auditing research strikes a responsive chord with me. Therefore, I offer some suggestions which may be relevant to you in your research activity.

**EXTENT OF RECENT CHANGES IN AUDITING**

First, there have been remarkable changes in auditing within the last few years. Some of our "older" partners say they have seen more change in the last few years than in their total previous careers. A substantial increase in general interest and concern with auditing has been evident in such matters as the sensitive payments revelations, the discovery of so-called "corrupt practices," the executive perquisites, and the very useful development of corporate audit committees. Twenty years ago, we could practice auditing in peace and quiet, and unless we made a serious mistake, sufficient to arouse the ire of a client company, we had little fear of litigation, serious criticism, or even any significant interest. Now, we find ourselves in the public eye, even attacked by congressional spokesmen.

**THE POSSIBILITY OF REGULATION**

Out of this critical interest in accounting has grown a strong demand for regulation of our profession. Public accounting has always been a self-regulated profession with ethical standards as high as those of any other profession. The profession has had a record of continued progress in its service to clients and has served as an essential element in the functioning of our capital markets and in the growth and efficiency of industry.

Certainly, additional progress is needed, and I see strenuous efforts to make the profession more responsive and responsible. Yet it may well be that we had become complacent, and that at least some of the recent criticisms have been constructive and productive in encouraging us to reconsider our goals and standards. But does that mean that we should become a regulated profession? Would our economy be better served by certified government accountants than by certified public accountants?

What will happen to professional accounting if it ceases to be a self-regulated activity? I do not know, and I do not think anyone else does. We occasionally surmise what the effect might be and we are much concerned. We observe the drag on business which many of the recent congressional regulatory efforts have become, and we all hope that
will not happen to accounting. Personally, I have very strong faith in private enterprise. I believe that competition, combined with the profit motive, brings great benefits in innovation and efficiency. Competition among professionals is not the same as competition within industry, but if the differences are recognized realistically, it can result in some of the same social benefits. Competition in auditing has three aspects—quality of service, fees, and financial reporting standards.

My view is that we should be expected to compete with other firms as vigorously as possible on the basis of quality of service, and Ernst & Ernst is prepared to do so. We mistrust competition solely on a fee basis because reduction of fees can so easily lead to reduced quality. The line between a truly efficient audit and one that is economized beyond the point of adequate scope is a fine one. Cutting corners to meet a time budget or to please a client is an almost unavoidable reaction to fee pressures. The ultimate results of excessive price competition are detrimental to the profession and to the economy.

We should and do refuse to compete on the basis of principles. Only a very few companies propose that their CPAs must compromise their principles to hold the company as a client. The impression that clients constantly call us and threaten to transfer to another public accounting firm if we do not go along with their chosen procedure is a false one. Rarely does this happen. Certainly, we have differences of opinion with some of our clients, and we must be prepared to defend our views. Such disagreements typically occur not as blatant disregard for standards but as issues subject to more than one interpretation which have not yet been resolved at the authoritative standard-setting level.

DIFFICULTIES IN AUDIT RESEARCH
A second thought I would share with you tonight is an expression of sympathy for the difficulty of the task you face as researchers. Speaking with relatively little knowledge of research methodology but with an intimate knowledge of the nature of auditing, I would assume that audit practice is a difficult subject to research effectively. We are required by our professional ethics, by the rights of our clients, and by our own best interests to maintain confidentiality for our audit programs, work paper content, and all associated information. Thus, there is a dearth of empirical data about audits available for research purposes.

Also, knowledgeable people in auditing are extremely busy. We
never have all the people we need to meet the total demands upon us, and consequently, in choosing between meeting clients' needs and meeting research requests from people like yourselves, the latter come out second best as do some of our own administrative requirements. But I do not mean to appear calloused on that score, so let me explain more fully. Some auditors, especially those who have developed unique industry techniques, have a strong proprietary attitude toward them. These auditors seek to protect through the claim of confidentiality what they regard as an earned, competitive advantage. Other auditors are concerned lest access to audit information by people outside their organizations might constitute a direct or indirect disadvantage in the event of litigation.

RESEARCHERS' NEED FOR AUDIT EXPERIENCE

A far more important reason that we often cannot be of much help to you in your research is that you need some hands-on knowledge of auditing to establish any useful communication with professional auditors. Unless you have some practical understanding of audit programs, procedures, and work paper content, you may ask for information we do not have. Without a knowledge of our filing and storage practices, you may request examples that are extremely costly to collect. If you do not understand the terms we use, the sources of our information, the extent of our audit procedures, and the importance of on-the-job judgments, you may reach erroneous conclusions. Your most sober proposals may seem unreasonable to us unless you have some appreciation for our point of view.

We see this problem in a very practical sense in our need to train your students whom we hire each year. If sitting through an auditing course or reading up on the literature were sufficient, we could bring those young people in at a much higher level. We find they need significant on-the-job experience under close supervision to obtain the understanding necessary for successful professional accomplishment. You need substantial experience in auditing to be fully effective in auditing research.

There are some ways to overcome these difficulties, of course. Faculty internships, sabbatical leaves spent with a public accounting firm, long-term consulting affiliations with one or more firms, continuing service commitments on professional committees, and even mid-stream career changes are all possibilities. If you would join firms such as ours for a period of a few years, I think at the end of such a period, you could return to your educational careers, if you were to
so choose, prepared to make a much better contribution to your teaching and research.

Another possibility for effective research would be for you to work with someone who could provide the public accounting experience which you lack. My understanding is that accounting research organizations are being established at some universities. Without minimizing the difficulties involved, the possibility of audit research teams which include both academics with considerable training and experience in research methodology and practitioners with comparable levels of training and experience in practical problems has considerable appeal. We are hopeful that something of this kind can be worked out, and we would be happy to see some of our people serve in such a capacity.

**IMPORTANCE OF CAUTION IN REACHING CONCLUSIONS**

Now let me leave you with a few thoughts which I hope may influence the purpose and nature of your auditing research. First, do not rush into print. Check your conclusions with great care. Competent professionals have thought just as objectively and intelligently about auditing’s apparent weaknesses and deficiencies as you can. If they have not made the changes that you may think are necessary, there could be a very good reason why they have not. To assume on the basis of any single project that you can improve in some major way what has served with reasonable success over a long time could be a mistake.

**AVOID REVOLUTIONARY PROPOSALS**

Second, please do not seek to bring about a revolution. We need all the progress we can get. There is no question about that. Contrary to what you may believe, auditors are not loathe to make changes and many of the major firms are at the present time making a serious and extensive review of their audit procedures and practices in an attempt to make them more efficient, and to take advantage of some of the developments in quantitative methods and statistical sampling. Yes, we need progress, but disruptive change is costly, and is seldom an efficient road to progress. My experience is that meaningful or substantive progress tends to come in evolutionary steps. If you come to us with well reasoned proposals for reasonable changes, we will be happy to meet you more than half way. If you come to us with something calling for change so drastic that the costs appear to exceed the benefits, you will not receive our enthusiastic support.
COMMUNICATE IN UNDERSTANDABLE TERMS

My last and strongest suggestion is that you strive in your research to have some real impact on the profession. Help us to find those places where we need to and can make improvements. As I said, progress comes slowly. It comes hard; but it has to come. The task we practitioners face in meeting the needs of clients gives us next to no time at all for research unrelated to immediate problems. Very few of us develop the patient, contemplative approach of a scholar. Even fewer have any knowledge of the latest research techniques and methodologies. We have to look to you for help with research. But you cannot help us if your research methods and conclusions are written in what to us is a foreign language. If you find mathematical methods and symbols useful in your research, by all means use them. But, when you finally reach the point where you recommend changes in practice—and unless you do reach that point from time to time, you can have no significant impact—speak to us in terms that we can understand. If you cannot make your conclusions and reasons understandable, we cannot comprehend either their meaning or their significance; hence, they have no impact.

I recognize how important publication is to you, both to advance your academic career and as a way of life. I do a little speaking from time to time myself, occasionally get quoted in the business press, and have even been published. My goal in these communications can be expressed as “Whom have I informed and influenced?” Unless I have communicated and hopefully converted someone to my way of thinking, the speech and its publication constitute an exercise in poor investment of time and other resources.

Progress in the sense of benefit to society comes when society can understand the possibility of the benefit. Better one researcher who can communicate with practitioners for the advancement of the profession than a dozen who converse only with one another for their own satisfaction and advantage. So, do your research with whatever tools you need, but then find a way to explain its purposes and results to practical accountants like myself. Without the second step, the first one will yield very low returns.

The public accounting profession fills an important role in our market economy—and fills it reasonably well. But it could do better with the right kind of help. The considerable talents here in this research symposium could become a meaningful force for progress or represent an expenditure of time and energy of no significant influence. I hope you will choose the former.
Session

SIX
Evaluating the Competence of Internal Audit Departments

THOMAS E. GIBBS and RICHARD G. SCHROEDER

In December 1975, Statement on Auditing Standards No. 9, "The Effects of an Internal Audit Function on the Scope of the Independent Auditor's Examination" was issued. This release requires the independent auditor to evaluate the competence, objectivity and performance of internal auditors in determining the degree of reliance to be placed upon the work of the internal audit staff.

The potential use of internal auditors in attest activities has been of interest to CPAs for a number of years. For example, as far back as 1956, the Research Committee of the Chicago Chapter of the Institute of Internal Auditors conducted a survey of the cooperation between independent accountants and internal auditors which disclosed a considerable degree of cooperation. A 1962 paper (Tiedemann, 1962, p. 157) presented before the Philadelphia Chapters of the Federal Government Accountants Association and the Pennsylvania Institute of Certified Public Accountants stated that "There is probably no phase of the public accountant's work that is not affected by a good system of internal audit." He further stated, "...the public accountant must evaluate the system of internal audit in the same way he evaluates other aspects of the system of internal control. He must be satisfied that all of the requirements for effective internal auditing have been met."

Similarly, Sayad noted that
The extent to which the independent accounting will be willing to accept the work of the internal auditor . . . will depend upon the evaluation of the system of internal control, the qualifications and effectiveness of the audit staff and his judgment of the various other factors to be considered. (Sayad, 1963, p. 168).

Later, Haase noted:

The extent of the internal auditor’s participation in the year-end audit typically depends upon
1. The number and availability of internal auditors
2. The extent of their technical proficiency and training
3. Their relative independence
4. The willingness or ability of the outside auditors to delegate certain responsibilities. (Haase, 1973, p. 44).

Finally, in 1973, two years prior to the issuance of SAS No. 9, Statement on Auditing Standards No. 1, Section 320.74 stated: “Independent auditors should consider the procedures performed by internal auditors in determining the nature, timing and extent of their own tests. The work of internal auditors should be considered as a supplement to, but not as a substitute for, tests by independent auditors.” This provision of SAS No. 1 has now been superseded by the provisions of SAS No. 9.

All of these previous statements have been normative, indicating the need for cooperation between independent accountants and internal auditors. At best, previous work has been description of the types of cooperation which might take place and has provided only superficial analysis of the problem of external auditor evaluation of the internal audit function. What is lacking is a description of the important criteria used by external auditors in arriving at judgments regarding the internal audit departments and a formal process to be used in evaluating the internal audit staff.

At present, the independent accountant is faced with abiding by the provisions of SAS No. 9 which require him to evaluate competence, objectivity, and performance; however, no formal guidelines are provided by which to make this evaluation. In the 1976 Peat, Marwick, & Mitchell monograph Research Opportunities in Auditing (1976, p. 36), this concern was noted when it was stated:

In evaluating findings, an auditor is often faced with the problem of deciding the appropriate degree of reliance on others. For example, what is the appropriate degree of reliance on internal auditors? . . . Previously, the auditor lacked formal criteria for making these judgments.

The provisions of SAS No. 9 clearly point out that the internal audit staff work cannot be used as a substitute for work performed by
the independent accountant. In essence, SAS No. 9 requires the independent accountant to consider the adequacy of the internal audit staff as a part of the evaluation of internal control. A highly “reliable” internal audit staff would be one indication of adequate internal control.

SAS No. 9 requires the independent accountant to make evaluations of the competence, objectivity, and performance of the internal audit staff. This requires three separate assessments which ultimately lead to an overall opinion or judgment. Each of the concepts requiring assessment is reviewed briefly below.

COMPELENCE

The competence of an internal auditing department is a function of four interacting variables:

1. Complexity of the parent organization;
2. Capabilities of the individual auditors;
3. Internal organization of the department; and
4. Administration of departmental operations.

Competence is a relative concept, not an absolute one. An internal auditing department can be considered competent to carry out the objectives outlined above when it has the requisite skills in adequate proportions to assess, evaluate, and recommend improvements for the internal control system of the organization.

OBJECTIVITY

The term objectivity “has been used in reference to the phenomena which the accountant observes, the methods he uses, his mental attitude, and the results or information he promulgates.” (Wojdak, 1970, p. 93). The internal auditor's perception of his own ability to act independently of the individuals responsible for the functions being audited can be seen as the major factor determining the internal auditor's objectivity. The mental attitude mentioned by Wojdak, at least in this instance, seems nearly synonymous with the Mautz and Sharaf (1961) notion of practitioner-independence. These authors have outlined three dimensions of practitioner independence which apply as well to the internal audit environment as to the independent auditor:

1. Programming independence: Freedom from control or undue influence in the selection of audit techniques and procedures and in the extent of their application. This requires that the auditor have freedom to develop his own
program both as to steps to be included and the amount of work to be performed, within the overall bounds of the engagement.

2. Investigative Independence: Freedom from control or undue influence in the selection of areas, activities, personal relationships, and managerial policies to be examined. This requires that no legitimate source of information be closed to the auditor.

3. Reporting Independence: Freedom from control or undue influence in the statement of facts revealed by the examination or in the expression of recommendations or opinions as a result of the examination. The relationship of reporting to the examination has been neatly expressed in the following: "You tell us what to do and we'll tell you what we can write in our report; you tell us what you want us to say in our report and we'll tell you what we have to do." (Mautz and Sharaf, 1961, p. 206).

**PERFORMANCE**

Even a competent and objective auditor can produce poor work or work not acceptable for reliance by the independent accountant. The evaluation of performance requires both process and outcome measurements. Process measurements involve the method of selection of audit areas by the internal audit staff and include an evaluation of the adequacy of internal audit scope in the audit process. Outcome measurements involve assessments of the results of previous recommendations, the effect of internal audits on the state of the internal control system, and attitudes of component departments toward compliance with internal audit recommendations.

As noted above, the provisions of SAS No. 9 require the independent accountant to make three separate assessments. These assessments are depicted in exhibit 1.

As can be seen from exhibit 1, the evaluation of performance is only undertaken after a judgment has been made as to the relative levels of competence and objectivity in evidence in the department under study. Once performance has been evaluated, the independent accountant will be ready to make an overall evaluation of the internal audit staff, and the degree of reliance to be placed upon the staff's work.

This study addresses itself only to the determination of internal auditor competence. The justification for this limited scope is due to the time and resources which would be required for the development and testing of the three separate models for competence, objectivity and performance. Additionally, work on the area of objectivity and independence is currently being supported by the Institute of Internal Auditors, and once the measurement criteria for competence and
EXHIBIT 1. Assessment of the Reliance to be Placed upon Internal Audit Staffs

Gather evidence

Measure and evaluate competence

Measure and evaluate objectivity

Competent and objective

NO

Gather additional evidence

Measure and evaluate Performance

Adequate Performance

NO

Form overall judgment and establish reliance to be placed on internal audit work

Do not rely on internal audit Work
objectivity have been developed, it will then be possible to proceed to measures of performance evaluation.

RESEARCH OBJECTIVES
The primary research objective of this project was to develop a structured model to be used by the independent auditor in assessing the competence of an internal audit department. This objective was further subdivided into four individual goals. Accomplishment of these goals would provide the information necessary for the development of a competence model. These goals are:

Goal 1—Develop a list of criteria or components of internal auditor competence as perceived by CPAs and internal auditors.
This list of competence criteria form the possible information element which could conceivably be used by an independent auditor in forming the overall judgment of an internal audit department's competence. These criteria serve as the potential cues or stimuli in the independent CPA's judgment formation process.

Goal 2—Reduce the number of competence criteria to a manageable number which identify the major structural components of competence.
With the assistance of a group of "experts" composed of CPAs, internal auditors, and academic auditing instructors, the research combined and eliminated redundant criteria. This procedure was necessary so as to provide a manageable number of criteria for inclusion in the Goal 3 methodology.

Goal 3—Determine the importance of the criteria in the external auditor's judgment formation process.
As indicated earlier, SAS No. 9 provided relatively little guidance regarding the specifics of evaluation methods and measures to be used. This project was not done to the exclusion of relevant literature and research findings; there was simply very little information compiled on the subject.

The researchers concluded early in the project that experienced practitioners could likely provide substantial insight into the matter. The methodology utilized in this project was intended to search out and refine the current "thinking" on internal auditor competence evaluations. This approach was selected over the alternative of logically developing a normative model of competence evaluation.

Associated with each major goal of this research was one or more
tasks of the methodology. These tasks were grouped into three stages. Task 1 was termed the *Component Development Stage*. This stage was designed to expand the researcher’s knowledge on internal audit competence evaluations through open-ended surveys. Task 2 was termed the *Component Screening Stage*. This stage was used to analyze, summarize, and synthesize the identified “potential” criteria. The intended result of this stage was a relatively short list of criteria believed to be the most important factors of internal audit competence. Finally, Task 3 was termed the *Model Development Stage*. Based upon the results of the first two stages, this last stage was intended to determine the relative importance of the identified criteria to the assessment of competence. Each stage of the research project is discussed in greater detail below:

**METHODOLOGY**

**Task 1**

A survey of CPAs and internal auditors was used to construct the list of competence criteria upon which internal audit departments should be evaluated. This survey was directed to major offices of the "Big 8" public accounting firms and a sample of professional internal auditors. A contact partner was identified within each of the "Big 8" firms, and all communications were directed through this person. The contact partner was requested to identify twenty-five partners or managers with "substantial experience in audits of firms employing internal audit departments." The only requisite was that the audited firm use the internal audit function, even if the external auditors did not make use of the function in the audit process.

Additionally, a random survey of internal auditors who were members of the Institute of Internal Auditors was used to develop the list of competence criteria believed to be important to external auditor use of internal audit work.

**Task 2**

A select group of CPAs, internal auditors, and academicians was used to reduce the size of the criteria set by eliminating and/or combining the redundant criteria. A modified Delphi form of analysis was used with the expert panel. This method required the experts to identify, during several rounds of questioning, the most relevant criteria for the assessment of internal audit department competence.
Task 3

Scenarios were developed which depicted two levels (satisfactory and unsatisfactory) of the five competence criteria developed in Task 2. The importance of the various criteria were assessed using an experimental design which requires a sample of certified public accountants to state a judgment of the internal audit department's competence based upon the cues or stimuli provided in the scenarios. The data were then analyzed by using a “Lens” type model similar to that used by Slovic in his study of the judgment formation process of security analysts. This model may be illustrated as follows:

![Diagram]

Where: \( Y_s \) = The individual's judgment of the state of the variable under consideration. (The distal variable, which in this case is competence)

\( C_1, C_2, C_3, \ldots, C_n \) = Cues (Items of information which may be used to judge the current state of the distal variable)

The data generated from this model are indicative of the relative importance of the criteria to an independent auditor's judgment of internal audit department competence.

Each of the “Big Eight” accounting firms supplied a list of twenty-five managers or partners with substantial experience in auditing organizations with large audit staffs. In many of the cases, the same respondents used in Task 1 were called upon again for Task 3. These respondents were required to state a judgment of the internal audit
department competence based upon the levels of the criteria indicated in the internal audit department scenario. Scenarios were developed using two levels of each of the five criteria finally selected. This design was completely crossed resulting in thirty-two different scenarios which each respondent was asked to evaluate.

Since the primary purpose of this research was the determination of criteria weighting by an "experienced" sample of independent auditors, rather than the determination of individual auditor cue utilization, the entire sample of responses from the 146 respondents was analyzed as a group. Many Lens Model research projects have been primarily concerned with human information processing (HIP) and individual expert cue utilization. Our primary focus was not human information processing, but instead, the consensus of a group of experienced auditors on the weightings of selected "competence" criteria. Since the cues used in the Lens Model were measured at only two discrete levels, ANOVA was the appropriate analytic technique.

RESULTS

Task 1

The national offices of eight large national public accounting firms were contacted, contact partners were identified, and each of these firms agreed to participate in the Task 1 survey. Each firm selected twenty-five partners or managers with "substantial experience in audits of firms employing internal audit departments" to participate in the study. Questionnaires were sent through the contact partners to each of these 200 participants. The questionnaire requested the auditors to list the criteria currently used and potentially useful in evaluating the competence of internal audit departments. The questionnaire was intentionally open-ended since it was designed to "explode" the researcher's knowledge of internal audit evaluation, but the instructions did note that competence was but one of the attributes to be evaluated under SAS No. 9. The intent of this information was to focus the auditor's attention upon this portion of an internal audit department evaluation. Of the 200 questionnaires initially sent, responses were received from 143 participants for a response rate of 71.5 percent.

Additionally, the 1976 directory of the Institute of Internal Auditors was used to select a random sample of 500 internal auditors. The Task 2 survey was designed to assess the criteria which were perceived
by internal auditors to be useful in evaluating internal audit department competence. These criteria were used in compiling the list of components of internal auditors' competence as perceived by internal auditors. As with the CPAs, this questionnaire was also open-ended, but did instruct the auditors to consider the criteria in light of a peer review of another internal audit department. One hundred eleven usable responses were received for a response rate of 22.2 percent. A criteria was chosen for further study if it was mentioned by at least 10 percent of the total sample of CPAs and IAs.

As would be expected, there was significant duplication of criteria in the responses. The final list of fifty-four criteria was used as input to the next Task. These criteria are listed in exhibit 2. Up to this point in the research project, the researchers had attempted to expand their knowledge of the components of internal audit department competence. At this point, the Component Development Stage was concluded, and the research entered the Component Screening Stage.

From exhibit 2 it can be seen that the respondents were having difficulty separately evaluating the attributes of competence, objectivity and performance. For example, the components "the quality of internal audit work papers," and "the nature of the audit report" are probably more closely connected to performance, and independence is more closely connected with objectivity.

To remedy this apparent confusion, Task 2 specifically identified the three SAS No. 9 evaluations during each step, thereby focusing attention on the attributes of objectivity and performance, as well as competence. In using this technique, the researchers felt that respondents could more clearly understand what was being requested of them. Exhibit 3 presents a sample question from the first round of the Delphi questionnaire.

As previously noted, Task 2 was termed the Component Screening Stage. Thirty-nine individuals were selected to serve on an expert panel (known as the Delphi group). This panel consisted of thirteen partners or managers of international CPA firms with experience in auditing companies with relatively large internal audit departments, twelve internal audit managers of large corporations, and fourteen academicians with current teaching and research interest in the field of auditing.

A modified form of Delphi analysis was used in three rounds of questioning. After each round, the group was provided feedback and asked to respond to a specific charge. The group membership was not
Exhibit 2: Competence Criteria from Model Development Stage

1. An existing continuing education program
2. Competitive compensation levels
3. A high degree of communicative ability among the internal audit staff
4. The level of cost savings generated by the internal audit department
5. The educational background of the internal audit staff
6. The internal audit staff's training and experience in EDP
7. The internal audit staff's prior experience in auditing with CPA firms
8. The extent of certification of the internal audit staff (CPA, CIA, CMA)
9. The internal audit staff's role in testing new and existing computer based systems
10. The internal audit staff's previous experience within the company
11. The existence of follow-up procedures to evaluate actions on audit reports
12. The existence of hiring policies and procedures for the internal audit department
13. The independence of the internal audit department
14. The ability of the internal audit department to investigate any area of company activity
15. Existence of methods and procedures for evaluating internal control
16. The ability of the internal audit staff to maintain favorable relations with other departments
17. The internal audit staff's knowledge of the company's operations, processes, and procedures
18. The internal audit staff's knowledge of new trends and techniques in auditing
19. Acceptance of internal audit staff findings and recommendations by auditees
20. Top management's readiness to act upon the internal audit department's recommendations
21. The existence of up-to-date audit manuals
22. The existence of both financial auditing and operational auditing teams in the internal auditing department
23. Top management's support of the work of the internal auditing department
24. The relationship of and coordination between internal auditors and independent CPAs
25. The number of supervisory levels within the internal audit department
26. The existence of a responsibility statement with approved objectives for the internal audit department
27. The relative size of the internal audit department in relation to the organization as a whole
28. The organization of the internal audit department
29. The relative significance of areas and departments being audited
30. The existence and use of audit programs
31. The existence of an annual audit plan
32. The career opportunities available within the department
33. The internal audit staff's membership in professional organizations
34. The existence and use of formal performance evaluation procedures
35. The form, content and nature of internal audit department reports
36. The level at which the internal audit staff reports
37. The existence of review procedures within the internal audit department for audits and reports
38. The existence of formal reporting procedures
39. The internal audit department's staff rotation policies
40. The existence of mandatory reply procedures within the company to internal audit department reports
41. The distribution policy for internal audit staff reports
42. Quantity and quality of supervision within the internal audit department
43. Adequate scope of internal audit department audits
44. Internal audit staff's training in and use of statistical sampling methods
45. The existence of audit schedules and work assignments
46. The internal audit department's degree of compliance with professional standards
47. The existence of standards for indexing, cross referencing, and controlling workpapers
48. The existence of a formal training program for new internal audit staff personnel
49. The utilization of time budgets on audits
50. A system of on the job training within the internal audit department
51. The turnover rate within the internal audit department
52. The existence of documentation in internal audit department workpapers
53. The quality of internal audit department workpapers
54. The career opportunities available within the organization for the internal audit staff

Exhibit 3. Sample Question from Round One Delphi Questionnaire

<table>
<thead>
<tr>
<th>The internal audit staff's previous experience within the company</th>
<th>VI</th>
<th>I</th>
<th>SI</th>
<th>U</th>
<th>SU</th>
<th>U</th>
<th>VU</th>
</tr>
</thead>
<tbody>
<tr>
<td>As a measure of Competence</td>
<td>VI</td>
<td>I</td>
<td>SI</td>
<td>U</td>
<td>SU</td>
<td>U</td>
<td>VU</td>
</tr>
<tr>
<td>As a measure of Objectivity</td>
<td>VI</td>
<td>I</td>
<td>SI</td>
<td>U</td>
<td>SU</td>
<td>U</td>
<td>VU</td>
</tr>
<tr>
<td>As a measure of Performance</td>
<td>VI</td>
<td>I</td>
<td>SI</td>
<td>U</td>
<td>SU</td>
<td>U</td>
<td>VU</td>
</tr>
<tr>
<td>VI</td>
<td>Very important</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>Important</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SI</td>
<td>Slightly important</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>U</td>
<td>Uncertain</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SU</td>
<td>Slightly unimportant</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>U</td>
<td>Unimportant</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VU</td>
<td>Very unimportant</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

disclosed, although some members, notably those in like professions, were aware of others involved in the project.

The initial round of the Delphi process involved completion, by each of the members, of a questionnaire which listed the fifty-four criteria mentioned by at least 10 percent of the respondents from
Task 1. The instructions requested the respondents to indicate the extent to which they believed a criterion to be a separate indicator or measure of competence, objectivity and performance. The instructions also listed that the responses should be independent of one another, that is, that a high or low score on one criterion should have no effect on the remaining two. As would be expected, the results of this round were upwardly biased in favor of importance; however, components were selected for further screening if their average importance score exceeded the group mean score for that attribute by one standard deviation. This process resulted in 11 components for competence, 9 for objectivity and 13 for performance.

Round two of the Delphi process was intended to produce better information regarding the relative importance of the remaining criteria. The researchers believed the upward bias recognized in round one of the Component Screening Stage could be reduced by forcing a ranking of the remaining criteria. In subsequent questionnaires, respondents were told that the criteria which they were being asked to rank were those found to be most important as measured by previous scores.

Kendall’s Coefficient of Concordance ($W$) was used to determine the degree of consensus in ranks among the Delphi group members in each round. In round two, the $W$ scores were competence .288, objectivity .422, and performance .109. Even though each of these scores was significant at the .01 level of probability, it was decided that a higher level of agreement was desirable. And a third round of the Delphi process was undertaken.

In round three of the Delphi process, each respondent was provided with the mean rank for each criterion on each attribute. The respondents were asked to consider the average ranks from round two and to respond again to the same charge as in round two, that is, rank each of the criteria in order of importance to the assessment of the given attribute. Even though several criteria shifted in importance, as measured by the mean rank after round two compared with the mean rank after round three, there was substantially greater agreement in the rankings after round three. This can perhaps best be attributed to the fact that the extreme ranks became more pronounced after round three, even though there was still a rather substantial “grey area” at the mid point of the mean ranks. The attributes achieved the following $W$ scores at the conclusion of round 3: competence, .516; objectivity, .605; and performance, .235. All of these scores were again significant at the .01 level of probability.
At this point, it was determined that consensus had been achieved, and the five competence variables which had received the highest scores were selected for further study. These variables were: the existence of a continuing education program; the adequacy of the educational background of the internal audit staff; the internal audit staff's knowledge of company operations, processes and procedures; the internal audit staff's knowledge of new trends and techniques in auditing; and the quantity and quality of internal audit department supervision.

No further analysis was attempted on the components of objectivity and performance since these variables were outside the scope of the original project. The inclusion of these variables in the questionnaire was only meant to focus attention upon competence.

The final output of round 3 of the Delphi process in Task 2 was a ranked list of eleven competence criteria. To simplify the experimental design and to recognize the difficulty inherent in "processing" eleven cues in evaluating the competence of a department, the researchers decided to reduce the cues to five. Again, for simplicity, only two levels of each variable were used in developing the test scenarios. A completely crossed 2× design was used resulting in thirty-two different scenarios. An example of one scenario is presented in exhibit 4. For each of the internal audit department scenarios, the respondents were asked to score the department's competence on the four-point scale also shown in exhibit 4.

**Exhibit 4. A Representative Internal Audit Department Scenario**

The internal audit continuing education program is *satisfactory*.
The educational background of the internal audit staff is *satisfactory*.
The internal audit staff's knowledge of the company's operations processes, and procedures is *unsatisfactory*.
The internal audit staff's knowledge of new trends and techniques in auditing is *unsatisfactory*.
The quantity and quality of supervision within the internal audit department is *satisfactory*.

**Competence Scaling for Internal Audit Department Scenarios**

☑ Extremely competent
☑ Competent
☑ Marginally competent
☑ Incompetent
The contact partners in each of the eight firms were sent twenty-five copies of the Model Development Questionnaire to be distributed to the selected partners and managers. Of the 200 questionnaires, 146 usable responses were obtained for a response rate of 73 percent. ANOVA was used to analyze the main effects and the two-way and three-way interaction effects of the five criteria (or cues) on the dependent variable, competence. Input to the ANOVA model consisted of 4672 individual decision cases (146 respondents x 32 scenarios/respondent). Exhibit 5 indicates the results of this analysis.

From exhibit 5 it can be seen that each of the five criteria are significant in the evaluation of the dependent variable, competence. Also, eight of the ten two-way interaction effects would also be considered significant (the 1 x 2 and 2 x 4 effects are exceptions). While the 1 x 3 x 4 and 2 x 3 x 5 interaction effects are more significant than the other three-way effects, none would generally be considered significant.

While the significance of the F test indicates whether or not there is a systematic interrelation between the dependent variable and independent variable(s), additional measure, the $\omega^2$ statistic provided additional insight by indicating the proportion of variance in the dependent variable accounted for by the independent variable or by the interaction effect of two or more independent variables. The third column of exhibit 5 indicates the $\omega^2$ statistics for the main and interaction effects.

Not only are the main effects statistically significant, but they also account for 60.2% of the variation in competence judgments. More importantly, the internal audit staff's knowledge of the company's operations, processes, and procedures accounts for 24.4 percent of the variance while the quantity and quality of supervision accounts for 19.1 percent. Exhibit 5 also discloses that the continuing education program is apparently viewed as the least important of the criteria in determining internal audit department competence.

Of the two way interactions, only the 1 x 5, 2 x 3, and 3 x 5 interaction effects account for more than 1 percent of the variance in the competence score.

The summed omega squared ($\omega^2$) score does indicate that nearly 70 percent (68.5 percent) of the variation in the external auditors' evaluations of competence was based upon the independent variables and their interactions. The unexplained variation is based upon differences in the individual auditors' evaluation processes as well as inconsistencies within individual auditor's own judgments.
Exhibit 6 presents the comparative rankings from the Component Screening and Model Development Stages of this research.

**LIMITATIONS OF THE STUDY**

As with any research design which uses artificial situations and decisions, there are numerous limitations. Among the most obvious, and perhaps the most important are the following:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Level of Significance (F)</th>
<th>$\omega^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (Continuing education)</td>
<td>&lt; .01</td>
<td>.029</td>
</tr>
<tr>
<td>2 (Educational background)</td>
<td>&lt; .01</td>
<td>.089</td>
</tr>
<tr>
<td>3 (Company operations)</td>
<td>&lt; .01</td>
<td>.244</td>
</tr>
<tr>
<td>4 (New audit trends &amp; techniques)</td>
<td>&lt; .01</td>
<td>.049</td>
</tr>
<tr>
<td>5 (Supervision)</td>
<td>&lt; .01</td>
<td>.191</td>
</tr>
<tr>
<td>1x2</td>
<td>&lt; .05</td>
<td>.000</td>
</tr>
<tr>
<td>1x3</td>
<td>&lt; .01</td>
<td>.003</td>
</tr>
<tr>
<td>1x4</td>
<td>&lt; .01</td>
<td>.001</td>
</tr>
<tr>
<td>1x5</td>
<td>&lt; .01</td>
<td>.028</td>
</tr>
<tr>
<td>2x3</td>
<td>&lt; .01</td>
<td>.010</td>
</tr>
<tr>
<td>2x4</td>
<td>&lt; .05</td>
<td>.000</td>
</tr>
<tr>
<td>2x5</td>
<td>&lt; .01</td>
<td>.007</td>
</tr>
<tr>
<td>3x4</td>
<td>&lt; .01</td>
<td>.007</td>
</tr>
<tr>
<td>5x5</td>
<td>&lt; .01</td>
<td>.022</td>
</tr>
<tr>
<td>4x5</td>
<td>&lt; .01</td>
<td>.005</td>
</tr>
<tr>
<td>1x2x3</td>
<td>&gt; .50</td>
<td>.000</td>
</tr>
<tr>
<td>1x2x4</td>
<td>&gt; .50</td>
<td>.000</td>
</tr>
<tr>
<td>1x2x5</td>
<td>&gt; .10</td>
<td>.000</td>
</tr>
<tr>
<td>1x3x4</td>
<td>&lt; .05</td>
<td>.000</td>
</tr>
<tr>
<td>1x3x5</td>
<td>&gt; .50</td>
<td>.000</td>
</tr>
<tr>
<td>1x4x5</td>
<td>&gt; .50</td>
<td>.000</td>
</tr>
<tr>
<td>2x3x4</td>
<td>&gt; .50</td>
<td>.000</td>
</tr>
<tr>
<td>2x3x5</td>
<td>&gt; .05</td>
<td>.000</td>
</tr>
<tr>
<td>2x4x5</td>
<td>&gt; .25</td>
<td>.000</td>
</tr>
<tr>
<td>3x4x5</td>
<td>&gt; .25</td>
<td>.000</td>
</tr>
</tbody>
</table>

1. Reliance upon "experienced" auditors in the Component Development Stage, Tasks 1 and 2. Recognizing the difficulty that the respondents had with the questionnaires, it would perhaps have been better to proceed directly to an "expert panel." This would have resulted, perhaps, in less arbitrary elimination of potentially important or meaningful criteria.
2. Failure in the Component Screening Stage and the Model Development Stage to control for intra organizational and environmental conditions. Such variables as size of organization, complexity, degree of decentralization, financial condition, etc., were not controlled.

3. Failure to adequately define both the variables and the levels of the variables in the Model Development Stage. While we received no feedback indicating a lack of understanding, this loss of control could diminish the impact of our findings.

<table>
<thead>
<tr>
<th>The Internal Audit Continuing Education Program</th>
<th>Round 3 component screening stage</th>
<th>Rank based upon $\omega^2$ from model development stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Educational Background of the Internal Audit Staff</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>The Internal Audit Staff's Knowledge of the Company's Operations, Processes, and Procedures</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>The Internal Audit Staff's Knowledge of New Trends and Techniques in Auditing</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>The Quality and Quantity of Supervision within the Internal Audit Department</td>
<td>5</td>
<td>2</td>
</tr>
</tbody>
</table>

4. The number of criteria remaining at the Model Development Stage and the dichotomous levels used. These represent only a crude scenario of an internal audit department. Clearly, actual situations or more robust cases would have provided more realistic evaluation settings, but the level of difficulty and the time involved to complete a questionnaire would have increased.

**SUMMARY AND CONCLUSIONS**

Based upon this study, it can be concluded that knowledge of the company's operations, processes and procedures and the quantity
and quality of supervision are the most important factors, of those studied, affecting an external auditor's judgment of internal audit competence. Of more benefit to auditing research and to public accounting is the conclusion that a structured model can be developed which will systematize the internal audit evaluation process. This research was primarily exploratory. A large group of potential "competence" criteria was developed. An "expert" group was used to assist in summarizing and ranking the importance of these criteria. Finally, a sample of experienced auditors demonstrated that competence could be "roughly" assessed on the basis of a relatively crude set of indices.

The potential "competence" criteria should be studied further. The eleven considered most important by the Delphi group are possibly a good starting point, as are the top five used in the Model Development Stage. Scoring models for each of the criteria must be developed before the criteria scores can be combined into an overall assessment of competence. But, the simple model developed in this study points out that such models of the competence evaluation process are workable—at least in an artificial setting.

A further conclusion of this research is that the competence, objectivity and performance evaluations outlined in SAS No. 9 are not well understood by practitioners. It was clear in Task I that competence evaluation was being used as a synonym for evaluation. Once this factor was pointed out to the Delphi panel, the understanding appeared to improve.

Aside from the limitations outlined above, the three stage (Component Development, Component Screening, Model Development) procedure appeared to function well. Face to face communication during each of the three stages would have improved the communication and possibly added to the value of the results. A similar procedure could be employed to determine the major criteria in the external auditor's remaining SAS No. 9 evaluations—objectivity and performance.

REFERENCES


———, Statement on Auditing Standard No. 9: The Effect of an Internal
Audit Function on the Scope of the Independent Auditor's Examination


Discussant's Response to "Evaluating the Competence of Internal Audit Departments"

JOHN O. MILTON

In agreeing to being a discussant on Messrs. Gibbs and Schroeder's paper, I must confess that I viewed the title as being synonymous with Statement on Auditing Standards (SAS) No. 9. I was therefore surprised to find that the scope of the authors' project was limited to competence as defined in SAS No. 9, Par. 6. Objectivity, or independence, of internal auditors as well as methods for evaluating results were therefore excluded from the study. Consequently, the study was concerned only with the evaluation and ranking of such criteria as hiring, training, and supervising. After reading the paper, it was apparent that the authors' approach to evaluating the effectiveness of the internal audit function differs from mine; that is a difference which arises prior to interjection of a structured model into the evaluation process.

The authors consider the process of internal audit evaluation to be composed of three separate assessments in arriving at an overall judgment. That is, they believe competency should be considered separately from the determination of objectivity and the evaluation of performance. Further, as reflected in exhibit I, they suggest that performance should not be judged until competency and objectivity have been evaluated. However, the question exists whether SAS No. 9 requires segregation of these assessments, and whether it is feasible to do so. SAS No. 9 quotes Section 32.085 of SAS No. 1 regarding the role of client personnel with respect to accounting control: "reason-
able assurance that the objectives of accounting control are achieved depends on the competence and integrity of personnel, the independence of their assigned functions, and their understanding of the prescribed procedures.” What appears to be contemplated is a single assessment comprised of three independent variables—competence, objectivity and performance. To assist in this decision, paragraphs 6 and 7 were included in SAS No. 9 to provide clarification about the terms “competency” and “objectivity.”

From an audit efficiency standpoint, it does not make sense to separately consider the three components. For example, if the internal audit function is used solely for operational control reviews, why should the external auditor bother to determine competency and objectivity or evaluate performance? Likewise, if the external auditor finds the internal audit department’s documentation standards to be deficient, no reliance will be placed on the internal auditor’s work regardless of the competency and objectivity standards. A serious deficiency in any one of the three components will result in the practitioner placing no reliance on the internal audit function as would be the case with any internal accounting control.

Viewed in this light, it is understandable why the fifty-four criteria, produced in Task 1 with the help of external auditors, included many performance and objectivity criteria. The practitioner tends to view competence in a wider spectrum and is not concerned with categorizing the evaluation process. Consequently, the authors were seemingly not justified in concluding that “competence, objectivity and performance evaluations outlined in SAS No. 9 are not well understood by practitioners . . . many of the identified criteria were clearly unrelated to competence.” Rather, the study itself implies that it is difficult to isolate competence from objectivity and performance.

The authors inadvertently recognized the difficulty in segregating competence from objectivity when discussing the concept of competence. They state that the competence of an internal audit department is a function of four interacting variables. Two of these are the internal organization and administration of the department operations. I have difficulty in distinguishing these variables from the SAS No. 9 definition of objectivity appearing in paragraph 7 which states, “when considering the objectivity of internal auditors, the independent auditor should consider the organizational level to which internal auditors report the results of their work—and to which they report administratively.” The four interacting variables for evaluating competence appear to be arbitrary, with no foundation being provided by
the authors. The critical elements of behavior and judgment are absent.

We at Coopers & Lybrand have taken the following approach to evaluating the internal audit function—i.e., the three components are evaluated together. Specifically, the steps which are taken in order to obtain this satisfaction may be summarized as follows:

- preliminary assessment of internal audit
- liaison with internal audit
- evaluation and testing of internal audit work

The preliminary assessment phase integrates the assessments of competency and objectivity into one decision which is also influenced by many performance-related factors. Such a judgment encompasses the following considerations:

(a) The standing and responsibilities of the person or group to whom the head of internal audit reports;
(b) The responsibilities assigned through written company policy;
(c) The professional competence of the internal audit department with regard to the complexity of the accounting function. For example, where the circumstances require an evaluation of computer controls, the internal audit department should have an adequate knowledge of computer audit techniques;
(d) The extent of the control and supervision of work. For example, the work carried out by the internal audit department should be supervised by more senior personnel within the internal audit department;
(e) The extent, if any, to which access to records, documentation and personnel is restricted;
(f) The adequacy of the evidence of the work done with particular reference to quality of audit programs and working papers;
(g) The extent of the audit coverage; and
(h) The nature and frequency of, and response to, the issued reports.

While this list is not exhaustive of the criteria considered, it does illustrate the breadth of the decision being made. It should also be noted that this is a preliminary assessment; final judgment will be based on reviewing the actual work performed. As a practitioner, I have found it more efficient to make this comprehensive assessment rather than attempt to separately evaluate competency, objectivity, and performance. This is primary because this recognizes and includes the interaction between the three components.

The ultimate decision to rely upon the work of the internal auditor...
involves judgment. The reduction of such a judgmental process to a uniform mathematical model is difficult to achieve. As evidenced by the fifty-four criteria produced in Task 1, there exist a sizable number of variables which are potentially important. In practice, I consider many more criteria than the five developed by the panel of experts to evaluate competence. Thus, any decision model would probably have to consider a much larger set of factors.

Further, I am uncertain whether any weights assigned to the variables would uniformly apply to all situations. It seems possible that the relative importance of such criteria might vary depending on the complexity of the accounting system and the related internal accounting controls, the client's size and industry, the client's financial stability, etc. Therefore, is it cost effective to build and test such a uniform assessment model? It is understandable why the authors reduced the number of variables in the scoring model to five; however, the elimination of important criteria to make such a model feasible, produces a risk that some potentially important factors may be left out. As a practitioner, I need the flexibility to consider any and all variables which impinge on the decision at hand.

An additional point should be raised regarding the methodology used in the Task 3 scenarios: the description of each criteria as either satisfactory or unsatisfactory. The practicing auditor frequently encounters situations where only partial satisfaction can be achieved. Accordingly, I may be satisfied with most criteria, but the dissatisfaction with one criteria, e.g., supervision, may preclude reliance on the entire internal audit function. Therefore, I suggest that future application of this scoring model should consider expanding the degrees of satisfaction. Additionally, since satisfaction is a product of each auditor's experience and judgment, would there need to be formal guidelines for each criteria as to how much satisfaction should be assigned in a given situation?

Messrs. Schroeder and Gibbs denote the lack of a description of the important criteria used by external auditors in arriving at judgments regarding internal audit departments, and a structured model to be used in evaluating same. Thus, they appear to have intended to formulate both a descriptive model and a normative model.

However, it does not appear that the project's research design was congruent with these intentions. For instance, why were internal auditors included in Task 1 and why was the panel of "experts" composed of two-thirds nonexternal auditors if the purpose was to describe the independent auditor's judgement process? I would think
it entirely likely that the internal and external auditors might each view the ranking of the criteria differently. It appears that to have achieved this objective, the panel should have been entirely composed of a sample of external auditors. Utilization of that panel’s end product in Task 3 would then have provided some insight as to which criteria are used in the evaluation process and their relative importance.

While the research results do indicate the external auditors’ treatment of those five criteria listed, the exclusion of many variables which I deem important (including performance and objectivity criteria) indicates that the external auditor’s judgment has not been adequately described. This is not to say that the authors’ approach was inappropriate, but that it was incomplete. The background of the experts and the exclusion of relevant variables seems to have impaired any definite conclusions about current practice.

Further, the authors note that the Task 3 scoring model was simplistic. It is possible that this enabled the responding external auditors to reach a reasonable degree of consensus (68.5 percent). It is not clear, however, whether the expansion of the degrees of satisfaction and the number of criteria considered will produce a similar overall level of concordance. Thus, another obstacle could exist which would impede formalization of a decision model.

The observations also relate to the paper’s conclusion regarding normative models: “a structured model can be developed which will systematize the internal audit evaluation process.” As noted previously, it is difficult to specify how such a model would be structured, how the acceptability of the individual criteria should be described, and how many criteria should be included. To resolve such issues implies that a description of current behavior first needs to be derived. The authors seem to have implicitly accepted this as they did not attempt to develop a decision model in the paper, but settled on determining whether or not their approach could identify important criteria used by external auditors.

Practitioners always appreciate guidance in performing their work. A descriptive model might provide a framework for discussion on what criteria should be considered. While it is possible a decision model may not result from such efforts, it is possible clearer audit guidelines could evolve.

In summary, the following comments should be noted:
• separability of competence, objectivity, and performance assessments does not occur in our practice;
• there is such a sizable number of relevant criteria (inter-related among the three components) that a description of the evaluation process should first be made;
• such a description should probably be obtained solely from the group being described—i.e., practitioners; and
• satisfaction is not a clear-cut decision; accordingly, future research should consider providing standards or a framework for evaluation.
Discussant’s Response to “Evaluating the Competence of Internal Audit Departments”

EDWARD J. JOYCE

The authors have attacked a research problem of enormous dimensions. We really do not know much about the relationship (normative or descriptive) between the internal audit function and the scope of the independent audit. The authoritative pronouncements of the Auditing Standards Executive Committee have been far from helpful in this matter. Little empirical evidence exists.

Largely because of this lack of authoritative and empirical help, the authors have had to make a number of the choices necessary in any research design on a largely arbitrary basis. This was unavoidable. However, I find myself in basic disagreement with many of those choices.

The research project was conducted in three stages:

1. Component Development Stage;
2. Component Screening Stage; and
3. Model Development Stage.

In the Component Development Stage, 148 managers and partners from “Big 8” public accounting firms and 111 internal auditors provided lists of competence criteria in an open-ended questionnaire. In the Component Screening Stage, a 89-person panel (partners, managers, internal auditors, and academicians) was employed to reduce the fifty-four competence criteria provided by the respondents in the preceding stage to five criteria. Finally, in the Model Development Stage, the five criteria from the second stage were dichoto-
mously scaled and manipulated in a $2^5$ factorial design. The responses of 146 respondents (again managers and partners from "Big 8" firms) were pooled, and a composite ANOVA model was formulated from these responses. My comments on each of these stages follow immediately.

**COMPONENT DEVELOPMENT STAGE**

I do not think it was necessary to survey 143 managers and partners and 111 internal auditors to obtain the five criteria the authors ended up with in the third stage of the project. In my judgment a review of the literature and the assistance of several practitioners would have provided essentially the same results. In fact, paragraph 6 of SAS 9 alone suggests four of the five criteria the authors obtained for stage 3.

I do not think that we, as audit researchers, should view the time of people who assist us in our work as a free good. We have an ethical responsibility to use them efficiently. Frankly, I do not see anything resulting from the Component Screening Stage that warrants the use of 143 managers and partners and 111 internal auditors.

**COMPONENT SCREENING STAGE**

I have a number of reservations concerning what the authors have done to reduce the fifty-four criteria obtained in the Component Development Stage to five criteria for the model they estimate in the last stage. The authors have obviously gone to a lot of trouble to avoid making this selection on their own. I think that is laudable. Resorting to a panel of experts—as the authors have done—is a good idea.

The panel members were asked to rate each of the fifty-four criteria from the first stage on a seven-point scale according to its individual importance as a measure of competence. After several iterations of something like a Delphi technique were conducted to reduce the disagreement among the panel members' ratings, the authors took the five competence criteria with the highest individual scores as the input to the third and final stage.

I feel somewhat uneasy about the Delphi technique that was employed here, but let us assume, for the sake of argument, that it provided the authors with exactly what they wanted: the five competence criteria with the strongest univariate relationship to competence. This is fine if there are no real world correlations among the five competence criteria which are to be used as predictor variables in the third stage. It is well known in multiple regression analysis, for
example, that the linear combination of predictors that best fits the criterion may not consist of the predictors with the strongest univariate relationships with the criterion. For example, if \( x_1 \) and \( x_2 \) both have a correlation of, say, .8 with \( Y \) but \( x_1 \) and \( x_2 \) are correlated 1.0 with one another (i.e., they are completely redundant), then we gain nothing by having both \( x_1 \) and \( x_2 \) in the model. Given that we know one of these two predictors, knowledge of the other tells us nothing more about \( Y \).

If we examine the five competence criteria the authors end up with (exhibit 4), we see that the first, second, and fourth criteria are related—and in my judgment, likely to be highly intercorrelated in the real world. If we wish to build a model to be used as a formal process for evaluating competence—and we cannot include all 54 competence criteria in the model—then we should make sure that those we choose are not highly redundant.

I think the procedures the authors have employed in this stage have resulted in a model with highly redundant criteria.

**MODEL DEVELOPMENT STAGE**
A serious problem is encountered here because the first stage of the experiment has documented that fifty-four variables are relevant for the task of assessing competence and only five of them are included in the model. Now it can be argued that human beings probably cannot simultaneously process more than about five variables. I think that is probably true. Two or three variables probably account for the lion's share of the systematic variance in an individual's judgments and any other variables simply increase the random variance. That does not mean five variables are sufficient, however. The five factors that auditor A uses may be quite different from the five that auditor B uses. The Component Screening Stage demonstrated considerable disagreement among the panel members as to the important competence criteria.

It appears that the authors settled on five competence criteria because, in a completely crossed factorial design with dichotomously-scaled predictors, thirty-two judgments are required of each subject. Adding just one more dichotomously-scaled predictor increases that number to sixty-four.

A larger number of independent variables could have been accommodated in the model, however, if the authors had chosen a fractional rather than a completely-crossed factorial design. For example, it is
possible to estimate the relative weights of eight dichotomously-scaled predictors from thirty-two judgments in a 1/8 replication of a 2^8 design. The price of doing this is the need to assume that interactions are negligible—which they are, empirically. Interactions in cue utilization by individuals are sometimes statistically significant, but rarely explain much of the variance. Does going from five predictors to eight really help when fifty-four are relevant? Well a little, certainly, but probably not enough.

The last statement raises a more fundamental issue. Is an experiment appropriate here? Can experimental methods do what the authors want them to do—namely, develop a model for evaluating competence? I have some doubts. The advantages of a well-conducted experiment are precise manipulations of the experimental variables and control of extraneous variables. The cost of obtaining these advantages is nontrivial, however. In the case of an experiment one gets to manipulate a few variables at the expense of not manipulating all the others. So, unless those fifty-four or so relevant competence criteria can be reduced to perhaps nine (the maximum number that could comfortably be accommodated in a factorial design), I am skeptical of the fruitfulness of a strictly experimental approach to the resolution of this problem.

If I were attempting to study this issue, I would approach it in roughly the following fashion:
1. Attempt to assess the real world intercorrelations among the fifty-four criteria from the authors' first stage.
2. Use a factor analytic technique to reduce the dimensionality of the data set. Hopefully, nine or fewer salient dimensions would be isolated.
3. For each of these salient dimensions, choose a competence criterion which loads highly on that dimension.
4. Manipulate the levels of the competence criteria in an orthogonal and/or representative design. Capture the policies of expert auditors by observing their reactions to the manipulated stimuli.
5. Formulate a composite model of the auditors' policies and use this to provide general guidelines for evaluating competence criteria.

**CONTRIBUTION OF THE STUDY**
Although the general tone of my comments above is critical, I do think the authors' study has made a contribution. First, we now have a documented list of fifty-four competence criteria that experts con-
sider relevant for evaluating the competence of internal audit departments. The importance of this should not be overlooked. Subsequent research should benefit from this.

Second, the results of the Component Development Stage cast some doubt on the conceptual usefulness of distinguishing among the attributes of competence, objectivity, and performance as suggested by SAS 9. The authors reported that some of the fifty-four competence criteria overlapped the other dimensions. Apparently, there is some confusion among expert auditors on the meaning of the concepts, or at least how to operationalize them. The authors' work suggests further research into this issue is warranted.
To Blow the Whistle or Not: An Employee's Dilemma with Internal Control Implications

DOUGLAS A. JOHNSON

Do conventional internal control procedures provide reasonable assurance that irregularities such as embezzlement, conflicts of interest, illegal payments or frauds will be deterred or detected on a timely basis? Auditors and responsible managers have always been concerned with this question; however, it has taken on a special significance in light of the current high expectations for corporate accountability reflected in the report of the Commission on Auditor's Responsibilities (1978). The final report prepared by this independent task force includes a recommendation for management to issue a report on the financial statements and to express an opinion on the adequacy of "the company's accounting system and controls over it, including a description of the inherent limitations of control systems and a description of the company's response to material weaknesses identified by the independent auditor." (Commission on Auditor's Responsibilities, 1978) This question also warrants particular attention from managers attempting to comply with the Foreign Corrupt Practices Act of 1977 which in addition to prohibiting "corrupt" bribes to foreign officials, sets new requirements for the effectiveness of internal accounting controls. The Securities and Exchange Commission (SEC), in Accounting Release No. 242 (1978; p. 3633), advises management that because the Act became effective upon signing, it is important that issuers subject to the new requirements review their accounting procedures, systems
of internal control and business practices in order that they may take any actions necessary to comply with the requirements contained in the Act.

There is considerable uncertainty as to how management should conduct this internal accounting control review. In response to this uncertainty, the American Institute of CPAs' Special Committee on Internal Control issued a preliminary report which provides some guidance. In their report (1978, p. 9), they express the opinion that "an overall evaluation of the company's internal accounting control environment is a necessary prelude to the evaluation of a company's internal accounting control procedures and techniques." The Committee supports their concern with the control environment by arguing that "a poor control environment would make some accounting controls inoperative for all intents and purposes because, for example, individuals would hesitate to challenge a management override of a specific control procedure." They also suggest that management must demonstrate leadership and take actions which establish the proper control environment:

...the role of the board and particularly its audit committee in establishing an appropriate control environment cannot be overemphasized. ... That leadership involves creating an appropriate organizational structure, using sound management practices, establishing accountability for performance, and requiring adherence to appropriate standards for ethical behavior, including compliance with applicable laws and regulations. (AICPA, 1978, p. 9)

In the short run, management is forced to take a very subjective approach toward the evaluation of the effectiveness of internal accounting control as the theoretical models, research evidence and practical procedures that would be necessary for a more rigorous treatment do not yet exist. The objective of this paper is to begin to fill this void by: (1) focusing attention on some of the behavioral contingencies operating in the control environment which influence employees' willingness to report irregularities, (2) suggesting a theoretical framework and research methods which have potential for probing the behavioral dimensions of the control environment, (3) presenting the results of a preliminary study designed to measure the perceived seriousness of not reporting an irregularity in order to evaluate the effect of reporting consequences, (4) presenting results indicating the perceived impact of several innovative strategies for increasing the probability of reporting an irregularity, and (5) suggesting further research which can provide evidence on the reliability, validity, and generalizability of these results and contribute to the
development of more satisfactory approaches for the evaluation and the management of the control environment.

WHISTLE BLOWING AND INTERNAL CONTROL
Systems designers, auditors and managers involved with the design and review of internal control systems have relied heavily on the strategy of segregation of duties to deter or detect irregularities. It has generally been assumed that employees made functionally independent by the design of the organization structure will (always) report irregularities. These assumptions about the willingness of employees to report irregularities, however, overlook the possibility that an employee who becomes aware of an irregularity may face a dilemma due to perceived or real adverse consequences associated with whistle blowing. Thus, it is argued that we must better understand the nature and generality of the behavioral contingencies motivating whistle blowing behavior before we can adequately evaluate the internal accounting control environment.

Behavioral Assumptions of Internal Control
An organizational structure which provides segregation of functional responsibilities is a basic internal control strategy. If the responsibility for the custody of assets and the responsibility for record keeping are vested in different individuals, it is generally assumed that a fraud can only be concealed through collusion and that fear of being rejected (reported) will deter collusion among independent employees. The rationale for separation of duties to achieve functional independence among employees was explained in Montgomery's Auditing, a traditional authority on internal control as follows (Lenhart and Defliese, 1957, p. 517):

The value of "independence" on the part of persons performing many of the procedures of internal check is based upon the assumption that an independent person will report to knowledgeable authority deliberate errors, falsifications or improper use of documents, forgeries, or other irregularities coming to his attention. To be "independent" in this sense, the employee must be free to report such matters both from the standpoint of the duties assigned to him and that of his position in the line of the organization. When there is such freedom, failure to report to proper authority should occur only if the employee is incompetent, or if there is collusion between employees presumed to be independent.

A few accounting writers have clearly recognized the possibility that employees may not always be motivated toward reporting in-
regularities. For example, Mautz and Sharaf (1961, p. 145) in explaining the vulnerability of the auditor’s approach to the review of internal control caution that:

To borrow a phrase, “internal control is people.” A system of internal control is made up of people and procedures, procedures which people are expected to perform and report in normal fashion. But unknown to the reviewers, the pressures which motivate the people in the “system” may change sufficiently that they cease to act in an expected fashion, whereupon the internal control procedure loses its effectiveness.

The behavioral assumptions implicit in internal control have been best recognized by Carmichael (1970). After formally stating eight behavioral hypotheses of internal control, including the assumption that employees will report irregularities, Carmichael suggests that theory and research evidence on the operant organization casts doubt on the validity of these assumptions. He also concludes that there is a need for research on the internal control objective of safeguarding assets as most existing research is concerned with the productivity objective of internal control.

The Whistle Blower’s Dilemma
To better understand the motivation of corporate employees to report (not report) irregularities, it is necessary to consider the behavioral consequences associated with the available alternative actions. The employee who becomes aware of an irregularity may face a dilemma in that none of the alternatives open to him seem to be attractive. Generally, if the employee decides to intervene by blowing the whistle, he cannot expect much in the way of rewards or reinforcement, but must do so on the basis of moral values or a sense of duty to his company, profession or society. Depending on the mechanism available, the employee must be prepared to accept some risk to his job and career due to possible retaliation by offenders. It is also possible that management or his peers may think less of him and perhaps even consider him a trouble maker or a traitor. This would seem to be a very real risk if the offense had been committed to achieve a management goal (e.g., productivity). At the very least the whistle blower knows that getting involved may mean that his attention and energy will be diverted away from his normal responsibilities and that his performance may suffer accordingly. On the other hand, if the employee decides not to report an irregularity he may suffer from a sense of guilt in that he is indirectly responsible for subsequent occurrences of the irregularity and may be found legally guilty of “aiding
and abetting" the crime should it be discovered and all the facts become known.

If not reporting an irregularity is considered an offense, it would seem that an employee must be able to justify the irregularity as necessary or acceptable under the circumstances. Alternatively, a belief that the company would not take effective or appropriate corrective action if the irregularity was reported could also encourage the employee not to incur the costs of blowing the whistle.

There is also a third alternative action. Faced with two potentially unattractive alternatives, the employee may attempt to solve the dilemma by electing not to decide for the present time and seek more information or perhaps advice about the offense or how it could be reported safely.

If becoming aware of an irregularity puts the individual in a no win position, then there is also another solution. Employees may consciously or subconsciously attempt to avoid recognizing suspicious events or danger signals so they never have to face the dilemma.

The above analysis is, of course, highly speculative; however, it should be apparent that the contingencies associated with whistle blowing behavior may not always favor reporting the irregularity. Hence, it would seem desirable to obtain additional evidence about whistle blowing before relying on the assumption that employees will report irregularities. It would also seem desirable to obtain evidence that might help managers or auditors to form an opinion about the adverse consequences resulting from whistle blowing that are associated with company policy and procedure for receiving tips and dealing with offenders.

**THE RESEARCH APPROACH**

Do conventional internal control systems provide reasonable assurance that irregularities will either be deterred or detected? One research approach to this question would be to focus attention on the offender, i.e., the employee or manager who violates company policy or breaks the law. Traditional sociological theories of "white collar" crime (Sutherland, 1940; Cressey, 1953) take this approach and do provide some insights into the factors which may influence an employee to commit an irregularity. However, these theories are not rich in implications for control system evaluation.

The research approach adopted for this study recognizes that it is difficult to study deviant behavior directly as offenders are not likely
to cooperate with data gathering attempts. Instead, attention is focused on the willingness of corporate employees to intervene by reporting irregularities. This approach is consistent with theories of ethical risk taking developed by Rettig and his associates (1968) and economic theories of crime (Ehrlich, 1974). Here it is assumed that white collar crimes are not crimes of passion but rather rational decisions in which the offender has weighed the subjective utilities associated with the act against the subjective disutilities of being caught. The rationale then is that the attitudes of corporate employees toward reporting irregularities, as perceived by a potential offender, influence the estimated probability of being caught, and hence the decision to commit the irregularity.

Thus, it would seem fruitful to measure the moral values and attitudes toward reporting corporate irregularities held by corporate employees under various conditions and to use these measurements as an input into the evaluation of the risk of irregularities in that environment. Ultimately, procedures developed for monitoring the attitudes and values of employees through time could become part of an approach for evaluating the control environment and the effectiveness of the corporate internal control system. Comparing such measurements across organizations may also suggest the existence of corporate policies or practices which do not support or encourage employees to report irregularities, and thus signal the need for special attention in the review of the internal control system.

AN INVESTIGATION OF THE CONTROL ENVIRONMENT

In order to begin to investigate the attitudes of corporate employees toward reporting irregularities, a field study was conducted in June 1978, at a division of a large international corporation. In this exploratory study, a group of corporate accountants evaluated a set of twenty-one cases, each describing an irregularity which had been discovered by another employee. Moral judgments of the seriousness of the offense, the seriousness of not reporting the offense, and the probability of reporting under various conditions were utilized to evaluate the following research questions:

(Q1) Is the failure to report an irregularity considered to be a serious offense?

(Q2) Can alternative mechanisms for reporting irregularities, which modify the consequences of whistle blowing, increase the likelihood that an irregularity will be reported?
Research Methods

Magnitude scaling procedures are utilized in this study to measure the seriousness of irregularities (offenses) and the seriousness of not reporting an irregularity. Magnitude scaling was developed in the 1950s and 1960s by S. S. Stevens and his associates, working in psychophysics (1975), as a procedure for measuring the subjective magnitude of metric stimuli such as sound, light or force which can be measured objectively. The procedure is based on “cross modality matching” in which a judge estimates the magnitude of a stimuli using another stimuli as standard. This extensive work with metric stimuli has established that the procedure can produce high quality ratio scaled measurements.

More recently, the application of magnitude scaling techniques to nonmetric social stimuli such as attitudes and values has been heralded as a major breakthrough for the social sciences (Hamblin, 1974). Stevens (1975, p. 267), after describing an impressive list of experiments using nonmetric stimuli, concludes:

For those who must build their science on a consensus based on one or another expression of human judgment, the way stands open for effective ratio scaled quantification provided the experimental subjects are given unconstrained freedom to match numbers, loudness, length of line or some other variable directly to the items of interest.

In the past, most social scientists, including accountants, have used simple categorical scales, frequently of the Likert or semantic differential types. Now it appears possible to develop and refine attitudinal theories using more accurate measurement methods. Shinn (1974, p. 154) strongly recommends that the use of category scales be “severely curtailed.” He explains:

Instead I prefer the routine use of magnitude or ratio-estimation methods, since they provide data which easily meet the assumptions underlying many statistics and provide additional ratio level information which may be useful in some contexts. As we move from linear to nonlinear hypotheses in much social theory, this additional capability will assume greater importance.

Magnitude scaling procedures have been used to scale attitudes and a variety of other social stimuli (Stevens, 1975; Hamblin, 1974), Ekman (1962), and more importantly, Sellin and Wolfgang (1964), pioneered the use of magnitude scaling in the study of deviant behavior. The work of Sellin and Wolfgang is particularly relevant to the present study as they have established the scalability of the construct “offense seriousness.” The Sellin and Wolfgang study was a large-scale project conducted over a three-year period by a large research
team. The objective was to develop an index of crime which reflected the seriousness of crimes as perceived by the general population. In the study, a wide variety of subjects evaluated the "seriousness" or gravity of 141 different offenses based on the magnitude of the perceived harm resulting from the offenses. Their results showed a remarkable social consensus which transcended differences in age, occupation, and geographical region. Wellford and Wiatrowski (1975) provide a review of many of the replications and extensions which have been stimulated by this study.

The Cases
Each of the twenty-one stimuli developed for this study were verbal statements describing an offense committed by an employee of a large corporation, which had been discovered by another employee. Each such case was typed on a separate page. The standard for comparison was a case describing a "material" overstatement of net income (20 percent over actual) resulting from "phony" transactions recorded by the corporation's controller. Following Stevens' dicta, this standard offense was assigned a value of ten. The subject's task was to assign a number to represent how serious each offense seemed relative to the standard. The subject also estimated the seriousness of the failure to report each irregularity in the same manner and in addition, each subject estimated the number of times out of ten that they would expect a person in this situation to report the irregularity.

In the second phase of each case, the subject was again asked to make judgments of the seriousness and probability of not reporting the offense. However, this time the subject was asked to assume that the company had recently developed a new ethics code which clearly prohibited the offense and which provided one of several special reporting conditions (SRP). Five types of reporting conditions utilized were: (1) Anonymous Reporting—Employee was periodically required to sign a statement indicating that he had neither violated the code of ethics nor was he aware of any violations; (2) Signed Statements—Employee was periodically required to sign a statement indicating that he had neither violated the code of ethics nor was he aware of any violations; (3) Questionnaire Reporting—Employees receive questionnaires asking if they have noticed anything suspicious lately; (4) Special Investigation—Audit committee pledges to investigate reported irregularities thoroughly and to deal with offenders severely; and (5) Reward—Cash reward of up to $10,000 would be paid for information leading to the detection of offenders.

Four types of offenses were represented in this set of cases: embezzle-
ments, conflicts of interests, questionable payments and general offenses. Pretesting indicated that this selection of offenses provided a wide spectrum of offense seriousness. Both offenses committed for personal gain and offenses committed for the benefit of the corporation were presented in the set of cases. The primary factors differentiating the embezzlement cases and the conflict of interest situations were the behavioral contingencies of reporting. The questionable payments cases included both domestic and foreign payments made either to influence politics or to obtain business. The general offenses included tax fraud, financial statement fraud, price fixing, research falsification, and industrial pollution.

Administration
Twenty-nine professional accountants employed by a large corporation with international operations rated each of the twenty-one cases. The research instrument was administered on an individual basis in a conference room furnished by the corporation. Each subject was first personally assured by the experimenter that all responses would be entirely anonymous and that only aggregate results would be released to the employer or published. Further assurance of the anonymous nature of this experiment was provided by allowing the subject to deposit the completed instrument without any identification in a sealed box. The order of the cases in the research instrument was randomized for each respondent to avoid a systematic bias (e.g., fatigue) which might result from a fixed order of presentation. Written instructions were provided describing the purpose of the research, the magnitude scaling procedures, and the concept of offense “seriousness.” Most respondents required approximately sixty minutes to complete the task. In a debriefing interview following the completion of the task, the participants generally indicated that the cases were interesting and the task caused them to think. Most participants appeared to be interested in the task and many requested the results of the study.

The Seriousness of Not Reporting
How serious is the offense of not reporting an irregularity? Exhibit 1 presents the median estimate of the seriousness of (1) the offense (irregularity), (2) not reporting the offense, and (3) not reporting the offense assuming an ethics code and special reporting procedures (SRP) have been implemented for each of the twenty-one offenses.
### Exhibit 1. Median Seriousness Estimates

<table>
<thead>
<tr>
<th>Type of offense</th>
<th>Committed by</th>
<th>Discovered by</th>
<th>Reporting consequence of complication factor</th>
<th>Committing offense</th>
<th>Not reporting offense</th>
<th>Not reporting (with SRF)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Embezzlement:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. EHA</td>
<td>A. Controller</td>
<td>Accountant</td>
<td>Harm to family</td>
<td>20</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>2. EMB</td>
<td>A. Controller</td>
<td>Accountant</td>
<td>Loss of job</td>
<td>20</td>
<td>10</td>
<td>15</td>
</tr>
<tr>
<td>3. EVC</td>
<td>Controller</td>
<td>Accountant</td>
<td>Supervisor involved</td>
<td>15</td>
<td>10</td>
<td>15</td>
</tr>
<tr>
<td>4. EKD</td>
<td>Controller</td>
<td>Accountant</td>
<td>Join plot/records</td>
<td>20</td>
<td>15</td>
<td>20</td>
</tr>
<tr>
<td>5. EJE</td>
<td>Shipping mgr.</td>
<td>Accountant</td>
<td>None</td>
<td>15</td>
<td>10</td>
<td>15</td>
</tr>
<tr>
<td>6. EXF</td>
<td>Accountant</td>
<td>Accountant</td>
<td>Close friend</td>
<td>20</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>7. EZG</td>
<td>A. Tres.</td>
<td>Secretary</td>
<td>Supervisor involved</td>
<td>20</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>8. EAR</td>
<td>Accountant</td>
<td>Accountant</td>
<td>Join plot/records</td>
<td>20</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>Conflicts-of-interest:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. CVI</td>
<td>Purchasing mgr.</td>
<td>I. auditor</td>
<td>Harm to family</td>
<td>20</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>10. CUI</td>
<td>Trans. mgr.</td>
<td>Asst.</td>
<td>Loss of job</td>
<td>10</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>11. CVH</td>
<td>V.P.</td>
<td>Accountant</td>
<td>Superior involved</td>
<td>10</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>Payments:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. PTL</td>
<td>(D-bus)</td>
<td>Product mgr.</td>
<td>Asst.</td>
<td>Competitors</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td>14. PRN</td>
<td>(D-pol)</td>
<td>President</td>
<td>I. auditor</td>
<td>Loyalty rewarded</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>15. PSM</td>
<td>(F-bus)</td>
<td>V.P.</td>
<td>Accountant</td>
<td>Protect company</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td>16. PQQ</td>
<td>(F-pol)</td>
<td>Pres. (sub)</td>
<td>Accountant</td>
<td>Protect company</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Other Frauds:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17. TBO</td>
<td>(Tax)</td>
<td>Controller</td>
<td>Accountant</td>
<td>Follow directions</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>18. IOQ</td>
<td>(Income)</td>
<td>Accountant</td>
<td>Follow directions</td>
<td>10</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td>19. FOU</td>
<td>(Price fix)</td>
<td>V.P. (sales)</td>
<td>Sales mgr.</td>
<td>Supervisor</td>
<td>10</td>
<td>7</td>
</tr>
<tr>
<td>20. KGW</td>
<td>(Drug res)</td>
<td>R &amp; D mgr.</td>
<td>Asst.</td>
<td>Cover up</td>
<td>30</td>
<td>25</td>
</tr>
<tr>
<td>21. WIT</td>
<td>Superintendent</td>
<td>Asst.</td>
<td>Follow directions</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
</tbody>
</table>
evaluated. The median value is used to describe the responses as suggested by Stevens (1975) in his advice for exploratory studies.

The median seriousness estimate for the offense is provided as a basis for evaluating the seriousness of not reporting. The positions of the principals in the organization and the major consequences provided in the case are also reported in Exhibit 1 to allow the reader to begin to form an opinion as to their effect on the whistle blowing decision. The embezzlement and conflict-of-interest cases involve an amount which is described as large but not material in relation to the financial statements.

In comparison with the standard offense of misstatement of net income (given a value of ten), the ratings by these twenty-nine accountants indicate that not reporting an irregularity is considered to be a relatively serious offense. The medians range from a low of six for not reporting a suspected bribe to a high of twenty-five for not reporting the falsification of research data in order to gain approval to market a new drug. The seriousness of not reporting and the seriousness of the offense are highly correlated. In sixteen out of twenty-one cases, the median seriousness estimate for not reporting is slightly lower than for committing the offense. Averaging across cases the seriousness of not reporting is approximately 75 per cent as serious as committing the offenses.

The eight embezzlement cases have median seriousness ratings for not reporting of either ten or fifteen. These values are somewhat higher than those for other groups of offenses. The magnitude of these ratings for the conflict-of-interest violations are somewhat lower than those for embezzlements with the same reporting consequences. As a percentage of the seriousness of the offense, however, the conflict-of-interest ratings are higher than the embezzlement cases. Perhaps this indicates lower consequences of reporting these irregularities. Since conflict-of-interest violations are similar to embezzlements, at least to the extent that they both are committed against the employer and are for personal gain, the comparison of cases 1 thru 4 with cases 9 thru 12 provides further insight into the relationship between the consequences of reporting and the seriousness of not reporting an irregularity. For both embezzlements and the conflict violations, the threat of harm to family or an offer to collude were reported to be the most serious situations for not reporting. Apparently, actions taken against the employee are considered complications or offenses in their own right. Perhaps these actions are
seen as being so extreme that they challenge the employee to take action. Threats involving the loss of job, either overtly as in EMB and CUJ, or implicitly due to the involvement of the supervisor as in EVC and CVH were considered to be consequences of lesser magnitude. This may be an indication that accountants recognize that there is a risk of losing the job in any whistle blowing situation so that the overt threat did not add new information.

The ratio of median seriousness estimates for not reporting, to the median seriousness estimate of offense seriousness is higher for Payments and Other Offenses than for the Embezzlement and Conflict-of-Interest groups. The Payments and Other Frauds would generally benefit the company and they were apparently considered less serious than the offenses committed against the company. This higher ratio apparently indicates a higher responsibility for reporting when outsiders (society) are involved. Alternatively, this could indicate a belief that the company to some extent deserves to be “ripped off.”

Notice also, that the seriousness of not reporting when there are Special Reporting Procedures (SRP) was in all cases greater than or equal to the seriousness of not reporting without the SRP. In eleven out of twenty-one cases, the median seriousness estimate for not reporting via the SRP were as high as for committing the offense. This result suggests that there is considerable potential for enhancing the control environment through the adoption of the special reporting procedures.

**Probability of Reporting**

Can special reporting procedures which modify the consequences of reporting irregularities increase the probability that an employee who becomes aware of an irregularity will report it? The arithmetic mean of the subjective probabilities for reporting in the 21 cases with the SRP and without the SRP are presented in Exhibit 2. These subjective estimates were in response to the question, “How many times out of ten do you feel a person in this situation would report the offense?” Since all irregularities were not tested with each of the six SRP, and since the assignment procedure was arbitrary, caution should be exercised in conclusions about the relative merit of these possible modifications to the control environment. Several cases were tested using two SRP and they are in this respect more interesting as they provide a direct comparison under comparable conditions. The alternative SRP and the median score for these dual cases is indicated in parentheses for each comparison.
Exhibit 2. Perceived Probability of Reporting Irregularities* under Six Special Reporting Procedures (SRP)

<table>
<thead>
<tr>
<th>Anonymous reporting: (A)</th>
<th>Without SRP</th>
<th>With SRP</th>
</tr>
</thead>
<tbody>
<tr>
<td>EMB Embezzlement—job threat</td>
<td>5.41</td>
<td>7.88</td>
</tr>
<tr>
<td>EUC Embezzlement—supervisor involved</td>
<td>3.93</td>
<td>7.66 (Q = 7.10)</td>
</tr>
<tr>
<td>EJE Embezzlement—shipping dept. head</td>
<td>6.83</td>
<td>8.69 (Q = 8.58)</td>
</tr>
<tr>
<td>EZG Embezzlement—clerk suspects boss</td>
<td>4.21</td>
<td>7.62</td>
</tr>
<tr>
<td>CUJ Conflict-of-interest—job threat</td>
<td>4.01</td>
<td>7.52 (Q = 7.51)</td>
</tr>
<tr>
<td>PRN Domestic political contribution</td>
<td>4.66</td>
<td>8.00 (C = 6.59)</td>
</tr>
<tr>
<td>PSM Bribe to obtain foreign business</td>
<td>4.66</td>
<td>8.45 (C = 6.55)</td>
</tr>
<tr>
<td>TBO Tax fraud</td>
<td>4.03</td>
<td>7.32 (S = 6.62)</td>
</tr>
<tr>
<td>RGV Drug research fraud</td>
<td>6.45</td>
<td>8.28</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Questionnaire reporting: (Q)</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>EUC Embezzlement—supervisor involved</td>
<td>3.93</td>
<td>7.10 (A = 7.66)</td>
</tr>
<tr>
<td>EJE Embezzlement—shipping dept. head</td>
<td>6.83</td>
<td>8.88 (A = 8.69)</td>
</tr>
<tr>
<td>CUJ Conflict-of-interest—job threat</td>
<td>4.01</td>
<td>7.31 (A = 7.32)</td>
</tr>
<tr>
<td>PQO Foreign political contribution</td>
<td>3.38</td>
<td>6.21 (C = 6.23)</td>
</tr>
<tr>
<td>WHT Flagrant pollution violation</td>
<td>6.00</td>
<td>7.28</td>
</tr>
</tbody>
</table>

Signed statement: (S)

| EHA Embezzlement—threat to harm family                       | 4.89        | 6.38     |
| EKD Join and help cover up                                   | 5.34        | 7.00     |
| EXY Join since you didn’t get raise-friend                   | 3.72        | 7.10 (R = 7.17) |
| EAR Join and help conceal                                    | 6.24        | 7.55 (I = 7.86) |
| CUI Conflict with threat of harm to family                   | 6.18        | 6.48     |
| PTL Domestic bribes to obtain business                       | 4.72        | 7.90 (C = 7.90) |
| TBO Tax fraud                                                 | 4.08        | 6.62 (A = 7.52) |
| FKU Price fixing                                              | 4.86        | 7.10     |

Investigate (Get tough): (I)

| EAR Join and help cover up                                   | 6.24        | 7.86 (S = 7.55) |
| CWK Join and help cover up                                   | 4.82        | 6.79     |
| IOQ Help co. make forecast & merger—financial reporting fraud| 5.90        | 7.45 (R = 7.45) |

Reward: (R)

| EXF Join friend because didn’t get raise                     | 3.72        | 7.17 (S = 7.10) |
| EZG Clerk suspects boss                                      | 4.21        | 7.60 (A = 7.62) |
| IOQ Financial reporting fraud                                | 5.90        | 7.45 (I = 7.45) |

Code: (C)

| CYH Supervisor involved                                      | 5.00        | 6.34     |
| PTL Bribe to obtain domestic business                        | 4.72        | 7.00 (S = 7.90) |
| PRN Illegal domestic political contrib.                      | 4.66        | 6.99 (A = 8.00) |
| PQO Foreign political contribution                           | 3.38        | 6.28 (Q = 6.21) |
| PSM Bribe to obtain foreign business                         | 4.66        | 6.55 (A = 8.45) |

N = 29

*Mean of responses to question—"How many times out of ten do you feel a person in this situation would report the offense?"
The highest estimates for the probability of reporting occurred under the anonymous reporting procedure, SRP-A. Under SRP-A, employees were encouraged to contact the audit committee and to report suspicious events or problem areas that should receive the special attention of auditors. A special “hot line” telephone number was provided so that the whistle blower did not have to reveal his identity, thus reducing the risk of retaliation from the offenders. The average probability of reporting an irregularity under SRP-A was 7.95 which is substantially lower than “always” but still 58 percent higher than the probabilities for reporting the same offenses without this enhancement. In the limited comparison with alternative SRPs the anonymous procedure always produced the highest ratings.

The average rating for SRP-Q was 7.25 which was also above average and indicated a potential for improved reporting over current procedures as this was 50 percent higher than without this SRP. Under this procedure the respondent received a questionnaire from the audit committee inquiring if s/he had noticed anything suspicious or unusual lately. Although the probabilities under SRP-Q were relatively high, they were not as high as for the same cases under alternative SRPs.

The probabilities of reporting under SRP-S were 44 percent higher than without this innovation and averaged a moderately high 7.0. Under the signed statement procedure, key employees were required to periodically sign a statement indicating that they were not aware of any irregularities. With the exception of case PTL which involved paying bribes to obtain domestic business, the alternative SRPs were perceived as superior to SRP-S in encouraging reporting.

Only three cases were tested under SRP-I which appealed for employee cooperation in a program to avoid future irregularities. In SRP-I, the audit committee promised that they would vigorously investigate reports of irregularities and to deal with offenders effectively (severely). The objective of this approach is to assure employees that the company would not just go through the motions of investigating an alleged irregularity while expecting the whistle blower to continue to work with the offender so as to be particularly subject to retaliation. Hopefully, the belief that the company is interested and committed to doing something about violators would encourage employees that it is worth the trouble to report irregularities. The average probability of reporting under SRP-I was 7.36 which was approximately a 32 percent improvement over the ratings for these cases without this program.
Three other cases were utilized in testing SRP-R which offered a reward of up to $10,000 for information leading to the discovery of a major irregularity. The high 7.40 frequency of reporting under this plan was almost double the rate of reporting without the procedure. This would seem to indicate that these corporate accountants strongly believe that their peers could be motivated to report irregularities if the company made it "worth their trouble." Perhaps this conclusion should be tempered, however, as the alternative SRPs also produced similarly high estimates of the frequency of reporting in these cases.

SRP-C announced that the company had recently developed a formal "ethics code" which clearly indicated that the irregularity was contrary to company policy and reminded employees that it was their responsibility to report irregularities. The mean probability estimate for cases tested under this procedure was 6.55 which was 58 percent higher than the estimates without SRP-C. The cases utilized with this procedure primarily involved questionable payments, an area in which the clarification of corporate policy could be expected to be effective. Although SRP-C appears to have been considered a substantial improvement, SRP-A was still able to provide a sizable increment over the benefits from clarification alone.

Overall, the ratings by these twenty-nine corporate accountants are consistent with the view that almost any expression of interest or program to enhance the control environment would be expected to increase the chance that an employee who becomes aware of an irregularity will decide to report it to the appropriate authorities. These results are also consistent with the argument that modification of the consequences of reporting an irregularity should enhance the control environment by motivating employees to intervene when they become aware of suspicious activities. These results suggest that particular attention should be given to the feasibility of establishing procedures for anonymously reporting irregularities. It should, however, be pointed out that the evidence which has been presented in support of control environment enhancement through whistle blowing consequence management (WBCM) and for specific SRPs is far from conclusive.

In almost all research, limitation of time and resources lead to design compromises. These trade-offs make the study feasible, but are often translated into threats to internal and external validity. This study which probes a new area of research leaves a number of factors uncontrolled and these factors could have possibly influenced or biased the results. Since the results of this study are generally encour-
aging, it is suggested that these limitations should be interpreted optimistically, as opportunities for future research.

**Future Research**

Do conventional internal control procedures provide reasonable assurance that irregularities will be either deterred or detected on a timely basis? The results of this study seem to indicate that a relatively small group of corporate accountants employed by one corporation believe that not blowing the whistle on an irregularity is nearly as serious as committing the offense, but that many offenses are nevertheless not reported when there may be adverse consequences associated with doing so. Although these results are encouraging, the evidence is not sufficient to warrant a call for major changes in corporate policy. A major program of research which could provide more conclusive evidence as to the need for modifying existing control environments and evaluate the potential of whistle blowing consequence management procedures might address the following research questions:

1. What other factors affect the perceived seriousness of an irregularity and the willingness of employees to report violations? For example, do organization and industry characteristics affect the willingness of employees to report irregularities? Multivariate designs and explicit consideration of interactions would seem appropriate for extending this research.

2. How do the attitudes of other employee groups compare with the attitudes toward reporting irregularities held by accountants?

3. How realistic or accurate are employee estimates of the frequency of reporting irregularities? Or, to what extent can employee attitudes toward reporting irregularities be relied on to predict actual employee behavior?

4. Do employees who have committed irregularities indicate that other employees were aware of their offense but did not report them (why not)? Do known violators admit having considered the chance that other employees would not report them as a factor in their decision to commit their offense?

5. How do corporate policies for treatment of known offenders (punishment) affect the willingness of employees to report irregularities?

6. Does the rotation of audit personnel adversely affect the willingness of client personnel to report irregularities?

7. Do outside auditors manage auditor-auditee relationships in a manner that encourages client personnel to report irregularities?

8. How do employees' attitudes toward reporting irregularities relate to other attitudinal/motivational measures such as job satisfaction? Are there less obtrusive indicators?

9. Do employee exit interviews reveal a frequency of unreported irregularities consistent with the estimates of current employees? What can be done to encourage whistle blowing by terminated employees?
10. Are particular groups or types of employees especially sensitive to the existence of irregularities? How stable are irregularity reporting estimates over time? What type of a sampling plan might be effective for monitoring these attitudes and beliefs on a practical basis?

REFERENCES


Discussant’s Response to “To Blow the Whistle or Not: An Employee’s Dilemma With Internal Control Implications”

ROBERT S. ROUSSEY

Internal accounting control is receiving increased attention; initially, from the endorsement of the Cohen Commission Report by the Financial Executives Institute; secondly from the passage of the Foreign Corrupt Practices Act which resulted from the disclosure of the so-called “sensitive payments” made by major U.S. corporate enterprises; thirdly, from the notice by the SEC that it is preparing rules to require management reporting on internal controls. Thus, any research directed to the betterment of the internal control environment should be well received by corporate management and external auditors.

The research recently performed and being proposed by Douglas Johnson is intended to improve this environment. It relates to the problems associated with reporting of irregularities by employees which is referred to by him as “whistle blowing.” The first phase centered around the seriousness and probability of not reporting fraudulent offenses by various employees. The second phase focused on these same aspects but where extended and required conditions were established, referred to as “Special Reporting Conditions,” for reporting of these offenses. To me, the thrust of the research asks the following question. Is there a need to modify the existing internal control environment with respect to reporting of irregularities? In effect, are the attitudes and environmental situations such that new codes of conduct and new reporting conditions need to be research and implemented?
In my discussion, I will not take issue with any of the research methodology or findings, per se. While these are fertile grounds for discussion, I would like to focus my comments on whether there is a need for expanded methods to promote whistle blowing.

THE WHISTLE BLOWER'S DILEMMA
The research paper correctly and forcefully brings out the problems an employee faces in reporting irregularities. In today's society, rewards cannot be expected while risks to job, career, life and limb can be. This may not be a pleasant situation for employees faced with this dilemma, but it is one we can all understand.

EFFECTIVENESS OF WHISTLE BLOWING IN AN INTERNAL CONTROL ENVIRONMENT
As I have not seen, nor do I believe has anyone else, any statistical evidence as to the role that whistle blowing has played in uncovering fraud situations, it is difficult to determine how important it is or can be to prevent or detect fraudulent activities. However, it is my personal observation that it is not, nor will it be, a significant part of the internal accounting control environment. I reach this conclusion because of a number of factors. These are:

Change in Business Organizations
In the last decade or two, business entities have grown to enormous size and complexity. Entities having employees of 10,000 to 100,000 are commonplace throughout the world. This change in structure has created opportunities for improved internal control conditions while at the same time opened new avenues for fraudulent activities. I find it difficult to imagine, however, extensive fraud reporting requirements for these organizations even if confined to salaried employees only.

Change in Data Processing
The use of computers to process data of these new multinational, multibusiness companies has helped in making these corporate empires a reality. As such, the system of internal controls has had to change to keep pace with the changes in the processing of data. Centralization of data, on-line, real-time access to data, creation of data bases, telecommunication of data and data access as well as distributed processing of data has stretched the imagination of all people involved with internal accounting controls.
Change in Control Environment
The passage of the Foreign Corrupt Practices Act will probably have the single greatest impact on internal accounting controls. At the present time, many corporate organizations are having significant and detailed reviews made of their internal control systems. Corporate management, boards of directors, and audit committees have turned their attention to the requirements of the Foreign Corrupt Practices Act and are placing a new emphasis on internal controls. As part of the potential for management reporting on controls under this Act, corporate management is asking division and subsidiary management to report on controls at their locations. While management has always been involved and has been responsive to the need to change control systems as new developments have occurred, there is a new, more direct involvement by management with internal controls. This, I feel, will change the overall control environment even further and help to make it even more responsive to today's environment.

Difficulty in Recognizing an Irregularity
I think one of the most significant reasons for not putting whistle blowing high on the internal control technique list is due to the difficulty in recognizing a fraud situation. What may appear to be an unusual activity or transactions may not be and vice versa. A well-concealed fraud is very difficult to recognize or detect. In a recent newspaper article, a trained fraud investigator discovered a suspected fraud only after finding a suspicious vendor invoice after reviewing 25,000 invoices.

ROLE OF WHISTLE BLOWING IN AN INTERNAL CONTROL ENVIRONMENT
If reporting of irregularities is not a significant part of the overall control environment, does this mean it is not important, should not be emphasized or should not be considered for additional research?

To this question, I believe the answer is that it is important, should be emphasized and should be researched further. The real question is, however, to what extent and how? To put it into auditor's language, what is the extent, scope and timing of whistle blowing in the internal control environment?

I believe that reporting of unusual or potentially fraudulent activity or transactions is necessary in a properly functioning internal control environment. If unusual transactions or potential fraudulent activity are not reported and investigated, it would seem that there is a potential for weakness in the control system.
What do I mean, however, that reporting of these activities or transactions is necessary in a properly functioning internal control environment? Perhaps to do this I need to explain what I mean by internal control environment. In most companies, internal accounting control can be viewed from two aspects: first, from the overall or supervisory type of control exercised by management; and second, from the more specific level of control involving specific control techniques and the people who implement and monitor these techniques as well as the people who investigate and conclude on questionable matters brought out through the control system. This group of people encompasses management, the controllers' organization, the data processing department, internal auditors and others in the organization having responsibility for internal control functions.

In most organizations, this group of people is a relatively small percentage of the total number of employees. As such, in considering the need to have employees report irregularities, I believe the scope of such reporting should be directed more toward the employees involved with internal control activities: the employees performing some type of control function in a company. These are the people trained, in effect, to prevent and detect, through the application of various manual and computerized control techniques, the unusual as well as the possible fraudulent activities. To this group I ascribe the importance of bringing unusual items, including suspected irregularities, to the attention of the proper people in their organization. This is part of their job and part of what is expected of them. I do not believe the same degree of responsibility should be directed to all of the other members in an organization.

THE FRAUD ENVIRONMENT

Let us turn for a moment to the other environment being discussed: the fraud environment. We have all read accounts of the huge potential amounts of fraud that might exist in the business world. Occasionally, we read about specific large fraud situations. Most of these massive frauds involve some type of collusion with the larger frauds involving the highest echelons of management.

However, if we take all the U.S. business frauds reported within the last decade and match them against the gross product of U.S. businesses, what would be the percentage amount of these frauds? While I have not seen any specific statistics, I would expect the amount to be insignificant. What about the unreported business fraud? While these estimates are large, and suspect or at least questionable in my view,
they still would be only an insignificant part of total business turnover.

The point is this: is the actual or potential fraud that exists in the free enterprise system of such magnitude that we need to consider some of the actions mentioned in Doug Johnson's paper? Are we operating in an environment permeated by dishonest and corrupt people that would require such drastic action so as to require codes of ethics utilizing such things as anonymous reporting, signed statements by employees that they are not aware of fraud situations or rewards for reporting offenders?

As an auditor, I do not think we are operating in a fraud environment. On the contrary, I believe we are working in an environment where the greatest percentage of management and employees are honest and hardworking and that drastic fraud control action extending to all employees is not appropriate. Fraud exists and needs to be prevented and detected. The best way to do this, however, may not be through extensive fraud prevention methods of the type described in Johnson's paper.

It is interesting to contrast two situations: the so-called “sensitive payments” disclosure and the currently disclosed massive fraud situation at the General Services Administration (GSA).

The Foreign Corrupt Practices Act came into being primarily because of the sensitive payments disclosures. Do corrupt practices relate to fraud? Are these sensitive payments equivalent to fraudulent transactions? Whatever the answer, it was interesting to read the letter of Alexander Hehmeyer, an attorney in Chicago, written to the chief of the criminal division of the Department of Justice, which appeared in the October 17, 1978, edition of The Wall Street Journal. Mr. Hehmeyer questions how the Justice Department could “ever consider criminally prosecuting American business executives for doing what all sectors of our government told them was a customary and expected practice.” In the letter, he names officials of various government agencies, including Treasury and State, and describes how the businessmen on one trade mission were told of the under-the-table culture, of the amounts of payments expected in return for contracts and of how the payments were normally handled through commission agents. If U.S. government officials aided and abetted the sensitive payment issue, should all the blame and focus on internal control deficiencies be placed only on American business enterprise? Or, should some of the blame and focus be directed toward the U.S. Government?
The frauds being disclosed at the GSA, on the other hand, seem to imply a more rampant, widespread situation, as related to the number of GSA employers involved, than any I have read about in any business enterprise with the possible exception of Equity Funding. Is this the tip of the iceberg with respect to fraud in government operations? Do we need a U.S. government Foreign Corrupt Practices Act? Do we need codes of ethics, reporting requirements and a reward system for all U.S. government employees? Is the same necessary for state government employees? Is the same necessary for all individuals as related to knowledge of fraud in all business and government organizations?

Obviously, this would be going to great extremes and is not the answer. The point I want to make is that I do not believe fraud is so rampant in the business sector to warrant the establishment of extensive whistle-blowing techniques. On the other hand, it appears that additional emphasis has to be placed on internal control in the government sector.

LIMITING WHISTLE BLOWING RESEARCH TO INTERNAL CONTROL PERSONNEL

As I mentioned before, I believe reporting of unusual and potentially fraudulent activity is an important and necessary part of the internal control environment. At this time, I believe any emphasis on research should be directed toward reporting of this activity by personnel involved in the internal control system.

Why should this reporting activity not be extended to all personnel in an organization as demonstrated by some of the research work contained in Johnson’s paper? Some of the possible factors and questions for limiting this activity to personnel involved with internal control are discussed below. Perhaps these will stimulate further discussion.

Spotting the Fraud

Unusual or fraudulent transactions leave pointers or flags which indicate something out of the ordinary. I have been told by someone experienced in fraud detection that these pointers or flags are extremely difficult to spot or recognize. If internal auditors, external auditors and control personnel do not have the complete training necessary to recognize fraudulent transactions, should the average employee be expected to have the ability to recognize them?
Erroneous Reporting
If all employees are “required” to report unusual or fraudulent activity, will the requirement result in erroneous reporting? Will employees, particularly if anonymous reporting is allowed, report anything that appears suspicious? Will this cause suspicion to be placed on people not involved in fraudulent activity? In effect, could it cause many “wild goose chases”?

Fraudulent Reporting
Could extensive emphasis on fraud reporting encourage “grudge” reporting? Could any system that promises extensive investigation and follow-up encourage a flood of reporting by employees that are based on grudges or grievances against supervisors or other employees?

Cost/Benefit
Another important factor to consider is would the costs, both tangible and intangible, associated with any extension of a whistle blowing campaign to an entire organization be beneficial to an organization? Could a cost/benefit relationship be established?

WHAT TO DO
How does whistle blowing fit into the control environment and is further research necessary? My thoughts on this are briefly summarized below.

Control Environment
Companies need to continue to create an environment of honesty and control. Management example and direction are imperative to create this type of environment.

Control System
The internal control systems in today’s complex business environment have to be developed to meet the objectives of control as established by management. The use of control techniques to meet these objectives will require extensive consideration and use of a myriad of techniques.
Whistle Blowing
Reporting of unusual or potentially fraudulent transactions is only one part of a control system and should not be blown out of perspective. It should be emphasized within a reasonable perspective and apply to the personnel involved with internal controls.

Whistle-Blowing Research
Research can be conducted to help determine how employees in control positions can better deal with the problems associated with recognizing and reporting unusual and potentially fraudulent activity.

Government
Importance of internal accounting control should be emphasized for all government operations and Congress should consider passage of a Government Corrupt Practices Act.
Magnitude scaling procedures were developed originally for the purpose of determining the relationship between the magnitudes of physical stimuli, which exist in the environment and which can be measured independently, and the individual's subjective perception of those magnitudes, which exist solely within the individual. Examples of the topics originally investigated via such procedures are loudness, brightness and heaviness. This work was later extended to include investigation of the individual's subjective response to various types of stimuli for which independent measurement is impossible. Examples include studies of occupational preferences among United States university students, preferences for wrist watches among Japanese university students, the esthetic value of music, the pleasantness of odors, the importance of Swedish monarchs, and the seriousness of some 150 offenses ranging from murder to stopping in a no-parking zone to mail a letter (Stevens, 1966).

Following the lead of S. S. Stevens, who is responsible for much of the original development and application of magnitude scaling, and the lead of several investigators who have tried to scale the notion of offense seriousness, Doug Johnson is engaged in a line of research which focuses on the types of offenses or irregularities of interest to auditors, and he has provided us with an initial empirical study in this area. The potential for this kind of research to provide answers to questions in which auditors are interested must be evaluated jointly
by practicing auditors and by those who are cast in the role of research critic. If I can legitimately claim expertise in either area, it must be the latter. Therefore, the law of comparative advantage, plus my conviction that the internal validity of research must be firmly established before it makes sense to take any action based on that research, lead me to concentrate entirely on the manner in which Johnson’s research was conducted. My remarks will be critical in an effort to help Johnson and other researchers who undertake similar projects in the future.

I think Johnson’s research has two pervasive problems which are sufficiently serious so that a confident interpretation of the results is precluded. I will address each of these problems at some length. In addition, as with most research projects, there are several narrower issues about which I have some question, and I will enumerate some of these narrower issues. Finally, I will conclude my remarks with some very general comments about the reliability and validity of this type of research. For the most part, these latter comments will apply to all types of behavioral research, but I will couch them in terms of research on perceptual psychophysics.

A great deal of literature is available which suggests that individuals approach vague, complex or novel tasks, including experimental tasks, with a set of simplifying heuristics which often lead to systematic response biases. One heuristic which has been identified is that of anchoring and adjustment; that is, people have been found to make estimates by starting with some initial value that is adjusted to form the final estimate (Tversky and Kahneman, 1974). The initial value (the anchor) may be suggested by the context of the problem or it may result from a partial computation or some other event that occurs during the solution of the problem. Different anchors yield different estimates, and the extent of adjustment from an anchor typically is insufficient.

Two examples (from Tversky and Kahneman, 1974) should help to illustrate the anchoring phenomenon. In one study a wheel of fortune was spun in the subjects’ presence, generating a number between 0 and 100. The subjects were required to estimate the percentage of African countries in the United Nations by moving upward or downward from this arbitrary number to make their estimates. The median estimate for subjects who started with the number 10 was 25 percent, while the median estimate for those who started with the number 65 was 45 percent. In another study one group of subjects was required to estimate within 5 seconds, the product 8·7·6·5·4·3·2·1, while a second group was required to estimate the product 1·2·3·4·5·6·7·8. If subjects
anchor on an initial partial computation (i.e., multiplication from left to right), one would expect those subjects receiving the descending sequence to produce estimates that are higher than those of the subjects receiving the ascending sequence. The median estimate for the descending sequence was 2250, while the median estimate for the ascending sequence was 512. (The correct answer is 40,320.)

With this in mind, let us return to Johnson’s study. It certainly seems reasonable to suggest that his subjects would have viewed their task as somewhat vague and novel. It is very unlikely that they had ever thought seriously about quantifying their own subjective notions of relative seriousness if, in fact, they had even entertained the notion of seriousness prior to the experiment. Thus, Johnson’s experiment would seem to provide a prime opportunity for the use of simplifying heuristics. Indeed, much of the data he reports are consistent with this interpretation.

Note that Johnson’s subjects had the opportunity to use both types of anchors referred to previously—those provided by the context of the problem and those resulting from an event occurring during the solution of the problem. Johnson supplied his subjects with the first type of anchor when he required them to estimate the seriousness of various offenses relative to a standard offense with a value of 10. In addition to this anchor supplied by Johnson, the subjects may have supplied themselves with additional anchors since they were required to estimate the seriousness of not reporting the offense during the same experimental session that they estimated the seriousness of the offense itself.

If the subjects anchored on the number 10 and made relatively small adjustments from that anchor, we would expect their responses to cluster around 10. If, in addition, the subjects anchored on their own seriousness-of-offense estimates when estimating the seriousness of not reporting the offense, we would expect their seriousness-of-offense estimates and their seriousness-of-not-reporting estimates to agree very closely with each other. Finally, if we combine the tendency toward anchoring and insufficient adjustment with the tendency to simplify the task by responding in “round” numbers, we would expect to find among the subjects’ responses a large number of 10s, 15s, 20s, and so on.

Examination of Johnson’s exhibit 1 reveals that of the 63 median seriousness estimates reported there, 52 of them end in 0 or 5. Of these 52, there are nine 20s, twelve 15s, and twenty-eight 10s. Finally, the remaining eleven numbers are between 6 and 9, inclusive.

I think there is a distinct possibility that all the results reported in
exhibit 1 hinge on Johnson's decision to provide the subjects with a standard offense which was scaled as 10. Further, the rationale for this decision by the researcher is unclear. Johnson states that he was following a dictum of S. S. Stevens, the latter-day father of perceptual psychophysics. Stevens (1956) presented a list of nine dicta in his first published experiment based on magnitude scaling procedures—an experiment concerned with measuring subjective loudness. Stevens' dicta are simply suggestions concerning how to conduct such an experiment. Hamblin (1974) observes that Stevens' dicta never have been subjected to systematic investigation but that most are still followed in psychophysical experimentation. Stevens himself seemed to be uncomfortable with his own suggestions. As a matter of fact, Stevens did not require that the number 10 be used; instead, he suggested the use of a number that could be easily multiplied and divided—such as 10. He also suggested that both the standard stimulus and the number assigned to it be varied in future replications of the same experiment.

Toward the end of his original article, Stevens engaged in a lengthy discussion of some of his own work in which no standard stimulus was used. He concluded that the results in the no-standard conditions were at least as good as those in the conditions in which a standard was used. Further, he reported that some subjects had a strong preference for the procedure in which no standard was used. One of his subjects said the following:

I felt freer to use numbers over a wide range. I liked the idea that I could just relax and contemplate the tones. When there was a fixed standard I felt more constrained to try to multiply and divide loudnesses, which is hard to do; but with no standard I could just place the tone where it seemed to belong. (Stevens, 1956, p. 21)

Hamblin (1974) observes that the use of the no-standard procedure has become the preferred way of conducting such experiments. Hamblin also suggests that zero be defined for the subjects, although he gives no specific rationale for this suggestion. This suggestion does, however, have its counterpart in recent work on end-of-scale bias which shows that subjects are reluctant to "use up" the end of a scale. In the present context, end-of-scale bias could be an additional explanation for the clustering of median estimates around 10. I think the implications of the anchoring-and-adjustment phenomenon for the interpretation of Johnson's results are evident when one recognizes the importance of response biases in studies such as this one.

Turning to the second pervasive problem of Johnson's paper, the
principal advantage of the experimental method is that it allows the researcher to control certain variables and to manipulate other variables systematically. The purpose of such control and systematic manipulation is that it enables the researcher to establish associations between independent and dependent variables and that it may provide a basis for inferring causation. To say the least, the design of Johnson’s study does not enable him to enjoy the advantages of the experimental method. The variables included in his study are not manipulated in a systematic fashion, and this results in a great deal of confounding of effects.

One of Johnson’s primary concerns seems to be the type of offense which has been committed. These types include embezzlements, conflicts of interests, questionable payments and a miscellaneous category of five assorted offenses. Unfortunately, Johnson’s did not hold constant the sets of circumstances surrounding the occurrence of these various offenses. Exhibit 1 shows that as we move from one type of offense to another, or even to other offenses of the same type, several factors changed other than the offense itself. Intuitively, the most important of these other factors seems to be the potential consequences of reporting the offense (e.g., loss of job or harm to family). For the most part, however, the potential consequences of reporting do not vary systematically with the type of offense. Further, what Johnson calls a “complication factor” also varies unsystematically with the type of offense. Complication factors include such things as an offer of collusion and the involvement of one’s close friend, one’s supervisor, and one’s competitors. In addition to unsystematic variation of the types of offenses, the potential consequences of reporting, and complication factors, similar unsystematic variation relates to the person who committed the offense (e.g., the controller, the assistant treasurer, the shipping manager, the purchasing manager, the president, and the superintendent) and the person who discovered the offense (e.g., the internal auditor, the secretary, and the engineer). Finally, the primary beneficiary of the offense (e.g., the perpetrator himself or the company) varies as well.

Similar confounding occurs with respect to Johnson’s attempt to determine the impact of several special reporting procedures on the perceived seriousness of not reporting the offense and the perceived probability of reporting. The special reporting procedures were matched arbitrarily with the types of offenses which, for whatever reason, were rated as more or less serious. Thus, the effects due to the type of special reporting procedure are confounded with the serious-
ness estimates which themselves are a function of several other confounding forces operating in the experiment.

The principal result of this widespread confounding is that it is virtually impossible to know the subjects' perceptions of the relative seriousness of the various types of offenses, the relative seriousness of not reporting the offenses, or the probability of reporting. Neither can we isolate the effects of the various consequences of reporting, the complication factors, or the special reporting procedures. Hence, Johnson's interpretation of the major findings, and especially his strong statements about the relative effectiveness of the special reporting procedures, must, depending upon the extent of confounding, be considered as somewhere between extremely tentative and completely unjustified.

In addition to the two issues just discussed, I have several narrower questions and concerns about other aspects of this study. I shall list briefly three of these issues but, in the interest of time, shall not attempt to discuss them in detail. First, the fact that each subject made a series of estimates for each offense is troublesome. For example, Johnson suggests that the seriousness of not reporting and the probability of reporting were greater when a special reporting procedure (SRP) was assumed. Our confidence in this interpretation would be increased if the pre-SRP and post-SRP estimates had been made by different experimental groups. Further, what the subjects assumed about special reporting procedures when they made their pre-SRP estimates is unclear. Did the company for which the subjects worked have some type of reporting procedure which resembled one of those investigated? Second, Johnson was concerned with the question of whether "a person" would report the offense or irregularity. It is not clear which person (e.g., themselves, a close friend, etc.) the subjects should have assumed (or did assume) in making their estimates. Nor is it clear who was supposed to receive the reported information (e.g., the auditor, the audit committee, the police, etc.). I think it is important to specify the experimental setting in some detail, not only for the reader of the report but also for the subjects in the study. Such specification increases the likelihood that all subjects will respond to the same set of hypothetical facts. Third, there is simply too little information reported in the paper. I would have liked to have known, for example, the exact wording of the twenty-one offenses as presented to the subjects. In addition, the presentation of only means and medians tends to obscure any potentially enlightening relationships inherent in the data. Thus, even if Johnson's study had been executed
flawlessly, the exact nature of the implications that could be drawn from the findings would be unknown because of the sketchy presentation of results.

Finally, I would like to mention two issues relevant to this type of research which were not considered adequately by Johnson but which are extremely important if one contemplates further research along the lines pursued by Johnson. First, Johnson accepts with enthusiasm the magnitude scaling procedures developed by Stevens, and the other researchers referenced by Johnson evidence similar enthusiasm. There is, of course, another side, and this opposing viewpoint has been ignored by Johnson. A rather large literature exists which critically examines psychophysical measurement in general and Stevens' procedures in particular. As you might suspect, much of this literature revolves around the questions of whether sensory processes can be measured and, if they can, whether the various measurement methods produce reliable and valid results. A comprehensive critical evaluation of psychophysics, including the work of Stevens, is provided by Savage (1970). Johnson's paper would have been improved had he attempted to deal with the objections to psychophysical measurement. At the very least, he should have recognized the existence of a large literature that is skeptical of psychophysical measurement.

A second, and related, issue concerns the reliability and validity of psychophysical scaling procedures. By reliability I refer to the internal consistency or the consistency over time of the responses generated by psychophysical scaling procedures. By validity I mean the correspondence between the subjects' responses and some independently-measured criterion. Reliability and validity are old problems in virtually all areas of psychology, and we are not going to settle any of the related issues today. Since questions about the reliability and validity of psychophysical scaling procedures are often raised, however, I think it is important not to overlook the lessons that the past few decades of psychological research have to offer.

When validity cannot be shown directly, there often is a tendency to place great emphasis on reliability. However, while high reliability certainly increases the chances of obtaining valid measurements, reliability can sometimes be misleading. Reliability could be caused by several factors (e.g., a consistently-applied anchoring-and-adjustment response strategy) that are unrelated to validity.

An excellent example of the lack of relationship that sometimes exists between reliability and validity is provided by Garner (1954). Garner's study involves an investigation of loudness estimates—the
same topic studied by Stevens in his first published work using magnitude scaling procedures. Garner had three groups of subjects listen to 600 pairs of tones. The first tone of each pair always had a loudness of 90 decibels. The second tone of each pair varied from 55 to 65 decibels for Group 1, 65 to 75 decibels for Group 2, and 75 to 85 decibels for Group 3 (with 100 repetitions of each odd number of decibels in the range). The subjects' task was to estimate whether each variable tone was more than or less than half as loud as the standard tone. Garner's results showed that all three groups of subjects gave highly reliable but totally invalid responses. For each group the median and mean values used to discriminate “less than half as loud” from “more than half as loud” were almost exactly the midpoints of the ranges of variable stimuli they were given (60, 70, and 80 decibels for the three groups).

Further analysis showed that the subjects developed their strategies for making half-loudness estimates early in the series of 600 trials and applied these strategies very consistently over the entire series. In Garner's words, “... even though we have evidence that the [subjects] do not know what they are doing ..., they do in fact do something quite consistently, and judging from the reliabilities, feel quite sure that they know what they are doing.” (Garner, 1954, p. 222) He went on to state:

In brief, these experiments have shown that [subjects] can exhibit a very high reliability for a type of judgment for which there is no validity. When [a subject] says that A is half of B, we can believe that he will continue to say the same thing in this and in similar situations. But we cannot believe that A is in fact half of B. (Garner, 1954, p. 222)

Garner was able to make statements about validity because he studied a stimulus which could be measured independently of the subjects' responses. This happy circumstance does not exist in the case of seriousness estimates, making the direct validation of seriousness estimates an impossible task. Both Stevens and Garner have commented on the problem of validating sensations. Stevens had this to say:

Sensations do not come with numbers written on them, and when we try to assess the ratio between a pair of them we find ourselves up against a difficult task of appraisal. It is no wonder then that subtle constraints and biases can influence the result. This is another way of saying that the outcome is a function of method—as it always is in science. What we want, of course, is an unbiased method, one that on the average lets [the subject] make an estimate
that is neither too high nor too low. Since we do not know in advance what his estimate should be, we can apply no independent criterion of validity. (Stevens, 1956, pp. 24-25)

For our purposes, however, I think Garner’s comments are much more helpful.

Direct validity of the numerical responses is impossible, because we have no independent measure of the sensory process itself. This fact in itself has often been used as an excuse for not attempting a validation. There are, however, other ways of getting validation. Validation can be obtained by using converging operations to arrive at a single construct or concept. If two or more independent sets of data, involving basically different indicators of the nature of the sensory process, lead to the same sensory scale, then we have a form of validation. Such validation is probably the only meaningful kind in this and in other areas of psychology. (Garner, 1954, p. 223)

The type of validation that Garner was proposing in 1954 has become fairly commonplace in many areas of psychology today, perhaps primarily as a result of a very influential article by Campbell and Fiske (1959) in which the notion of convergent validation was developed more extensively. Relating this to Johnson’s paper, it seems clear that one’s confidence in the validity of seriousness estimates would be greatly enhanced by the demonstration that such estimates possess convergent validity. This means, of course, that various types of research methods, from traditional psychophysics and elsewhere, should be applied to the research question addressed by Johnson before we accept Johnson’s results as valid.

Earlier I referred to certain decision-making heuristics and biases which may distort a subject’s responses in studies such as this. Another bias which has been found to distort one’s evaluations is the hindsight bias (Fischhoff and Beyth, 1975). Generally, the hindsight bias entails that individuals who possess knowledge of what has actually occurred in the past will systematically misperceive how they would have acted before such knowledge became available. I think that one group that continually falls prey to the hindsight bias consists of discussants of research papers. As one who has conducted some experiments and has occupied the role of presenter instead of discussant, I have been in a good position to observe the occurrence of this bias. However, I usually have found a few bits of information among all the noise generated by the discussant which I could use to improve future experiments and which made listening to the discussant almost bearable. I hope that Johnson will have the same experience as he considers my comments on this paper.
REFERENCES


