ANALYSIS O... ...M I...
IMPUTATIO... ...D E...
1989 FARM C... ...AN...
SURVEY FOR ...WA...
CAROLINA

Stanley J. Hoge
Diane K. Willi...

... ...ION AND EDITING IN THE 1989
... IOWA AND NORTH CAROLINA, by
... Research and Applications
... Statistics Service, U.S.
... D.C. 20250. July 1991.

...

... ...ntation was conducted during
... document the extent to which
... ...diting and item imputation on
... ...es from the 1989 Farm Costs
... ...n Iowa and North Carolina.
... ...ted for edited, missing, and
... edit actions were coded.
... ...er edit values were used to
... ...diting on the estimates.
... ...data from one cell to another
... ...s of misplacement. A small
... ...were changed, as little as 5
... ...most of the change in the
... ...t changes had no effect on
... Recommendations include
... ...rpose of editing relative to
... ...vey estimates versus internally
... ...new editing strategy should be
... ...clear, consistent editing
... ...to be achieved by editing.

```
*********************************
...ed distribution to      *
...he U.S. Department       *
                           *
*********************************
```

...

... Iowa and North Carolina State
... ...ticipated in the imputation
... ...e people in headquarters who
... the support of the Economics
... ...nitiation. These included
... ...Statistical Methods Branch,
... ...l thanks to Kay Turner who
... ...upport and encouragement of
... ...ie Atkinson throughout the

Table of Contents

Summary

The purpose of this study was to define, document, and understand item nonresponse, imputation, and editing on the Farm Costs and Returns Survey (FCRS), as well as to evaluate their impact on the estimates. This study pointed out that item imputation is completely reliant upon and not separate from the editing process.

Completed questionnaires from the 1989 FCRS were reviewed in Iowa and North Carolina. "Prior-to-edit" data values were collected for edited, missing, and imputed (EMI) items, and reasons for the edit actions were coded. These reasons were categorized into errors of omission or misplacement, imputation for responses of "don't know" or refusals, correction of implausible or physically impossible data, and problems in allocating a reported total into requested component parts.

A very few reports accounted for most of the EMI's effect on the estimates of key expenditure items. In addition, half of the edit changes had essentially no effect on the major survey estimates selected. That is, many small changes were made which had little effect on the final results.

Nearly half of the edits moved respondent-reported data from one cell to another cell on the questionnaire. These edits had little or no effect on survey estimates. The same results occurred for detailed editing for incomplete allocations. While these types of editing provide for internally consistent individual records, they appear to add little to the quality of the aggregate estimates.

Based upon these conclusions, we recommend reconsideration of the volume of manual editing on the FCRS, as well as the purpose it serves. If priority is placed on the published estimates, the results of this study indicate limited gains from much of the editing. The current practice of detailed manual editing by clerks and statisticians to provide internally consistent multivariate records should be re-evaluated relative to the extensive staff and computer resources and costs associated with editing. Finally, an editing strategy should be developed and administered through clear, consistent editing instructions, which reflect priorities relative to the gains to be achieved by editing.

The new Agency owned State Statistical Office microcomputer local area networks provide a technological opportunity for the development of an interactive editing system, possibly including multivariate relationships, for the FCRS.

Introduction

Nonresponse may affect the results of any sample survey. Entire units may be lost to a sample through their inaccessibility or refusal to participate. Even when a response is obtained, individual items may be missing for a variety of reasons. Nonresponse is of special concern on the Farm Costs and Returns Survey (FCRS) due to its length and difficulty. The FCRS is a nationwide survey of farm expenditures and income conducted annually by the National Agricultural Statistics Service (NASS) in cooperation with the Economics Research Service (ERS).

Unit nonresponse on the FCRS in the form of inaccessibles and refusals is monitored and documented. It is accounted for in the estimates through adjustment of the sample expansion factors. However, item nonresponse has not previously been measured or formally documented. This project seeks to fill this void of information by defining the problem of item nonresponse on the FCRS, exploring its causes, measuring its existence in the survey process, and evaluating its impact on survey estimates.

Item nonresponse occurs when some responses to items on a questionnaire are not available. Causes of item nonresponse include item refusals, "don't know" responses, omissions, and values deleted in editing. Omissions may be due to respondent or interviewer errors, where answers are omitted by mistake or skip patterns not followed correctly, or to illegible answers. Values provided by respondents may also be found to be logically impossible or implausible and deleted from the sample.

Deletion of implausible values occurs during the editing process, when errors in survey responses are detected and a determination is made regarding how they are to be handled. The handling of errors includes checking for coding or keying errors, and perhaps recontact of the respondent to obtain the correct response. Besides correcting errors, the editing process also includes tasks such as validating totals, entering "office use" codes, and verifying that skip patterns were followed correctly.

Item nonresponse may be handled by deciding to designate the entire sampled unit as a nonrespondent, resulting in missing data for an entire record. However, loss of an entire unit from the sample is undesirable. An alternative is to impute acceptable values for missing or incorrect items.

Sande (1982) wrote "the real problem of imputation is the interaction with editing." Editing and imputation are inextricably linked. For the purposes of this research, the link between missing items, editing, and imputation must be clear. This is particularly important in two instances. The first instance is where missing items are allowed to remain missing. This action is equivalent to imputing a value of zero. The second instance is where respondent-provided values are implausible or incorrect and are "edited out". In this case the values are made missing first before being replaced with another value, an imputation.

1

Although this editing practice is considered one activity, it consists of three separate actions:

1) identification of incorrect values (including the value of zero when items are not obtained from the respondent);

2) deletion of implausible or incorrect values, causing an item to become missing; and

3) imputation of now missing values with new values.

This study originated with the purpose of identifying those items for which missing data were imputed by the office staff. As the link between missing items, editing, and imputation became clear, it was evident that all edited items should be reviewed. Thus, we coined the term EMI, meaning "edited, missing, or imputed," to refer to the items studied.

Purpose

FCRS questionnaires are manually reviewed at two levels before data are keyed and loaded to the mainframe computer for automated editing. Often each supervisory enumerator reviews the work of field enumerators, while survey statisticians in each State Statistical Office (SSO) review each questionnaire as it comes in from the field and prior to key entry. At least some "hand" editing and "manual" imputation are done by the SSO personnel. Data values can be corrected or deduced based on logical relationships with other data, before the data are objectively machine edited. Errors flagged by the machine edit may also be corrected by the survey statistician in the SSO. This editing or imputation for missing values is a necessary step in preparing the data for summarization. However, the process relies on the knowledge and judgment of field and office staff and may provide inconsistent results.

The consistency of imputed data may be enhanced by using an alternative editing strategy or automated item imputation routine. While discussion of specific imputation schemes is beyond the scope of this paper, ongoing research conducted by the Survey Research Branch may support the development of item imputation methodology for the FCRS. Recent multivariate correlation analysis of the FCRS data by Bargmann, Donaldson, and Turner (1991) may provide statistical models for imputing missing values or for multivariate editing algorithms. In addition, interactive editing, such as the Blaise system of the Netherlands Central Bureau of Statistics, would enable real time editing, imputation, and review of FCRS data. However, no systematic documentation of missing items on the FCRS has previously been available. Nor have we known the extent to which edited, missing, or imputed items affected the estimates from the FCRS.

Documentation of the EMI items on the FCRS and the evaluation of their impact on the estimates may be used to guide development of strategies

2

for dealing with item nonresponse. These may include a review of questionnaire design, data collection procedures, and alternative editing and imputation schemes.

Therefore the specific purposes of this project are:

- to define item nonresponse on the FCRS for research purposes and for data users in NASS and ERS;

- to document the extent to which item nonresponse occurs and for what reasons;

- to understand the causes of item nonresponse, through the analysis of reasons for survey statisticians' edit actions and "manual" imputation;

- to evaluate the impact of editing and imputation of data on selected major expenditure estimates published from FCRS data;

- to set the stage for further research into alternative strategies to impute missing or questionable responses on the FCRS.

Literature Review

As concern for survey data quality has increased and survey resources have become more limited, research has been undertaken by several survey organizations worldwide to review the effect of editing. A study of the World Fertility Survey (WFS) evaluated the results of machine editing (Pullman, Harpham, and Ozsever, 1986). Field and office editing were not part of this study, as it evaluated only the editing that was done on errors flagged by the computer.

Machine editing itself was shown to have very little substantive effect and no statistically significant effect. A series of univariate, bivariate, and multivariate analyses on both raw and clean data, that is, the data before and after editing, showed virtually no difference in distributions, estimates, or inferences. Insensitivity to editing was also found in an evaluation of a multiple regression equation that included variables most subject to editing. According to the authors, the effect of editing was almost always less than sampling error.

The authors concluded that the average delay of twelve months in the processing and release of WFS results was due to ineffectual machine editing. Although they recognized the necessity of a clean and internally consistent micro-data set to meet the expectations and needs of data users, they recommended the publication of preliminary reports before intensive data cleaning, as key figures differed from final values by only one to two percent. The authors attributed the limited effect of machine editing to the training of interviewers and the care taken during field and office editing, but fell short of questioning the effect of these earlier editing actions. They suggested that such a good job was done on the hand edit that a machine edit was superfluous.

The Australian Bureau of Statistics did several independent evaluation studies of editing on economic surveys with results that consistently supported revised editing methods (Linacre and Trewin, 1989). One of these studies was on the 1983/84 Agricultural Census. This study differentiated between substantive manual edits, which involve changes to data values reported by respondents, and nonsubstantive edits, such as improving legibility or rounding. More than 85 percent of the edits were substantive, with nearly three-quarters of the units receiving at least one substantive edit. Although large units were to receive more extensive manual checking than small units, this did not appear to be the case in practice. The study showed that for the selected variables, appropriate automated consistency checks by computer would have caught nearly all the errors identified by hand. There were many very small changes and few large changes.

Editing resulted in substantial changes in the value of some estimates, while others, despite extensive editing, were little affected. The study showed that restricting actual changes in items to large units accounted for substantially the entire effect of editing, and that computer editing for small units was sufficient. Finally, in a review of the three editing studies completed by the Bureau, the authors

4

concluded that "more emphasis should be placed on the accuracy of published aggregates in using resources, rather than ensuring the accuracy of each contributing record."

The Netherlands Central Bureau of Statistics studied editing on two social and two economic surveys, one of each was large while the other was small (Bethlehem, 1987). The purpose of the study was to better understand the editing process and to evaluate the benefits of editing relative to its costs.

The study found that editing problems in all four surveys were of the same nature. Specifically, the major findings included the following:

•   The data were handled by many different people from different departments, from interviewers and respondents to clerks, key entry operators, programmers, and subject matter specialists.

•   A lot of time was spent on clerical editing, simply cleaning up the forms in preparation for data entry, with little correction of errors or improvement of data quality.

•   The repetition of the editing cycle through data entry, computerized checking, and manual correction was very time-consuming.

The results of this study led to the development of the highly successful BLAISE system for survey processing at the Netherlands Central Bureau of Statistics. BLAISE integrates the survey process, from data entry to interactive editing, into a structured system.

The United States Bureau of the Census evaluated the editing and imputation procedures used in the 1982 Economic Censuses (Greenberg and Petkunas, 1986). They studied the procedures and their use by an automated audit trail which recorded each edit pass through the data, noted the original reported data, and identified the source of any imputation. Difficult, large, or unusual cases were targeted for manual review by a clerk or an analyst, who may also provide corrections or imputations.

The study found that for most of the variables, reported data were usually retained, and imputation for missing data was far more common than were changes to reported data. The proportion of cases imputed was greater than the proportion of the estimate due to imputation. For example, for one variable 20 percent of the cases received imputation for missing values, but these accounted for only 5 percent of the estimate for that variable. Approximately 5 percent of the cases where reported data were changed contributed more than 90 percent of the total change in the estimate.

Finally, while the authors found the "interplay of automated routines and individual review (to be) an effective strategy both in the use of resources and treatment of establishment data records" in editing and

imputation, they recommended consideration of an on-line interactive editing system as a more efficient, streamlined process.

The National Agricultural Statistics Service studied manual data imputation on the 1976 December Enumerative Survey (DES) in Oklahoma (Bosecker, 1977). All data edited in the DES as well as the sources of the edited data were observed and evaluated. However, data were considered imputed only when the response code indicated refusal or inaccessible; thus imputation on incomplete questionnaires was not studied.

Results showed that while nearly 10 percent of the DES questionnaires were refusals or inaccessible, only three of the twenty estimates exhibited amounts of imputation greater than 10 percent. Imputed data contributed nearly 12 percent of the tract indication for total land and approximately 11 and 13 percent of the weighted indications for bulls and replacement beef heifers, respectively. These results like those of the other studies, suggest that the impact of editing and imputation on most survey indications is small relative to the proportion of reports edited.

In summary, nearly all of these studies:

- had difficulty accounting for "simple" clerical edits, such as unit or rounding conversions, within their scheme of study of missing or edited data;

- found many small changes, which had little overall effect on estimates;

- recommended that more editing be done by computer, such as unit conversions, summation to totals, or changes within tolerance limits;

- suggested that emphasis be placed on editing to improve the quality of the estimate (i.e., to reduce the error in an estimate), rather than just to provide internally consistent individual records;

- recommended on-line interactive data entry and editing.

Methodology

Data for this project were collected immediately after the 1989 FCRS in the Iowa and North Carolina SSO's. All useable questionnaires from the survey were reviewed for items that field or office personnel indicated as missing or that were changed during the office edit process. These edited, missing or imputed (EMI) items and pertinent data were captured using an automated entry program. A total of 769 questionnaires were reviewed, 448 in Iowa and 321 in North Carolina.

Items reviewed were limited to the income and expenditure portions of the expenditure, farm operator resource (FOR), wheat cost of production (COP), and dairy COP versions. Administrative and office use items were considered beyond the scope of this project. The original value prior to the office edit, a reason code describing its EMI status and the complete ID were obtained for each EMI item identified. Thirteen different reason codes were used to categorize why items were changed or imputed. In order to fully describe item nonresponse all edit actions were captured, including clerical edits and updates of computer flagged errors. Every effort was made to categorize as many recurring situations as possible.

EMI's were divided into two groups. One group consisted of items where the "prior-to-edit" values were positive and the other group, where the "prior-to-edit" values were missing. This division provided a rough separation of missing items from edited items. Imputation can occur in either situation. Six reason codes were assigned to each group. A thirteenth reason code was added as an "other" category, which required additional comments to allow review and recategorization after the data collection in the SSO's. Detailed definitions of the reason codes appear in Appendix A.

Analysis of items where a value was changed during the edit process is quantifiable in terms of direct impact on the estimates, indices, and models. Such items, however, comprise only part of the problem of imputation. The other part of the problem is the impact of the failure to impute for items where the need for imputation is indicated. This second part of the imputation problem can only be inferred from its frequency of occurrence. This paper discusses both parts of the problem. Direct impact on the estimates is discussed by examining the relative differences between the final summarized values and the prior-to-edit values. The indirect effects are discussed through an examination of EMI item frequencies.

Edited, Missing, and Imputed Item's (EMI) Direct Impact on Estimates

Direct change was measured by taking the difference between the "prior-to-edit" value entered in the field and the final edited value. For the ith observation in stratum h, the difference was measured as

$$D_{hi} = E_{hi} - O_{hi}$$

where stratum h = 1,....,L and unit i = 1,....,n_h.

E_{hi} = Final edited response recorded by the office staff (the summarized value).

O_{hi} = Original field level response recorded by enumerator or supervisory enumerator.

Positive differences indicate the summarized item value was larger than the original response.

Net differences were summarized for the following major aggregate items:

Total Expenditures
Farm Services Expenses
Interest Expenses
Feed Expenses
Fuel Expenses
Fertilizer Expenses
Total Land Operated
Market and Storage Expenses.

Summarized differences were evaluated on two bases, effect on the edit and effect on the estimate. This type of analysis was motivated, in part, by Boucher (1991).

EMI effect on the edit was measured by ordering the expanded differences by record from largest to smallest, based on absolute value, and then dividing the cumulative difference at various points by the total summarized difference. This statistic relates the contribution of individual editing changes to the overall effect of editing. For the ith observation in the sample the EMI effect on the <u>edit</u> is measured as follows:

$$EFF_{edit} = \frac{\sum_{i=1}^{p} DEXP_i}{\sum_{i=1}^{n} DEXP_i} \times 100$$

where

DEXP$_i$ = the expanded difference for an individual record.

8

EMI effect on the estimate was determined by dividing the direct expansion less the cumulative differences, as described above, by the state level estimate. This statistic measures the overall effect of editing relative to its impact on the state expansions. The EMI effect on the estimate is measured as follows:

$$EFF_{estimate} = \frac{T - \sum_{i-1}^{p} DEXP_i}{T} \times 100$$

where

i = 1,.......n and p is an arbitrary breakpoint less than or equal to n,

and

T = the total direct expansion for the state.

Results

Impact of Relative Differences

Tables 1a and 1b represent the relative differences and their univariate significance levels in before and after editing expansions accounted for by edited, missing and imputed items on major summarized aggregate expenditures. Differences were adjusted to exclude contractor-reported data and items with only clerical edits (decimals, dollars and cents, etc.). Multivariate testing was not performed for the selected items even though the relative differences for total expenditures were significant in both states. The interdependency of total expenditures with its aggregate parts and the lack of estimation precision in the relative differences for the aggregate parts made this type of testing uninformative.

Relative differences for total expenditures were significant in both Iowa and North Carolina (with p-values < .05), even though a majority of the questionnaire items were not changed during the edit. The difference between total expenditures before and after editing was equal to zero for two thirds of the completed questionnaires in Iowa and nearly half of the completed questionnaires in North Carolina.

Although statistically significant, the net effect of editing and imputation on the estimate for total expenditures in Iowa was very small. If no EMI changes had been made the state level expansion would have been within one percent of the final estimate. In addition, the total EMI effect on the edit for total expenditures is accounted for by only 59 percent of the questionnaires that had changes. There was no change in the estimate after 5 percent of the largest differences were included. EMI effect on both the edit and estimate is shown graphically for both states in Appendix C.

TABLE 1a. Relative Differences Between Selected Pre-edit and Post-edit Expansions as Percentages of the Post-edit Expansions in Iowa, 1989 FCRS.

Variable	Relative Difference (%)	Significance Level 1/	Count of Q-naire's with Differences
Total Expenditures	1	.01*	162
Farm Services Expenses	1	.06	94
Interest Expenses	· 1	.84	9
Fuel Expenses	· 1	.38	16
Feed Expenses	2	.80	11
Fertilizer Expenses	· 1	.78	18
Total Land Operated	· 1	.42	5
Market & Storage Exp.	- 2	.85	66

1. The symbol "*" denotes significance level of .05 or less.

Total expenditures were affected to a much greater degree by the office edit and imputation procedures in North Carolina than in Iowa. This was primarily due to imputation of contractor refusals and "don't knows," unique to North Carolina. Less than 10 percent of the EMI's accounted for 94 percent of the editing effect for total expenditures. When these larger differences are accounted for, the net effect on the estimate is less than one percent. All of the EMI's in that first decile were due to contractor refusals and "don't knows" that required imputation. Thirty-six percent of the questionnaires with changes in North Carolina accounted for all of the EMI effect on the edit for total expenditures. Even though North Carolina's total relative difference for total expenditures was much larger than Iowa's, the two states demonstrated similar EMI effects on the estimates.

TABLE 1b. Relative Differences Between Selected Pre-edit and Post-edit Expansions as Percentages of the Post-edit Expansions in North Carolina, 1989 FCRS.

Variable	Relative Difference (%)	Significance Level 1/	Count of Q-naire's with Differences
Total Expenditures	22	.01*	196
Farm Services Expenses	< 1	.33	138
Interest Expenses	< 1	.35	10
Fuel Expenses	3	.40	20
Feed Expenses	46	.14	27
Fertilizer Expenses	2	.89	26
Total Land Operated	< 1	.98	10
Market & Storage Exp.	25	.14	102

1/ The symbol "*" denotes significance level of .05 or less.

Due to a lack of estimation precision, feed expenses and marketing and storage expenses indicated no significant difference in North Carolina, despite large real differences. Nonetheless, these items are of interest because of the large indicated effect of the editing process on them. Feed expenses had the largest relative difference of any of the summarized items, but this difference was accounted for by only 27 observations, or 8 percent of the completed questionnaires. Feed expenses contributed 71 percent of the relative change due to editing in total expenditures. As was the case with total expenditures nearly all of the edit effect was explained by the imputation of data for poultry and/or hog contractors that refused or were unable to provide the requested information. Once these contractor refusals and don't knows were accounted for, the relative difference in feed expense decreased to less than one percent. All of the EMI effect on the edit was attributed to item imputation in 63 percent of the questionnaires with a nonzero difference.

The questionnaire section in which marketing and storage expenses are asked was heavily edited in both Iowa and North Carolina. In fact, total crop marketing and storage expense was one of the top five most heavily edited items in both states (see Appendix F). Reasons for such frequent revisions were varied. A few of the reasons for EMI's on marketing and storage expense were 1) marketing expenses were omitted for certain commodities whose production require these expenses, 2) the value was not transferred or transferred incorrectly from the worksheet, 3) worksheet computations were done incorrectly, and 4) the respondent was unable to provide all the necessary information. Relative differences for North Carolina's marketing and storage expenses were similar to Iowa's. There doesn't seem to be one particular reason why items in this category were changed.

Ten percent of the differences for marketing and storage expenses in North Carolina expanded to 90 percent of the state's EMI effect on the edit. Almost half of these questionnaires were changed to cover marketing expenses related to tobacco allotments. However, there were also errors in computation on the worksheet and errors of omission, where a commodity warranting an imputation of marketing expense was reported elsewhere on the questionnaire. If editing for marketing and storage expense had been limited to the 10 percent of the EMI's with the largest expanded differences, the total difference in the survey estimate due to editing would have been less than 2 percent. Ninety-five percent of the editing effect is accounted for by less than 33 percent of the questionnaires changed. The estimate from the partially edited dataset would have been within 1 percent of the estimate from the fully edited data set, if only these edits were included.

Frequencies and Reasons for Editing

Frequency analyses of the EMI data were completed to document the reasons for editing and the results of edit actions in terms of the direction of change to reported data. The analyses also explored the relationship between editing and farm size as indicated by economic class and stratum. The results are reported in this section. Results of more detailed analyses of the distribution of EMI's, including consideration of selected individual items, appears in Appendix D. Only EMI's for operations qualifying as farms were analyzed. Clerically edited data and data reported by contractors were excluded.

Figure 1 displays the relative frequencies of the various reasons identified by the research review team for editing, missing, or imputed data item designation. The results show that more than half of the editing in both states is done to correct errors of omission and misplacement. Detailed definitions of the reasons and their coding scheme appear in Appendix A.

FIGURE 1. Reasons for EMI Data Items

Percent of EMI's
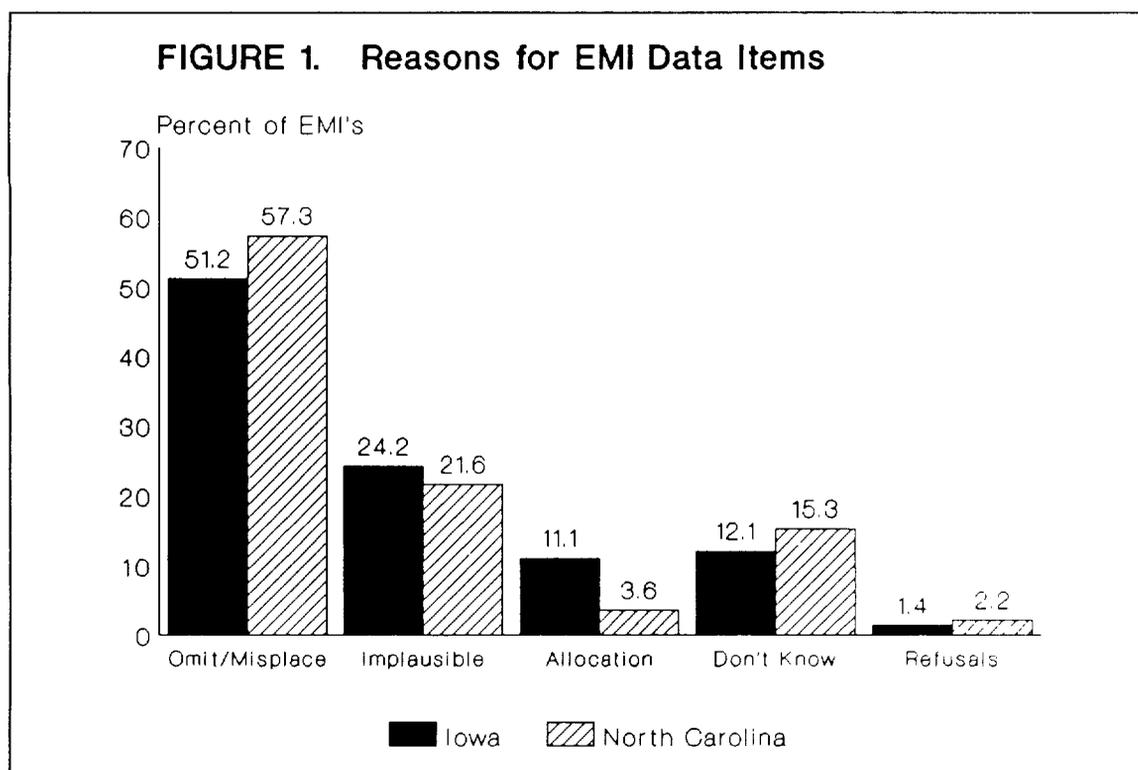
Omit/Misplace:	Reason Codes 1, 5, 11, and 13
Implausible:	Reason Code 2
Allocation:	Reason Codes 3 and 10
Don't Know:	Reason Codes 4, 7, and 9
Refusals:	Reason Codes 6 and 8

Data that had been placed in the wrong cell were classified as omitted from the correct cell. The item cell in which the data had been placed was assigned reason code 1, since its original value had been positive. The target cell receiving the misplaced data after editing was assigned reason code 1 if its prior-to-edit value was positive or reason code 5 if it was originally blank. While reason code 5 also included typical errors of omission due to overlooked questions or incorrectly followed skip patterns, these occurred infrequently. Therefore, the results suggest that much of our editing effort simply moves respondent-provided data from one cell to another on the questionnaire.

Reason code 11 denotes a special circumstance of omission in which individual items with positive responses were deleted in edit when an "R-box" was coded for a section. There were several R-boxes in the latter half of the questionnaire that were associated with sections of internally related items. A coded R-box indicated that at least one item in its group was missing, even though positive responses for others were to remain. In other NASS surveys, all items are edited to zero in a section that has been noted as incomplete. Confusion with the use of the R-box resulted in the loss of valid data, since the more familiar approach of editing out all data in an incomplete section was sometimes erroneously followed for the FCRS. A detailed look at the lost data items appears in Appendix D.

One-fourth of the EMI's in Iowa and one-fifth of the EMI's in North Carolina reflected the correction of implausible, illogical, or physically impossible reported data. "Implausible" does not necessarily mean "outrageous." It means the reported data appeared to be incorrect relative to other reported data. Reason Code 2 for "implausible" may have been assigned for errors as simple as reporting in incorrect units or a miscalculation, in addition to more substantive errors.

Allocation problems appear to have been more common in Iowa than in North Carolina. Allocation situations arise when the respondent provided the total of several items in one cell, but either refused or was unable to provide the requested individual items. Reason code 3 was recorded to indicate allocations that were completed through imputation by SSO staff. That is, SSO staff separated the respondent-provided total into the indicated parts.

In North Carolina, indicated incomplete allocations were imputed by the SSO staff. In Iowa, however, such imputations were completed only 28 percent of the time. The proportions of all EMI's represented by imputed allocations (reason code 3) were almost equal in both states. However, in Iowa, 72 percent of the allocation EMI's were left as obtained from the field, with the respondent-provided total in one cell and the indicated components remaining blank. All cells associated with an allocation which remained incomplete after editing were assigned reason code 10.

Twelve percent of the EMI's in Iowa were due to a "don't know" (DK) response; the corresponding figure was 15 percent in North Carolina.

The slightly higher percentage in North Carolina was largely due to frequent DK responses by contractee-respondents for contractor expense data.

In Iowa, there were 359 responses of DK indicated by the respondent or, in a few cases, noted as suspect by the enumerator. These accounted for 12 percent of the EMI's. Two out of every five DK's remained blank in the edited data set. In North Carolina, there were 483 indicated or enumerator-suspected DK's, accounting for more than 15 percent of the EMI's. Nearly one out of every three remained blank in the final edited data set.

Finally, item refusals appear to have been rare. Only 1.4 percent of the EMI's in Iowa and 2.2 percent of the EMI's in North Carolina were indicated to have been refusals. The greater frequency in North Carolina may be attributed to refusals by contractors to provide contractee expense data.

Detailed analyses of EMI's due to allocation problems, DK's, and refusals appear in Appendix D. These analyses identify individual items affected by EMI's. Appendix D also includes a detailed look at items affected by the miscoding of R-boxes.

Although expense data provided by contractors has been excluded from the general frequency analysis, a detailed examination of EMI contractor expenses provides an indication of the quality of these data. In North Carolina, over half of the positive entries for contractor expense data in the final edited data set was imputed for DK or refused items. Nearly all of the remaining positive data was reported by contractors rather than by the farm operator-contractees. That is, seldom was contractor expense data obtained from the farm operator, who from the standpoint of FCRS instructions is the preferred respondent. Consideration of selected individual contractor expense items is included in Appendix D.

Table 2 documents the results of the various edit actions taken to increase, decrease, or leave unchanged the value reported by the respondent. More than half of the time, edit actions resulted in an increased item value. In most of these cases a blank was replaced with a positive value. That is, a data value possibly reported elsewhere in the questionnaire is entered into a blank cell to correct an omission. The reported value decreased in roughly a third (28 percent in Iowa, 36 percent in North Carolina) of the EMI's. The primary reason for this action was to correct errors of misplacement. Decreased item values also occurred in allocation situations where a data value from a single cell was divided to fill more than one cell.

Corrections for implausible or illogical values were almost equally likely to have increased a reported value as to have decreased it. A detailed tabulation of edit action by reason code appears in Appendix E.

15

TABLE 2. Direction of Change of Data Values Resulting from Edit Actions by SSO Staff.

| | % of EMI's | |
Edit Action	IA	NC
Total number of nonclerical EMI's 1/	2939	3167
Reported value was decreased	28.5	36.4
Reported value > 0, edited value > 0	14.2	14.1
Reported value > 0, edited value = 0	14.3	22.3
Reported value was unchanged	16.3	8.3
Reported value = 0 = edited value	11.0	6.3
0 < reported value = edited value > 0	5.3	2.1
Reported value was increased	55.2	55.3
Reported value = 0, edited value > 0	41.2	42.8
Reported value > 0, edited value > 0	14.0	12.5

1/ This is the number of EMI's after clerical edits were removed and contractor-reported data were accepted and not considered editing.

The difference in relative frequencies between Iowa and North Carolina for EMI's where the reported value remains unchanged is primarily due to reason code 10 in Iowa. When the allocation of a total into parts remained undone, blanks remained blank and positive values remained unchanged. Other missing values that remained missing were due to DK's and refusals. Items where a reported positive value remained unchanged were likely keypunch errors that required corrective action, since the reported value was actually the final edited value, as well. There appear to be no more than 53 such errors in Iowa and 65 in North Carolina.

Table 3 documents the amount of editing done per list frame expenditure version questionnaire by economic class and stratum. The figures in this table offer a sense of where we are spending time editing. In both states, the average number of EMI's per questionnaire increased with increasing economic class and farm size as indicated by stratum-level summarization. However, expressed as a percent of the number of positive responses per questionnaire, the amount of editing in Iowa did not appear to vary much across economic classes. An exception occurred for the largest farms, where the average increased to 7.4 percent. However, the proportion of items edited did increase by stratum. In North Carolina, it is more evident that the amount of editing within a questionnaire increased with economic class. This is shown graphically in Figure 2 on page 18.

TABLE 3.   Average Number of EMI's 1/ and Average Number of Positive
Responses per Questionnaire, by Economic Class and Stratum, Expenditure
Version, List Only.

Iowa:

Economic Class	Mean Number EMI's/Q-naire	Mean Number +'s/Q-naire 2/	Mean Percent Items Edited Per Q-naire
$1,000 - $9,999	5.7	95.2	6.0
$10,000 - $39,999	6.1	104.7	5.8
$40,000 - $99,999	7.8	130.7	6.0
$100,000 - $249,999	7.4	132.3	5.6
$250,000 +	10.6	142.5	7.4
Strata			
All COP Strata 3/	8.1	137.6	5.9
75	5.0	111.9	4.5
80	6.8	123.6	5.5
85	5.5	108.0	5.1
90	9.4	137.6	6.8
95	11.8	132.8	8.9
Total	8.1	129.4	6.3

North Carolina:

Economic Class	Mean Number EMI's/Q-naire	Mean Number +'s/Q-naire 2/	Mean Percent Items Edited Per Q-naire
$1,000 - $9,999	5.1	77.5	6.6
$10,000 - $39,999	8.4	85.4	9.8
$40,000 - $99,999	13.4	101.4	13.2
$100,000 - $249,999	11.9	116.8	10.2
$250,000 +	15.7	120.5	13.0
Strata			
All COP Strata 3/	13.8	106.9	12.9
75	7.7	94.7	8.1
80	13.6	113.3	12.0
85	10.5	75.0	14.0
90	11.4	99.1	11.5
95	4/	4/	4/
Total	11.5	101.6	11.3

1/   After exclusion of clerical edits and contractor-provided data.
2/   A total of 853 items per questionnaire were reviewed.
3/   In Iowa, COP strata were 55, 56, 57. In NC, COP strata were 53-58.
4/   Included with strata 90.

A chi-squared test of independence indicated that the direction of edit actions in both states was not independent of farm size. That is, whether original data were increased, decreased, or left unchanged in editing did depend on farm size. In particular, in Iowa, the proportion of items edited to zero appeared to decrease as economic class increased. However, the proportion of missing items remaining missing in the edited data set increased with increasing economic class. In North Carolina, reported values were increased with slightly greater frequency on farms in both the smallest and largest economic classes. Values were imputed for missing items more frequently on the largest farms, while reported positive values were more likely to be increased on the smallest farms. These results suggest that we edit different-sized farms differently.

Figure 2. Mean Percent of Items Edited Per Questionnaire
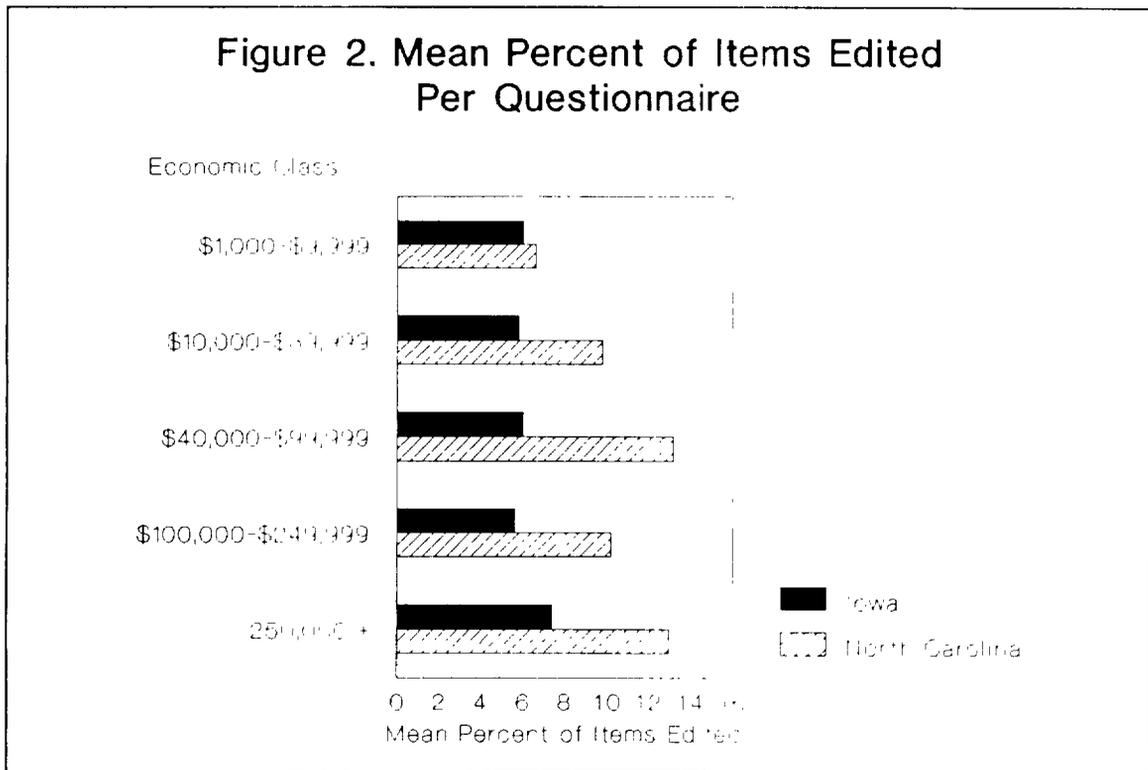
Discussion

Direct Impact Analysis

Results from the impact analysis showed that a few EMI's with large expanded differences accounted for virtually the entire effect of the edit process. In all situations studied, the total change in the estimate was accounted for by less than 50 percent of the largest EMI's. In some cases no further change was observed after 5 percent of the largest EMI's were applied.

North Carolina had several aggregate items analyzed which showed large EMI effects on the estimates. Few reports with large expanded EMI differences carry most of the EMI effect on both the edit and estimate. In some cases the larger differences resulted from identifiable and recurring situations. The larger differences in North Carolina's total expenditures and feed expenses were due entirely to contractor nonresponse.

Even though differences which resulted from editing marketing and storage expenses were statistically insignificant, due to a lack of estimation precision, they are of practical concern. Marketing and storage expenses were among the most heavily edited items that we observed in both Iowa and North Carolina; however, the total expanded difference due to editing relative to the state estimate was much greater in North Carolina. This between state difference in the extent of editing marketing and storage expenses reflected both agricultural differences and differences in the edit procedures themselves. Special handling of contractor data and tobacco allotments unique to North Carolina generated EMI's. Other situations requiring special handling, also more prevalent in North Carolina, usually caused items to be cross checked and resulted in additional imputation. The resulting blank to positive changes generally accounted for the differences of greater magnitude and had a greater effect on the estimates.

The substantial number of EMI's on marketing and storage expenses was largely attributable to the design of the questionnaire and how each state handled design related problems. Enumerators and respondents had difficulty filling out this portion of the questionnaire. Many times enumerators were not able to get enough information from the respondent to correctly complete the worksheet. In other cases the respondent indicated the type of expense but didn't know the amount, the respondent didn't understand exactly what was being asked, or the expense was simply omitted in error. Additional errors occurred when the enumerator failed to move a response from the worksheet to the response cell or when worksheet calculations contained mistakes. While North Carolina had several reports with expanded differences of greater size, both states had EMI's of this type.

Similarities between the editing philosophies of Iowa and North Carolina begin in the field. Both states instruct field enumerators to review their own work before turning in a questionnaire to the supervisor.

19

Supervisors edit all but the last few questionnaires, which are not field edited due to time constraints. Questionnaires are checked-in prior to an automated edit. They then receive a clerical edit and a detailed manual edit by the survey statistician. After the machine edit, additional changes can be made.

Differences exist in the manner in which survey statisticians manually review questionnaires. Iowa's office edit procedures tended to accept data reported by the enumerator and respondent as complete and correct. Data would only be edited or imputed when absolutely necessary to pass the automated edit. North Carolina would accept reports from the field as correct but not always complete, in part because of the unique problems with data collection for livestock contractors and crops that require special handling. The information recorded elsewhere on the questionnaire appears to have been be more frequently used to cross check items and impute data in North Carolina than in Iowa. Much of the EMI effect on the estimate is explained by the inability of the FCRS questionnaire to handle special situations, requiring SSO's to manipulate data more often during the manual review in the office.

Frequency Analysis

The frequency analysis section of this paper provides documentation of the editing activity in Iowa and North Carolina. Editing is a labor-intensive survey activity. The data presented here describe characteristics of that activity, the reasons that editing occurs, and the outcome of edit actions. They provide an accounting of what actually happened during the edit and review process in these two states. These data are offered as management information which may be used as an indication of data quality as well as to provide feedback on questionnaire design. They can also be used to guide the training of enumerators or SSO staff, preparation of manuals, or other aspects of the survey process.

This analysis documents differences and similarities in the editing practices of these two states. For instance, Iowa's SSO staff tends not to allocate respondent-provided totals into component parts for detail items. North Carolina's SSO staff tends to edit data at a higher rate. These differences raise a question of the value of consistency in edit practices across states. For example, specific instructions for resolving allocation problems would enhance consistent interpretation of these items. One possibility may be to specify the cell into which an unallocated total should be placed, paired with a coding scheme to indicate the cells of its missing components.

The effect of another source of inconsistency was also evident in the frequency results. Instructions for recording item nonresponse differ across surveys. On the 1989 FCRS, the existence of item nonresponse within some sections was to be indicated by coding an R-box. However, positive respondent-reported data in the section was to remain. This procedure contrasts with instructions for the Agricultural Survey

Program (ASP). For example, when the nonresponse cell in the ASP Livestock Section is coded to indicate incompleteness, all respondent-reported data are edited to zero. This inconsistency of instructions across surveys caused confusion among some SSO staff and resulted in the loss of some valid data on the 1989 FCRS.

Although refused items were rare, responses of "don't know" were not uncommon. The importance of missing items that remain blank, regardless of the reason, must be judged by several factors: 1) the use of the data, 2) the likely magnitude of the missing data, 3) the potential impact on the estimates, and 4) the proportion of the data that remains missing. Evidence in Appendix D supports greater concern for the "missingness" among rarer items, those items that receive relatively low positive response. The importance of this situation remains a judgment call based on prioritized use of the data. However, it should be noted that the results reported here are evidence of only the DK's and refused items that were discernable. We cannot know how indicative these results are of item nonresponse that remains unidentified.

The significance of the results on contract expenses lies in a comparison of instructed procedures with documented experience. Pretest experience with contractees in North Carolina during development of the 1989 and other FCRS questionnaires provided evidence that contractees received "settlement sheets" from their contractors containing much of the contract information requested by the FCRS. Thus enumerators were instructed to obtain contract data from the responding farm operators, the contractees, by using the settlement sheets. It is clear from the analysis of contractor expense data that enumerators were unable to obtain the requested data in this way. The results from this study showed that half of the contractor data was imputed due to DK or refused responses from the contractee and that half was provided by a proxy respondent, the contractor.

Frequency analysis provides evidence that, during the edit process, considerable effort is expended moving data around the questionnaire. This movement has little effect on total expense estimates, and may only affect published expense categories if data are moved across sections. It should be noted that even though individual items may lack substance in terms of published estimates, they may be of importance in other uses of the data. In particular, detail on many FCRS items may be needed to support the maintenance of price indices by the Economic Statistics Branch or for econometric models by ERS. However, in light of the degree of "missingness" for some of the rarer items as documented in this report, even in this context quality needs to be balanced against quantity. How good will the inputs to the indices and models be when they are based on data of which half has been manually imputed?

Conclusions and Recommendations

The purpose of this study was to define, document, and understand item nonresponse, imputation, and editing on the FCRS, as well as to evaluate their impact on the estimates. Our mission was to set the foundation for further research into alternative strategies to impute for missing or questionable items. This study pointed out that item imputation is completely reliant upon and not separate from the edit process. Item imputation is not just the replacement of missing items with positive entries, but also includes replacing values for implausibility, incompleteness, or allocation problems.

Many of our findings resemble those of the studies cited in the literature review. While we collected data on all edits, the need for isolating so-called clerical edits became evident during analysis of our data. Clerical editing is necessary to tidy the form of the data. In future studies of editing, clerical edits should be identified separately from other edits.

Analysis of the direct impact of edited, imputed, and missing items indicates that in Iowa and North Carolina a very few reports of total expenditures indicated large enough editing differences to account for most of the EMI's effect on the estimate. In addition, half of the edit changes made to the FCRS questionnaires had essentially no effect on the final results. That is, many small changes were made which had little overall effect on survey estimates.

Nearly half of the edits moved respondent-reported data from one cell to another on the questionnaire. The results of the impact analysis suggest that this type of editing has little or no effect on survey estimates. The same conclusion may be drawn regarding detailed editing for incomplete allocations. While these types of editing provide for internally consistent individual records, they appear to add little to the quality of the major survey estimates.

We have not studied the effect of item nonresponse on the estimates and data analysis published by the Economic Research Service based on FCRS data. ERS should examine the implications of editing on their uses of the data. The best models are of questionable value if they are based on data that can not be reported accurately.

Based upon these observations, we offer the following operational program recommendations:

- The volume of editing being done on the FCRS, as well as the purpose it serves, should be reconsidered. Recognizing an underlying assumption that editing improves data quality, the value added by editing must be considered relative to the final use of the data. If priority is placed on the published estimates, the results of this study indicate that there are limited gains from a large portion of the current editing system.

22

- Both the need and methodology for obtaining an FCRS data set containing internally consistent records should be examined. Extensive clerical and statistician labor and resources are spent doing detailed editing to provide internal consistency. Does the value of an internally consistent data set justify the cost in resources expended on editing? Is there a more cost efficient method to obtain internally consistent and complete records?

- A new editing strategy should be developed relative to prioritized use of the data. For example, if priority is placed on editing to improve the quality of the published estimates, interpretation of our results may suggest the following strategy:

  1) Clerical edits must be performed to clean the data.

  2) Imputation for DK's and refusals is necessary, as it provides the greatest impact on the survey estimates.

  3) Other types of editing especially those which move data around the questionnaire for internal record consistency and completeness should be performed as simply and consistently as possible, probably with automated routines.

- Editing strategy should be administered through clear, consistent editing instructions. These instructions should prioritize the kinds of editing, and indicate responsibility for editing. In other words, what kind of editing should be done by enumerators, supervisory enumerators, statistical assistants, survey statisticians, or the computer?

- An editing manual should be written to standardize editing practices across states as much as possible. This would enable consistency in editing procedures to handle comparable situations, such as allocation problems, and in interpretation of resulting data.

- Automated edit procedures based on statistical tolerance limits or regions should be developed and evaluated, so that those errors with the greatest potential impact on survey estimates may identified and given priority.

- Procedures that minimize the amount of manual editing required to ensure internally consistent records should be developed. These may include improved questionnaire design and enumerator training and the use of computer-assisted personal interviewing to collect data. Automated and manual procedures should be balanced within a pre-defined, standard editing strategy.

- Procedures and notation to identify and capture missing items in the FCRS data set should be developed and incorporated into the survey process. An audit trail of edited data which includes identification of the source of imputations should also be developed. Together these will provide valuable survey management

23

information and indications of data quality to evaluate data usefulness.

- Procedures should be developed to handle incomplete allocations. An automated imputation routine should be developed to complete the allocation of respondent-provided totals into their missing component parts. This procedure could be a simple "hot-deck" proration of reported totals to the detailed item cells for incomplete allocations, based on corresponding cell percentages from usable reported data.

- Editing procedures to handle incomplete records and to identify item nonresponse should be consistent across surveys so that valid data are not lost.

- Data collection procedures for obtaining contractor expense data need additional review. Field practices should be monitored and the source of these data recorded. Survey instructions must be realistic regarding the collection of the data from contractors as proxy respondents.

Additional research will be necessary to support several of the above recommendations. Specifically,

- Research should focus on directing editing strategy relative to 1) easily identified farm characteristics such as size, type, or stratum, and/or 2) magnitude of changes to reported data. Editing attention should be focused on particular records or on changes outside statistically determined tolerance limits or regions.

- Research should be continued on alternative editing and imputation strategies, particularly those that are automated and/or interactive, such as the Blaise system, and multivariate in nature. One possibility would be to examine the GEIS system of Statistics Canada for the FCRS. This system is designed to protect the univariate and multivariate integrity of the data and distributions. Automated editing should ease the burden upon all those involved in the editing process. Automation should not, however, be viewed as the solution to editing problems, but as an alternative resource allocation to fulfill a pre-defined, standard editing strategy. The new State Statistical Office microcomputer local area networks provide a technological opportunity for the development of an interactive editing system, possibly including multivariate relationships, for the FCRS.

References

Allen, J. Donald, "An Overview of Imputation Procedures," SMB Staff Report SMB-90-06, National Agricultural Statistics Service, USDA, August, 1990.

Anderson, Andy B., Alexander Basilevsky, and Derek P. J. Hum, "Missing Data: A Review of the Literature," Chapter 12 in Handbook of Survey Research, edited by Peter H. Rossi, James D. Wright, and Andy B. Anderson, Academic Press, Inc., London, 1983, pp. 415-441.

Bargmann, R., W.W. Donaldson, and Kay Turner, "Multivariate Analysis of Farm Costs and Returns Survey Data," unpublished internal staff report, National Agricultural Statistics Service, USDA, 1991.

Bethlehem, J. G., "The Data Editing Research Project of the Netherlands Central Bureau of Statistics," in the Proceedings of the Third Annual Research Conference of the Bureau of the Census, U.S. Bureau of the Census, Washington, D.C., 1987, pp. 194-203.

Bosecker, Raymond R., "Data Imputation Study on the Oklahoma DES," unpublished internal staff report, Statistical Reporting Service, USDA, October, 1977.

Boucher, Louis, "Micro-Editing for the Annual Survey of Manufactures: What is the Value-Added?" in the Proceedings of the 1991 Research Conference of the Bureau of the Census, U.S. Bureau of the Census, Washington, D.C., 1991, forthcoming.

Fellegi, I. P., and D. Holt, "A Systematic Approach to Automatic Edit and Imputation," Journal of the American Statistical Association, Volume 71, Number 353, 1976, pp. 17-35.

Greenberg, Brian, and Thomas Petkunas, "An Evaluation of Edit and Imputation Procedures Used in the 1982 Economic Censuses in the Business Division," SRD Research Report Number: Census/SRD/RR-86/04, Statistical Research Division Report Series, U.S. Bureau of the Census, 1986.

Hanuschak, George A. et. al., Federal Committee on Statistical Methodology, Data Editing in Federal Statistical Agencies, Statistical Working Paper No. 18, Office of Management and Budget, May, 1990.

Kalton, Graham, and Daniel Kasprzyk, "The Treatment of Missing Survey Data," Survey Methodology, Volume 12, No. 1, Statistics Canada, June, 1986, pp. 1-16.

Lemeshow, Stanley, "Nonresponse (in Sample Surveys)," in Encyclopedia of Statistical Sciences, Volume 6, John Wiley and Sons, Inc., New York, 1983, pp. 333-336.

Linacre, Susan J., and Dennis J. Trewin, "Evaluation of Errors and Appropriate Resource Allocation in Economic Collections," in Proceedings of the Fifth Annual Research Conference of the Bureau of the Census, U.S. Bureau of the Census, Washington, D.C., 1989, pp. 197-209.

Little, Roderick J. A., and Donald B. Rubin, "Incomplete Data," in Encyclopedia of Statistical Sciences, Volume 4, John Wiley and Sons, Inc., New York, 1983, pp. 44-53.

Naus, Joseph I., "Editing Statistical Data," in the Encyclopedia of Statistical Sciences, Volume 2, John Wiley and Sons, Inc., New York, 1982, pp. 455-461.

Pierzchala, Mark, organizer, "Summaries of the Presentations from the NASS Editing Research Forum," unpublished staff notes, National Agricultural Statistics Service, USDA, October, 1988.

Pullman, Thomas W., Trudy Harpham, and Nuri Ozsever, "The Machine-Editing of Large-Sample Surveys: The Experience of the World Fertility Survey," International Statistical Review, Vol. 54, No. 3, 1986, pp. 311-326.

Sande, Innis G., "Imputation in Surveys: Coping with Reality," The American Statistician, Vol. 36, No. 3, Part 1, August, 1982, pp. 145-152.

Definition of Reason Codes

I. Items with Positive Values

   Code 1   Error of misplacement, the original value is removed or corrected.

   Code 2   Implausible, imputed replacement of an illogical or physically impossible value.

   Code 3   Allocation, items are imputed when a respondent provided a total value but did not provide values for related components.

   Code 10   Allocation not done, respondent provided a total but didn't provide breakouts of detail items requested by the questionnaire.

   Code 11   Positive values that were deleted from a section that was coded as a refusal.

   Code 13   Machinery codes corrected in the office.

II. Items with No Entry

   Code 4   Don't Know, positive value imputed because the respondent was unable to provide the requested information.

   Code 5   Error of omission, positive value moved from an error of misplacement.

   Code 6   Refusal, value imputed due to the respondent's refusal to provide information for a specific item.

   Code 7   Identifiable Don't Know, respondent indicated that a positive value existed but was unknown. No value was imputed.

   Code 8   Identifiable refusal, enumerator recorded that the respondent refused this particular item but no value was imputed.

   Code 9   Other, reason for EMI's that did not fit into any their category.

   Code 12   Contractor provided data.

Definition of Selected 1989 FCRS Item Codes

Item
Code   Description

IC17   How many acres did this operation own?
IC19   How much annual cash rent was paid?
IC25   Annual rent received from acres rented to others?
IC27   Total acres operated under this arrangement?
IC35   Acres of all types of corn planted
IC36   Acres of corn for grain harvested
IC37   Total production of corn for grain
IC45   Total production of alfalfa hay
IC46   Amount of alfalfa hay used on this operation
IC49   Total production of other hay
IC50   Amount of other hay used on this operation
IC52   Acres of oats planted
IC68   Total production of soybeans
IC111   Acres of tobacco planted
IC112   Acres of tobacco harvested
IC120   Acres of all other crops planted
IC122   Number of acres used only for pasture
IC123   Acres in Set-aside or Acreages Reduction Program
IC127   Acres in other uses (Farmstead, woodland, etc)
IC128   Number of acres double cropped
IC129   Total number of acres
IC131   Residue from previous year look like this (Y/N)
IC132   Acres the residue looked as great as photos
IC136   Amount spent on seeds, plants, seed clean & treat
IC137   Amount spent for fertilizer, lime, and soil condition
IC138   Amount spent for crop chemicals & pesticides
IC141   Total cost for all fuels & oils
IC145   Farm share expense for repairs & parts for motor vehicles
IC149   Farm share expense for farm supplies & hand tools
IC150   Farm share expense for accessories for motor vehicles
IC151   Farm share expense for farm shop power equipment
IC156   Farm share expense for all other insurance
IC157   Farm share expense for inter & service fees on land,
            buildings, etc
IC158   Farm share expense for inter & service fees on operator loans
IC160   Farm share expense for real estate & property taxes
IC161   Percent of real estate & property tax for real estate only
IC164   1989 depreciation expenses for all capital assets
IC167   Amount spent on transportation items to this operation, etc
IC172   Peak number of workers on payroll on any one day
IC173   Total cash wages paid to all workers-excluding contract labor
IC175   Amount ($) of total cash wages paid to operator
IC176   Percent of total cash wages paid to operator
IC178   Percent of total cash wages paid to other household members
IC179   Amount ($) of total cash wages paid to everyone else
IC180   Percent of total cash wages paid to everyone else
IC189   Average hours per week the operator worked in May
IC190   Average hours per week the operator worked in June
IC191   Average hours per week the operator worked in July

IC192   Average hours per week the operator worked in August
IC215   Peak number of other livestock
IC217   Total expense for all purchased feed items
IC221   Amount spent for veterinary & medical services, etc
IC222   Amount spent for livestock and poultry chemicals
IC256   Quantity sold of market contr livestock, poultry, eggs
IC257   Unit code of marketing contr livestock, poultry, eggs
IC276   Quantity removed from production contr livestock, poultry
IC277   Unit code of production contr livestock, poultry
IC278   Price per unit of produce contr livestock, poultry
IC279   Total $ for production contr livestock, poultry
IC303   Weight per unit (lbs) market contr fruit, vegetables, other
         crops
IC309   Weight per unit (lbs) market contr fruit, vegetables, other
         crops
IC356   Total agricultural expense not recorded (excluding mark &
         storage charges, etc)
IC425   Landlord expense for other crop & livestock insurance
IC429   Landlord expenses for all other insurance
IC431   Landlord expenses for real estate & property taxes
IC436   Contractor expenses for hauling items to/from operation
IC450   Contractor expenses for comp rations & formula feed(s)
IC466   Contractor expenses for veterinary services & supplies
IC468   Contractor expenses for livestock, dairy, poultry chemicals
IC486   Contractor expenses for purchase of broilers, fryers, etc
IC539   Amount received for corn, barley, oats, sorghum (milo)
IC543   Amount $ received for all tobacco
IC552   Total crop marketing charges, store, check-offs, etc
IC556   Total dairy marketing charges, store, check-offs, etc
IC568   Total livestock marketing charges, etc (excluding dairy)
IC572   Total amount of cash pay received from State or FED program
IC573   Total face value of certificates received as payment
IC578   Amount received for PIK certificates sold in 1989
IC596   Amount of other farm related income
IC598   Cash income received code for off-farm wages/salaries
IC599   Cash income received code for interest & dividends
IC600   Cash income received code for other off-farm sources
IC605   Acres owned by this operation on 12/31/89
IC606   Market value of land per acre owned on 12/31/89
IC607   Total value of land only
IC608   Market value of operator's dwelling on 12/31/89
IC609   Operator's dwelling in town/city or suburban lot?
IC613   Total market value of buildings owned
IC614   Total market value of land & buildings owned
IC617   Value of all livestock & poultry on hand 1/1/89
IC618   Value of all livestock & poultry on hand 12/31/89
IC619   Value of breeding stock all livestock & poultry on 12/31/89
IC620   Value of crop storage on & off this operation 1/1/89
IC621   Value of crop storage on & off this operation 12/31/89
IC623   Value of all feed, fertilizer, chemicals, etc 1/1/89
IC625   Value of stocks in lending institution on 12/31/89
IC626   Code for total value of all other assets
IC629   Average interest rate on balance owed Prod Credit Association
IC630   Balance owed to Federal Land Banks on 12/31/89

IC631	Average interest rate on balance owed to Federal Land Banks
IC634	Balance owed to banks and savings & loans 12/31/89
IC635	Average interest rate on balance owed to banks and S&L
IC643	Average interest rate on balance owed to individuals
IC650	Percent of total debt outstanding had term 1 to 10 years
IC652	Code that represents total gross value of sales
IC654	Category that represents largest part of gross income
IC656	Code for type of farm ownership
IC662	Percent of net income operation & household receives
IC701	Acres - planned crop rotate strictly follow
IC703	Miles of cropland that borders the stream or river
IC704	Amount($) of seeds/plants spent on field crops or s-grn
IC705	Percent of seeds/plants spent for field crops or s-grn
IC707	Percent of seeds/plants spent for lentils, drybeans, etc
IC709	Percent of seeds/plants spent on other crops
IC718	Acres cust work harvesting of hay, straw, etc
IC719	Cost of cust work harvesting of hay, straw, etc
IC727	Cost of cust work harvesting of cotton
IC733	Average cost per gallon for diesel fuel
IC735	Percent of total fuel cost for bulk delivered gas
IC736	Average cost per gallon for bulk delivered gasoline
IC739	Average cost per gallon for purchased gasoline
IC740	Amount ($) of total fuel cost spent on LP gas (propane)
IC748	Farm share expense for electricity for home & farm
IC750	Farm share expense for telephone charges
IC751	Farm share expense for Federal crop insurance
IC752	Farm share expense for other crop & livestock insurance
IC757	Expense for employee workers compensation, Social Security, unemployment taxes
IC758	Social Security self-employment-taxes for operator & all partners
IC762	Percent of feed purchased; barley, corn, oats, wheat, etc
IC764	Percent of feed purchased hays and forages
IC765	Amount of feed purchased; compl rations & formula feeds
IC766	Percent of feed purchased; complt rations & formula feeds
IC767	Amount of feed purchased; protein meals & concentrates
IC768	Percent of feed purchased; protein meals & concentrates
IC769	Amount of feed purchased; supplements
IC770	Percent of feed purchased; supplements
IC771	Amount of feed purchased; all other ingredients
IC772	Percent of feed purchased; all other ingredients
IC779	Model year of 1st tractor purchased
IC790	Cost for 3rd tractor purchased
IC816	Size of 1st type of trucks, etc purchased in 89
IC819	Size of 1st type of car(s) purchased in 89
IC822	Model year of 1st type of truck, etc purchased in 89
IC825	Model year of 1st type of car(s) purchased
IC831	1st type of car purchased new or used
IC837	Net cost of car(s) purchased
IC852	Number of car(s) purchased
IC892	Total amount received from State or FED program (cert & cash)
IC893	Amount State/Federal for set-aside or Acreage Reduction Program
IC896	Amount State/Federal farm program for any other programs

EMI Effects on Edit and Estimate

The following graphs display the EMI effect on the <u>edit</u> and the <u>estimate</u> for selected aggregate items in Iowa and North Carolina. Each of these effects is a measurement of the summarized EMI differences and are described in detail in the Methodology Section, pages 7-9, of this paper.

Both measurements use the expanded EMI differences from each changed record. These individual record expansions are then ordered from largest to smallest based on their absolute value. Absolute values are used for ordering only. EMI effect on the edit is then calculated by using the cumulative sum of the expanded differences at various points and then dividing this sum by the summed total of expanded differences for all changes made during the edit. EMI effect on the estimate is determined by dividing the state direct expansion less the cumulative sum of the expanded differences by the state level estimate. Each of these measurements relates the contribution of individual editing changes to the overall effect of all changes.

For example, let's consider the graph of the EMI effect on Iowa's total expenditures on the next page. We can see from the line labeled "Estimate" that if no edit changes had been made the state level estimate for total expenditures would have been within one percent of the final expansion. We can also see by examining the line labeled "Edit" that ninety percent of the effect caused by edit changes is attributed to approximately thirty percent of the questionnaires that were changed.

In general, these types of graphs are useful to clearly demonstrate that there is a point of diminishing return where EMI changes have no effect. By using this type of graphic representation we are also able to isolate and categorize certain types editing problems, like the contractor refusals in North Carolina.

31

EMI Effect: Iowa
Total Expenses

Percent of Edit

EMI Effect: Iowa
Marketing and Storage Expenses

Percent of Edit

EMI Effect: North Carolina
Total Expenses

Estimate

Edit

Percent of Effect

Percent of Edit

EMI Effect: North Carolina
Marketing and Storage Expenses

Estimate

Edit

Percent of Effect

Percent of Edit
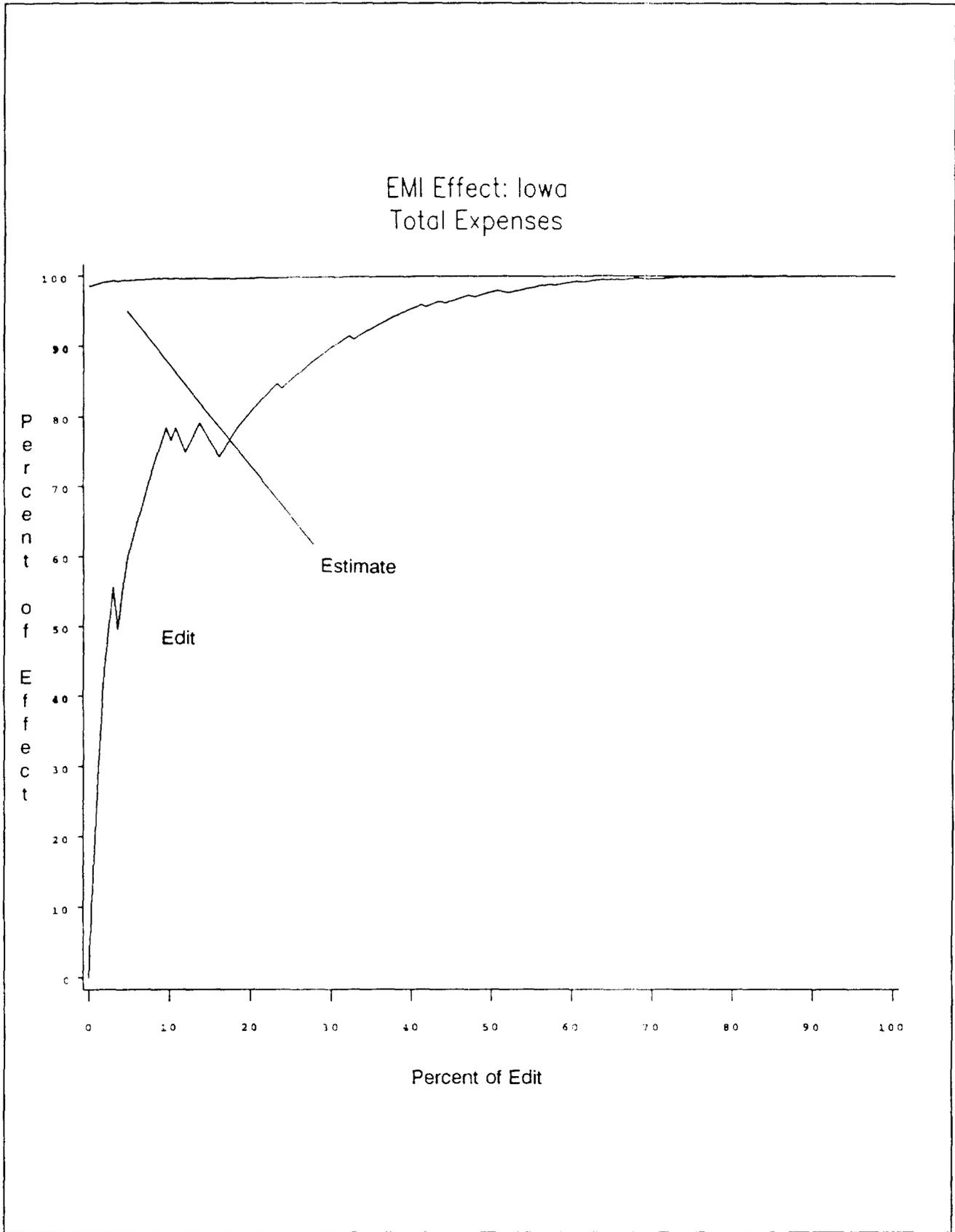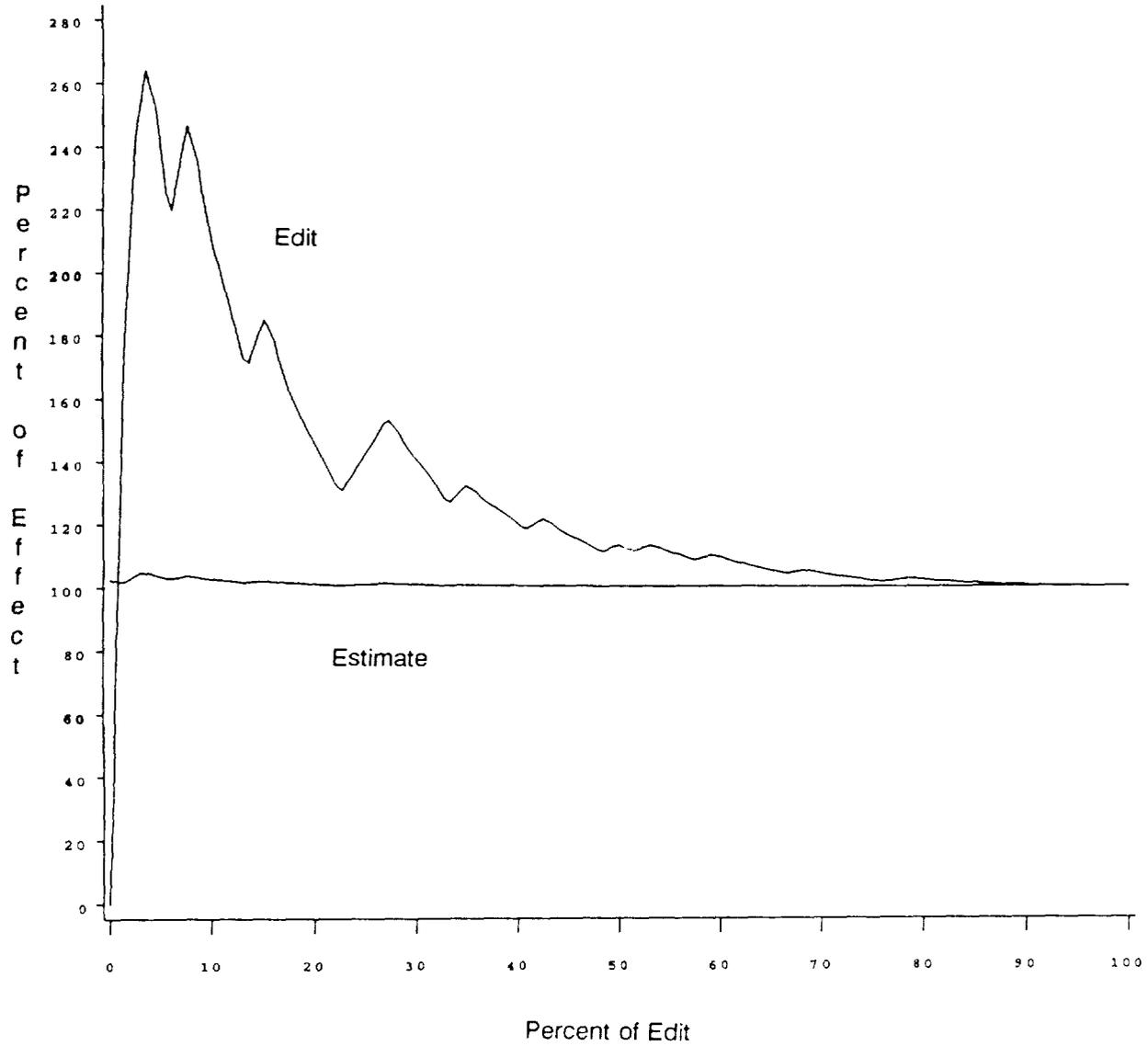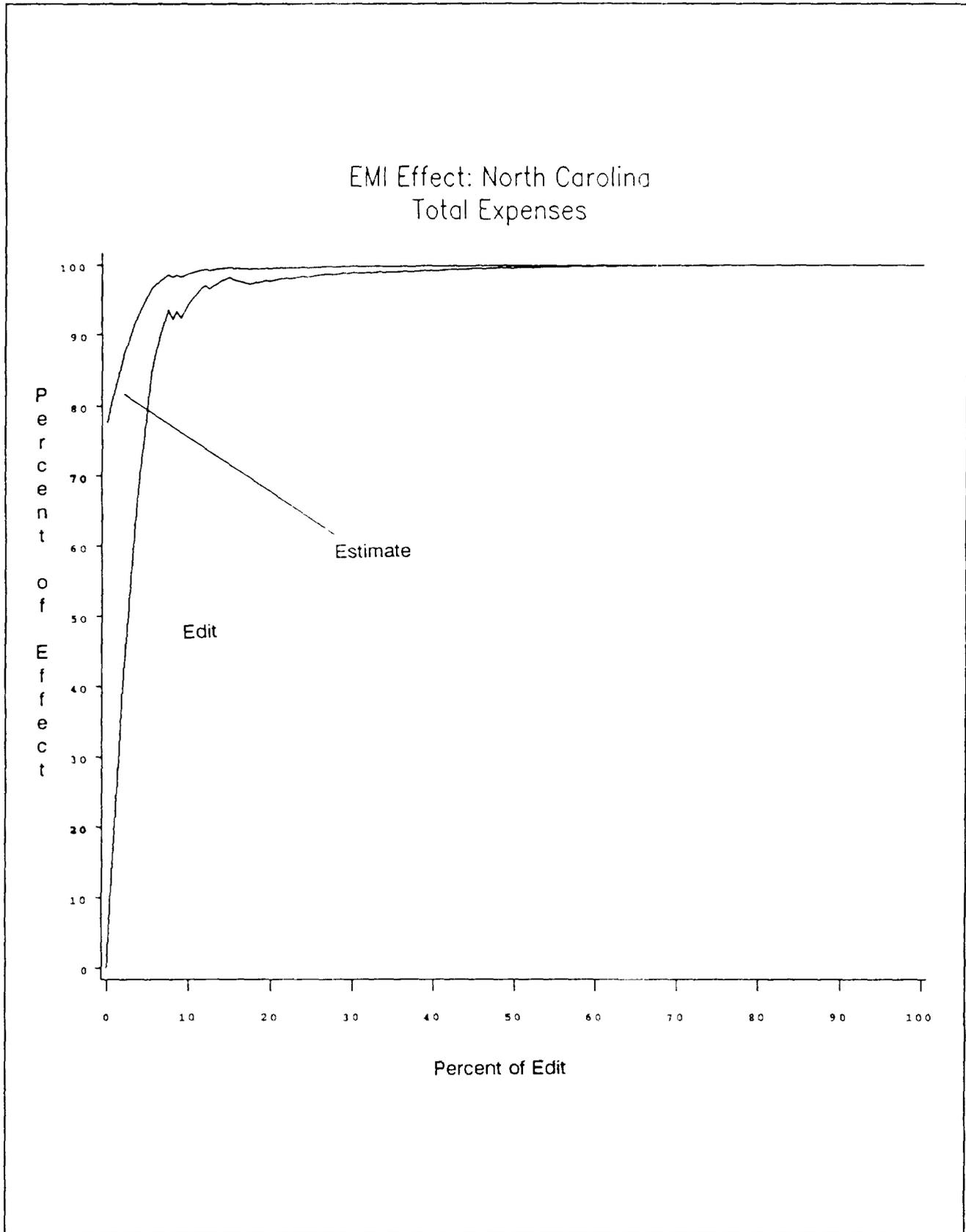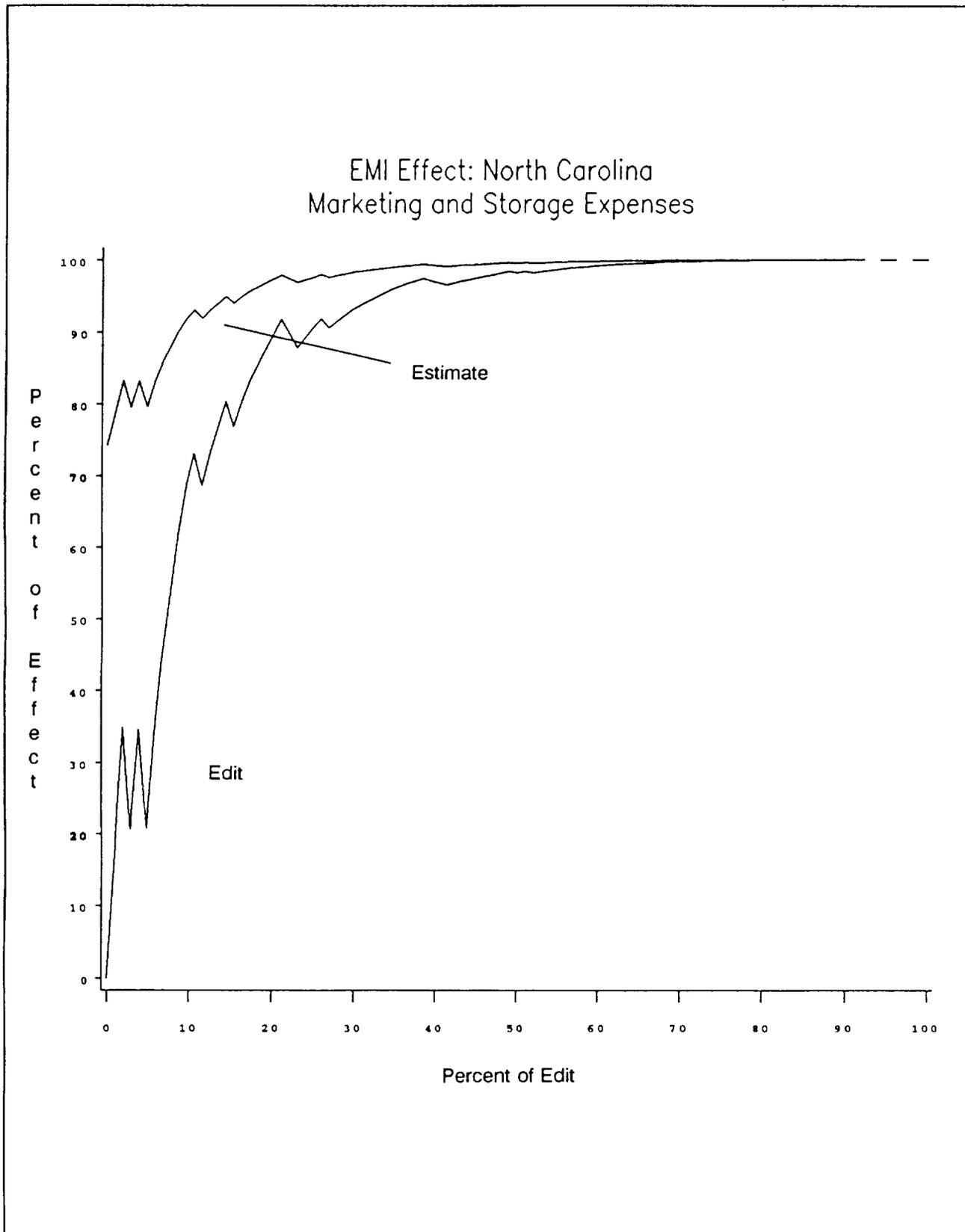
Detailed Frequency Analysis

General Results

Table D1 displays the frequencies of occurrence of the EMI data. The research team reviewed 448 completed, usable questionnaires in Iowa, identifying 3320 EMI items. Of these, 3292 came from 403 questionnaires for operations qualifying as farms. In North Carolina, 321 questionnaires were reviewed with 3428 EMI's observed. Of these, 3330 EMI's came from 304 questionnaires for operations qualifying as farms. All analyses following Table D1 consider only EMI's from farm operations.

TABLE D1. Distributional Characteristics of the Data for the Item Nonresponse Research Project.

	IA	NC
Number of completed, usable questionnaires that were reviewed	448	321
Number of questionnaires for operations qualifying as farms	403	304
Number of questionnaires for operations qualifying as farms that required no editing, no EMI's	35	12
Observed number of EMI's	3320	3428
Total number of EMI's for operations qualifying as farms	3292	3380
Number of clerical EMI's	353	117
Number of EMI's representing contractor-provided data	–	95
Number of items exhibiting non-clerical EMI's	410	451
Average number of EMI's per edited Expenditure Version 1/	7.7	10.1
Median number of EMI's per edited Expenditure Version 1/	6.0	8.0
Mode number of EMI's per edited Expenditure Version 1/	4.0	5.0

1/ After removal of clerical edits and contractor-provided data.

In Iowa, 353 EMI's, or nearly 11 percent of the edits could be considered clerical, that is, they merely corrected the recording of data for misplaced decimal points or they rounded responses in dollars and cents to the nearest whole dollar. These 353 clerical edits came from 33 questionnaires. In fact, 5 questionnaires contained only clerical edits. In North Carolina, 117 EMI's, or only 3.5 percent of the edits, were clerical. These edits were found on 63 questionnaires. All clerical edits were dropped from further analysis. Items for which half or more of the EMI's were clerical are noted in Appendix F.

In North Carolina, 95 EMI's represented data provided by contractors rather than by the respondent/contractee, of whom the data were actually requested. That is, these data were provided by a proxy respondent. Although they are considered valid, these data were originally recorded by the research review team as EMI with reason code 12, since edit action by the SSO staff was required to place the data in the appropriate cells. A more detailed examination of contractor-provided data will follow. Meanwhile, reason code 12 data have been removed from further analysis. Items for which half or more of the EMI's was due to reason code 12 are appropriately noted in Appendix F. EMI contractor expense items that were due to "don't knows" or refusals remain in the analysis data set.

In Iowa, 410 items exhibited EMI's. The three most heavily edited items accounted for 10 percent of the EMI's; 48 items contained 50 percent of the edits. These items exhibited 14 or more EMI's each. A little over half of the edited items accounted for 90 percent of the EMI's. On the other hand, 109 items were each edited only once, while 82 items each contained 10 or more EMI's. Only 36 items exhibited 20 or more EMI's.

In North Carolina, 451 items were edited or contained missing or imputed data. Like Iowa, three items accounted for 10 percent of the EMI's and 55 items contained 50 percent of the edits, exhibiting 14 or more EMI's each. Ninety percent of the EMI's were found in 255 items, just over half of those edited. Also similar to Iowa, 94 items exhibited only one EMI each. Ten or more EMI's were recorded for 89 items and, again like Iowa, only 34 items exhibited 20 or more EMI's.

While all versions of the FCRS questionnaires were reviewed, only the expenditure version had all non-administrative items observed for edited, missing, or imputed data. Therefore, means, medians, and modes are reported for the expenditure version only. In Iowa the average number of EMI's per edited expenditure version was 7.7, the median was 6.0, and the mode was 4.0. In North Carolina, the corresponding figures are 10.1 for the average, 8.0 for the median, and 5.0 for the mode. These results indicate the variability in the amount of editing per questionnaire within each state and between states.

Allocation Problems (Reason Codes 3 and 10)

Eleven percent of the EMI's in Iowa and less than 4 percent of the EMI's

in North Carolina reflect allocation problems. That is, the respondent indicated the value provided in one cell included expenses requested in other cells, but did not allocate this total into the requested parts. Like DK's, concern for allocation edits arises because of the need for the detail item data to support price index statistics. In Iowa, there were 328 EMI items identified as allocation problems; in North Carolina, the corresponding figure was 113. Data items coded as EMI allocation edits consisted of the following: 1) the item in which the respondent placed an unallocated total and 2) the detail items that the respondent indicated to be included in that total. The items of both types were recorded with the same reason code.

A detailed analysis of all allocation EMI's appears in Table D2. The table includes a listing of the item groups where allocation EMI's occurred most frequently, documenting differences between the two states in the groups of items that exhibited allocation EMI's.

Table D3 displays selected items that exhibited incomplete allocations (reason code 10) in Iowa. When the allocation of a total into the requested parts is not provided, the data offer some insight into which of the detailed items tend to receive more than their proper allocation and which tend to receive less. Thus, analysis of these data provides a sense of which individual items may be overstated, those items receiving the total, and which items may be understated, those items remaining blank.

For instance, consider the statistics for farm and motor supply items (item codes IC145, IC149, IC150, and IC151) presented in Table D3. Expenses for motor vehicle repairs and parts (IC145) were recorded as EMI reason code 10 a total of 19 times. Nearly all of those times, it contained a positive value. That is, the respondent indicated that the value recorded in IC145 included data requested elsewhere. Therefore, if unchanged, IC145 will likely be overstated.

On the other hand, expenses for accessories for machinery (IC150) were recorded as EMI reason code 10 only six times, all of which remained blank in the final data set. That is, the respondent indicated that the data requested in IC150 had been included with related data in another cell. IC150 is a rare item with only 27 positive values in the final data set. Thus concern may be justified for the understatement of IC150, since there could be at least 20 percent more positive responses than there are. Similar concern may be expressed for livestock and dairy chemicals (IC222), which could have a third more positive responses than were found in the final data set.

TABLE D2.   Detailed Analysis of Allocation EMI's (Reason Codes 3 and 10).

	IA	NC 1/
Number of Allocation EMI's	328	113
% of all EMI's (after removal of clerical edits and contractor-provided data 2/)	11.1	3.6
% of Allocation EMI's remaining unallocated in the edited data set	71.7	-

% of Allocation EMI's in the following item code groups:

	IA	NC
Fertilizer, crop chemical and pesticide expenses (IC137 or IC138)	10.4	17.6
Repairs, parts, tools, accessories, shop equipment, etc. (IC145, IC149, IC150, or IC151)	16.2	8.9
Vet expense and livestock chemicals (IC221 or IC222)	16.5	-
Electricity and telephone expenses (IC748 or IC750)	16.8	-
Wages paid to self, family & others (IC175, IC176, IC178, IC179 or IC180)	-	14.2
Seed expenses (IC136, IC704, IC705, IC707 or IC709)	-	11.6
Interest expenses, real estate vs. operating loans (IC157 or IC158)	-	10.6

1/   There were no items with reason code 10 in NC.
2/   There were no contractor-provided data in IA.

TABLE D3. Detailed Analysis of EMI's Reflecting Incomplete Allocations of Expenses (Reason Code 10) in Iowa - Selected Items.

Item (by Allocation Group)	Incomplete Allocations (Code 10)			Total Number of Positives in Edited Data Set
	Total Number	% Blank	% Positive	
Fertilizer expense (IC137)	9	0.0	100.0	378
Crop chemical & pesticide expense (IC138)	13	100.0	0.0	361
Motor vehicle repairs & parts (IC145)	19	5.3	94.7	413
Farm supplies, tools, etc. (IC149)	18	61.1	38.9	351
Accessories for machinery (IC150)	6	100.0	0.0	27
Farm shop power equipment (IC151)	7	100.0	0.0	89
Veterinary expense (IC221)	23	0.0	100.0	319
Livestock & dairy chemicals (IC222)	29	100.0	0.0	86
Electricity for farm and home (IC748)	21	0.0	100.0	309
Telephone charges (IC750)	20	100.0	0.0	231

Don't Knows (Reason Codes 4, 7, and 9)

Detailed analyses of the DK EMI's appear in Tables D4 and D5. In Table D4 we can see that the two states are surprisingly similar in the sections in which DK EMI's most commonly occur. In both states approximately one-third of the DK EMI's occur in the section requesting landlord and contractor expenses, and nearly 11 percent of the DK EMI's occur in the table of beginning and ending inventories in the Assets Section. DK EMI's are also common in the request for price per gallon of fuel in both states.

TABLE D4. Detailed Analysis of "Don't Know" (DK) EMI's (Reason Codes 4, 7, and 9).

	IA	NC 1/
Number of DK EMI's	359	483
% of all EMI's (after removal of clerical edits and contractor-provided data 2/)	12.1	15.3
% of DK EMI's remaining blank in the edited data set	44.0	32.5
% of DK EMI's in:		
Landlord & contractor expenses	34.5	31.7
Price/gal for fuels	18.4	8.7
Beginning & ending inventories	10.9	10.6

1/ There were no items with reason code 9 in NC.
2/ There were no contractor-provided data in IA.

Table D5 provides detail on DK EMI's for selected items in the major groups highlighted in Table D4. These items exhibited the most frequent DK EMI's within their respective groups. Table D5 indicates that 12 to 17 percent of the positive values in the edited data set for price per gallon of fuels (IC733 and IC736) and landlord real estate taxes (IC431) in Iowa were imputed. In North Carolina, nearly half of the positive values for landlord real estate taxes (IC431) were imputed and more than one-third of the positive values for contractor veterinarian expenses (IC466) were imputed. The imputation rate for price per gallon of the various fuels (IC736, IC733, IC739) ranged from 6 percent to 18 percent in North Carolina.

Finally, we should also be concerned about DK's that remain blank. In Iowa, for example, there quite possibly should be half again as many positive values for landlord expense for other insurance (IC429) as there actually are in the edited data set. DK EMI's for the inventory items (IC623 and IC625) tend to remain blank in both states as well.

41

TABLE D5. Detailed Analysis of "Don't Know" EMI's - Selected Items.

Number of DK EMI's

Item	Total	Imputed	Remaining Blank	Total Number of Positives in Edited Data Set
Iowa:				
Landlord real estate taxes (IC431)	55	44	11	254
Landlord, other insurance (IC429)	28	3	25	53
Price/gal, diesel fuel (IC733)	31	27	4	227
Price/gal, bulk gasoline (IC736)	31	28	3	222
Beginning inventory of feed, seed, supplies, etc. (IC623)	12	2	10	234
North Carolina:				
Landlord real estate taxes (IC431)	91	89	2	191
Contractor vet services (IC466)	11	10	1	28
Price/gal, gas purchased off operation (IC739)	18	17	1	97
Price/gal, diesel fuel (IC733)	15	13	2	115
Price/gal, bulk gasoline (IC736)	9	7	2	109
Ending inventory, stock in FLB's, PCA's, or coop's (IC625)	9	1	8	58

Item Refusals (Reason Codes 6 and 8)

Items that were indicated as having been refused by the respondent were rare, accounting for only 1.4 percent of the EMI's in Iowa and 2.2 percent of the EMI's in North Carolina. In Iowa, there were only 42 identifiable refused items, 29 of which received imputation, 13 remained blank. These refused items appeared on only 11 questionnaires and tended to be grouped together. Twenty of the imputed, refused items were aggregate expense items on a single dairy version questionnaire.

In North Carolina, there were 67 items identified as having been refused by the respondent or contractor proxy respondent. They occurred on 26 questionnaires. Fifty-nine EMI refusals received positive imputation, while only 8 remained blank in the edited data set. Unlike the EMI refusals in Iowa, there was a pattern to the refused items in North Carolina, which centered around contractor expense data. Half of the refused items occurred on 7 questionnaires and were solely contractor expense items. Another 15 percent of the refused items were found in IC626, "all other farm assets." The remaining refused items tended to be grouped in pairs or in sections on individual questionnaires.

R-box Miscoding (Reason Code 11)

A closer look is taken at the data lost due to reason code 11 in North Carolina because it occurred more frequently there than it did in Iowa. The 1989 FCRS survey follow-up suggested that other states exhibited confusion with the R-box instructions as well.

There were 126 EMI's with reason code 11 in North Carolina, 37 on 10 expenditure version questionnaires and 89 on 11 FOR version questionnaires. On the expenditure version, reason code 11 occurred in 4 sections: cash sales of crops, off-farm income, the assets page, and the loan balances items of the Assets and Liabilities Section. Most of the expenditure version reason code 11 EMI's occurred on the assets page.

Nearly all of the reason code 11 EMI's on the FOR version occurred in the "Loans, Interest Rates, and Terms" portion of the Assets and Liabilities Section. This portion consisted of a table with columns requesting each of the following: the balance of the loan, the interest rate, the term of the original loan, and the scheduled principal paid during 1989 for loans from a variety of sources. It is clear from the pattern of EMI's in this section that respondents had difficulty providing data on the scheduled principal paid in 1989 for their loans, because all of the data that were edited out occurred in the first three columns of this table. While items in the "Balance" and "Interest Rate" columns contained 91 positive responses in the edited data set, an additional 58 positive responses had been edited to zero when the R-box was coded. Likewise, items in the "Term of Original Loan" column contained 62 positive responses in the edited data set, while an

additional 25 positive responses were deleted in edit. Thus there was nearly half again as much data reported as was actually made available to the data user for this section.

Contractor Expense Data (Reason Code 12)

Contractor-provided data appeared on 22 questionnaires and for 23 items. Four of the items, IC276-IC279, had to do with the quantity of livestock raised under contract that were removed from the operation. Sixteen of the items were in the contractor expense section.

The final edited data set contained 175 positive responses to a total of 20 items in the contractor expense section. These 20 items exhibited 176 EMI's. (The number of EMI's may be greater that the number of positive responses in the final edited data set, because EMI's include items edited to zero and blanks remaining blank.) Only five of the 176 EMI's for contractor expenses remained blank in the edited data set. Of the EMI's with positive values, 46.7 percent represent data provided by contractors that were edited in, while 51.5 percent were DK or refused items for which the SSO statisticians imputed positive values using well-documented procedures.

A detailed analysis of response to selected contractor expense items in North Carolina appears in Table D6. The five items listed exhibited the greatest frequency of EMI's among all contractor expense items. For each of these items, nearly all of the positive data in the edited data set were provided by a proxy respondent, that is, the contractor, or were imputed for DK or refused EMI items.

TABLE D6. Detailed Analysis of Selected Contractor Expense Items in North Carolina.

| Item | Total Number of Positives in Edited Data Set | Number of EMI's | | | |
		Total	Provided By Contractor	Imputed For DK	Imputed For Refusal
Hauling (IC436)	22	21	8	9	4
Feed (IC450)	21	20	13	4	2
Vet (IC466)	28	29	10	10	7
Livestock chemicals (IC468)	16	15	5	5	5
Poultry purchases (IC486)	22	21	12	6	3

44

Edit Actions by Reason for EMI Items
Iowa (19)

		Description of Edit Action					
		Decreased		Unchanged		Increased	
		Reported value>0 edited value>0	Reported value>0 edited value=0	Reported value=0 edited value=0	Reported & edited values = and > 0	Reported value=0 edited value>0	Reported value>0 edited value>0
Reason for EMI							
Error of Misplac- ement Code 1	Frequency	81	328	2	32	6	102
	% by Reason	19.33	78.10	0.62	20.38	0.50	24.82
	% by Edit Action	14.70	59.53	0.36	5.81	1.09	18.51
Implaus- ible Code 2	Frequency	318	67	3	19	2	298
	% by Reason	75.89	15.95	0.93	12.10	0.17	72.51
	% by Edit Action	44.98	9.48	0.42	2.69	0.28	42.15
Allocat- ion Edit Code 3	Frequency	16	.	10	2	53	10
	% by Reason	3.82	.	3.11	1.27	4.38	2.43
	% by Edit Action	17.58	.	10.99	2.20	58.24	10.99
Don't Know Positive Edit Code 4	Frequency	.	.	.	.	201	.
	% by Reason	.	.	.	.	16.61	.
	% by Edit Action	.	.	.	.	100.00	.
Error of Omission Code 5	Frequency	.	.	19	.	905	.
	% by Reason	.	.	5.90	.	74.79	.

Edit Actions by Reason for EMI Items
Iowa (19)

		Description of Edit Action					
		Decreased		Unchanged		Increased	
		Reported value>0 edited value>0	Reported value>0 edited value=0	Reported value=0 edited value=0	Reported & edited values = and > 0	Reported value=0 edited value>0	Reported value>0 edited value>0
Reason for EMI							
Error of Omission	% by Edit Action	.	.	2.06	.	97.94	.
Refusal Positive Edit	Frequency	.	.	.	.	29	.
	% by Reason	.	.	.	.	2.40	.
Code 6	% by Edit Action	.	.	.	.	100.00	.
Don't Know No Edit	Frequency	.	.	136	.	.	.
	% by Reason	.	.	42.24	.	.	.
Code 7	% by Edit Action	.	.	100.00	.	.	.
Refusal No Edit	Frequency	.	.	13	.	.	.
Code 8	% by Reason	.	.	4.04	.	.	.
	% by Edit Action	.	.	100.00	.	.	.
Other	Frequency	.	.	18	.	4	.
Code 9	% by Reason	.	.	5.59	.	0.33	.
	% by Edit Action	.	.	81.82	.	18.18	.

Edit Actions by Reason for EMI Items
Iowa (19)

		Description of Edit Action					
		Decreased		Unchanged		Increased	
		Reported value>0 edited value>0	Reported value>0 edited value=0	Reported value=0 edited value=0	Reported & edited values = and > 0	Reported value=0 edited value>0	Reported value>0 edited value>0
Reason for EMI							
Allocation No Edit Code 10	Frequency	4	3	121	104	2	1
	% by Reason	0.95	0.71	37.58	66.24	0.17	0.24
	% by Edit Action	1.70	1.28	51.49	44.26	0.85	0.43
Value Edited Out Code 11	Frequency	.	22	.	.	.	.
	% by Reason	.	5.24	.	.	.	.
	% by Edit Action	.	100.00	.	.	.	.
Machinery Codes Code 13	Frequency	.	.	.	.	8	.
	% by Reason	.	.	.	.	0.66	.
	% by Edit Action	.	.	.	.	100.00	.
Total	Frequency	419	420	322	157	1210	411
	% by Reason	100.00	100.00	100.00	100.00	100.00	100.00
	% by Edit Action	14.26	14.29	10.96	5.34	41.17	13.98

47

Edit Actions by Reason for EMI Items
North Carolina (37)

		Decreased		Unchanged		Increased	
		Reported value>0 edited value>0	Reported value>0 edited value=0	Reported value=0 edited value=0	Reported & edited values = and > 0	Reported value=0 edited value>0	Reported value>0 edited value>0
Reason for EMI							
Error of Misplac- ement Code 1	Frequency	129	478	9	45	11	117
	% by Reason	28.86	67.71	4.52	69.23	0.81	29.55
	% by Edit Action	16.35	60.58	1.14	5.70	1.39	14.83
Implaus- ible Code 2	Frequency	282	101	1	17	9	275
	% by Reason	63.09	14.31	0.50	26.15	0.66	69.44
	% by Edit Action	41.17	14.74	0.15	2.48	1.31	40.15
Allocat- ion Edit Code 3	Frequency	36	1	8	3	61	4
	% by Reason	8.05	0.14	4.02	4.62	4.51	1.01
	% by Edit Action	31.86	0.88	7.08	2.65	53.98	3.54
Don't Know Positive Edit Code 4	Frequency	.	.	.	.	326	.
	% by Reason	.	.	.	.	24.08	.
	% by Edit Action	.	.	.	.	100.00	.
Error of Omission Code 5	Frequency	.	.	16	.	884	.
	% by Reason	.	.	8.04	.	65.29	.

48

Edit Actions by Reason for EMI Items
North Carolina (37)

| | | Description of Edit Action | | | | | |
| | | Decreased | | Unchanged | | Increased | |
		Reported value>0 edited value>0	Reported value>0 edited value=0	Reported value=0 edited value=0	Reported & edited values = and > 0	Reported value=0 edited value>0	Reported value>0 edited value>0
Reason for EMI							
Error of Omission	% by Edit Action	.	.	1.78	.	98.22	.
Refusal Positive Edit	Frequency	.	.	.	.	59	.
	% by Reason	.	.	.	.	4.36	.
Code 6	% by Edit Action	.	.	.	.	100.00	.
Don't Know Edit	Frequency	.	.	157	.	.	.
	% by Reason	.	.	78.89	.	.	.
Code 7	% by Edit Action	.	.	100.00	.	.	.
Refusal No Edit	Frequency	.	.	8	.	.	.
Code 8	% by Reason	.	.	4.02	.	.	.
	% by Edit Action	.	.	100.00	.	.	.
Value Edited Out	Frequency	.	126	.	.	.	.
	% by Reason	.	17.85	.	.	.	.
Code 11	% by Edit Action	.	100.00	.	.	.	.
Machine-ry Codes	Frequency	.	.	.	.	4	.

Edit Actions by Reason for EMI Items
North Carolina (37)

		Description of Edit Action					
		Decreased		Unchanged		Increased	
		Reported value>0 edited value>0	Reported value>0 edited value=0	Reported value=0 edited value=0	Reported & edited values = and > 0	Reported value=0 edited value>0	Reported value>0 edited value>0
Reason for EMI							
Machine-ry Codes	% by Reason	.	.	.	.	0.30	.
Code 13	% by Edit Action	.	.	.	.	100.00	.
Total	Frequency	447	706	199	65	1354	396
	% by Reason	100.00	100.00	100.00	100.00	100.00	100.00
	% by Edit Action	14.11	22.29	6.28	2.05	42.75	12.50

Count of EMI's for Selected Item Codes
Iowa (19)

Item Code	EMI Count	Positive Count 1/	Zero EMI Count 2/	Total Count 3/	Percent of Total	Reason Code 4/
IC129	120	443	0	443	27.09	2
IC669	89	433	0	433	20.55	5
IC614	71	379	0	379	18.73	2
IC552	68	310	1	311	21.86	5
IC431	59	254	1	255	23.14	4
IC609	56	317	23	340	16.47	5
IC568	49	295	0	295	16.61	5
IC652	46	443	1	444	10.36	2
IC607	43	360	0	360	11.94	2
IC635	38	232	5	237	16.03	2
IC132	37	127	0	127	29.13	5
IC127	36	431	1	432	8.33	5,2
IC892	36	261	1	262	13.74	5
IC356	35	57	9	66	53.03	1
IC750	34	231	1	232	14.66	10
IC128	33	69	5	74	44.59	5
IC222	33	86	0	86	38.37	10
IC613	33	363	0	363	9.09	5
IC736	33	222	0	222	14.86	4
IC161	32	324	2	326	9.82	5
IC221	32	319	1	320	10.00	10
IC145	31	413	0	413	7.51	10
IC733	31	227	0	227	13.66	4
IC599	30	268	1	269	11.15	2
IC606	30	360	0	360	8.33	2
IC748	30	308	1	309	9.71	10
IC429	28	53	0	53	52.83	7
IC893	28	245	0	245	11.43	2,5
IC138	26	361	0	361	7.20	10
IC137	25	378	0	378	6.61	10
IC149	23	351	0	351	6.55	10
IC303	22	37	0	37	59.46	5
IC572	22	365	1	366	6.01	5,1
IC141	21	425	0	425	4.94	2
IC173	21	293	1	294	7.14	5
IC654	21	443	0	443	4.74	5
IC703	21	123	0	123	17.07	*

1/ Positive count is the number of questionnaires with summarized positive values.
2/ Zero count is the number of questionnaires where an item was edited to zero.
3/ Total Count = Positive Count + Zero Count.
4/ Code(s) of the reason(s) for at least half of the EMI's.
* At least half of the EMI's were clerical in nature.

Count of EMI's for Selected Item Codes
Iowa (19)

Item Code	EMI Count	Positive Count 1/	Zero EMI Count 2/	Total Count 3/	Percent of Total	Reason Code 4/
IC718	21	94	4	98	21.43	5
IC825	21	27	12	39	53.85	1
IC131	20	131	4	135	14.81	5
IC217	20	338	0	338	5.92	10,2
IC643	18	98	2	100	18.00	2
IC45	17	232	0	232	7.33	4
IC123	17	296	1	297	5.72	*
IC158	17	248	1	249	6.83	10,3
IC46	16	195	0	195	8.21	5,2
IC626	16	380	2	382	4.19	2
IC35	15	340	0	340	4.41	*
IC52	15	246	0	246	6.10	5
IC157	15	232	1	233	6.44	10,5
IC167	15	198	1	199	7.54	5
IC578	15	265	0	265	5.66	2,5
IC617	15	316	1	317	4.73	5,4
IC771	15	142	4	146	10.27	5
IC120	14	27	7	34	41.18	1
IC573	14	275	0	275	5.09	3,4,2
IC770	14	114	1	115	12.17	5
IC779	14	66	2	68	20.59	2
IC539	13	277	1	278	4.68	1,5,2
IC600	13	99	0	99	13.13	2
IC662	13	44	12	56	23.21	1
IC701	13	161	0	161	8.07	*
IC764	13	120	0	120	10.83	5
IC813	13	27	11	38	34.21	1
IC819	13	27	9	36	36.11	1
IC822	13	53	3	56	23.21	2
IC37	12	369	0	369	3.25	5
IC598	12	207	1	208	5.77	2
IC623	12	234	0	234	5.13	9
IC719	12	94	8	102	11.76	2,1
IC765	12	119	2	121	9.92	5
IC768	12	173	1	174	6.90	5
IC769	12	102	1	103	11.65	5
IC772	12	153	2	155	7.74	5
IC816	12	53	3	56	21.43	5

1/ Positive count is the number of questionnaires with summarized positive values.
2/ Zero count is the number of questionnaires where an item was edited to zero.
3/ Total Count = Positive Count + Zero Count
4/ Code(s) of the reason(s) for at least half of the EMI's.
* At least half of the EMI's were clerical in nature.

Count of EMI's for Selected Item Codes
Iowa (19)

Item Code	EMI Count	Positive Count 1/	Zero EMI Count 2/	Total Count 3/	Percent of Total	Reason Code 4/
IC831	12	27	12	39	30.77	1
IC837	12	27	12	39	30.77	1
IC852	12	27	11	38	31.58	1
IC27	11	353	0	353	3.12	*
IC156	11	403	0	403	2.73	10,3
IC179	11	159	3	162	6.79	1,5
IC309	11	21	0	21	52.38	5
IC620	11	367	0	367	3.00	2,4
IC735	11	105	5	110	10.00	1,4,3
IC752	11	164	0	164	6.71	1,10
IC790	11	83	0	83	13.25	5
IC36	10	369	0	369	2.71	*
IC68	10	288	0	288	3.47	5
IC172	10	293	0	293	3.41	5
IC189	10	345	0	345	2.90	2
IC190	10	344	0	344	2.91	2
IC191	10	344	0	344	2.91	2
IC192	10	342	0	342	2.92	2
IC279	10	11	2	13	76.92	1
IC425	10	17	0	17	58.82	7
IC650	10	137	0	137	7.30	5,2
IC740	10	116	5	121	8.26	1
IC751	10	216	0	216	4.63	7,3
IC896	10	14	5	19	52.63	1

1/ Positive count is the number of questionnaires with summarized positive values.
2/ Zero count is the number of questionnaires where an item was edited to zero.
3/ Total Count = Positive Count + Zero Count
4/ Code(s) of the reason(s) for at least half of the EMI's.
* At least half of the EMI's were clerical in nature.

Count of EMI's for Selected Item Codes
North Carolina (37)

Item Code	EMI Count	Positive Count 1/	Zero EMI Count 2/	Total Count 3/	Percent of Total	Reason Code 4/
IC669	113	316	0	316	35.76	5
IC431	104	191	1	192	54.17	4
IC127	101	319	0	319	31.66	2
IC614	76	308	5	313	24.28	2
IC552	75	158	2	160	46.88	1,2
IC652	73	322	0	322	22.67	2
IC129	56	322	0	322	17.39	2
IC654	49	284	0	284	17.25	5
IC607	47	301	6	307	15.31	2
IC161	44	304	2	306	14.38	2,5
IC568	37	118	3	121	30.58	5
IC626	36	265	7	272	13.24	2,6
IC613	33	305	5	310	10.65	5,1
IC167	32	91	2	93	34.41	5
IC657	32	322	0	322	9.94	2
IC605	30	301	7	308	9.74	1,2
IC466	29	28	1	29	100.00	12,4
IC606	28	301	7	308	9.09	2,1
IC609	28	294	9	303	9.24	5
IC19	26	127	0	127	20.47	2
IC137	26	277	0	277	9.39	3,5
IC356	26	1	26	27	96.30	1
IC635	26	92	8	100	26.00	11,2
IC160	25	317	0	317	7.89	2,5
IC556	24	66	0	66	36.36	5,1
IC599	24	141	2	143	16.78	2,4
IC617	23	181	8	189	12.17	1,5
IC131	21	64	16	80	26.25	1
IC436	21	22	0	22	95.45	4,12
IC486	21	22	0	22	95.45	12
IC572	21	128	2	130	16.15	5
IC138	20	221	0	221	9.05	3
IC164	20	152	7	159	12.58	7
IC450	20	21	0	21	95.24	12
IC758	20	24	11	35	57.14	1,7
IC27	19	322	0	322	5.90	*
IC276	19	38	0	38	50.00	12,5
IC739	19	97	0	97	19.59	4

1/ Positive count is the number of questionnaires with summarized positive values.

2/ Zero count is the number of questionnaires where an item was edited to zero.

3/ Total court = Positive Count + Zero Count

4/ Code(s) of the reason(s) for at least half of the EMI's.

* At least half of the EMI's were clerical in nature.

Count of EMI's for Selected Item Codes
North Carolina (37)

Item Code	EMI Count	Positive Count 1/	Zero EMI Count 2/	Total Count 3/	Percent of Total	Reason Code 4/
IC608	18	257	6	263	6.84	1
IC705	18	97	1	98	18.37	3,5
IC50	17	120	0	120	14.17	5
IC132	17	62	0	62	27.42	5
IC165	17	121	9	130	13.08	2
IC180	17	97	3	100	17.00	3,5
IC596	17	65	3	68	25.00	1
IC600	17	82	1	83	20.48	2
IC618	17	191	10	201	8.46	1
IC141	16	156	0	156	10.26	2,5
IC158	16	101	0	101	15.84	5,3
IC458	16	16	0	16	100.00	4,12
IC619	16	156	3	159	10.06	5,2
IC765	16	69	0	69	23.19	5
IC892	16	55	1	56	28.57	5
IC17	15	308	0	308	4.87	2
IC468	15	16	0	16	93.75	12,6
IC631	15	55	5	60	25.00	11,7
IC733	15	115	0	115	13.04	4
IC136	14	255	0	255	5.49	2
IC278	14	9	7	16	87.50	1
IC598	14	153	5	158	8.86	1,2
IC634	14	70	8	78	17.95	11
IC764	14	78	1	79	17.72	5
IC769	14	43	4	47	29.79	5
IC49	13	118	0	118	11.02	*
IC157	13	157	0	157	8.28	3
IC256	13	0	13	13	100.00	1
IC257	13	0	13	13	100.00	1
IC766	13	54	5	59	22.03	5
IC112	12	90	0	90	13.33	2
IC128	12	77	5	82	14.63	1,2
IC179	12	81	6	87	13.79	1,3
IC217	12	176	0	176	6.82	1,5
IC539	12	62	0	62	19.35	2,1
IC543	12	79	0	79	15.19	2,5
IC621	12	157	2	159	7.55	7,2
IC629	12	53	4	57	21.05	11

1/ Positive count is the number of questionnaires with summarized positive values.
2/ Zero count is the number of questionnaires where an item was edited to zero.
3/ Total Count = Positive Count + Zero Count
4/ Code(s) of the reason(s) for at least half of the EMI's.
* At least half of the EMI's were clerical in nature.

Count of EMI's for Selected Item Codes
North Carolina (37)

Item Code	EMI Count	Positive Count 1/	Zero EMI Count 2/	Total Count 3/	Percent of Total	Reason Code 4/
IC757	12	73	0	73	16.44	5
IC768	12	52	6	58	20.69	5
IC771	12	43	5	48	25.00	5
IC25	11	40	0	40	27.50	2
IC111	11	90	0	90	12.22	2
IC277	11	29	0	29	37.93	5
IC656	11	284	0	284	3.87	5
IC767	11	43	5	48	22.92	5
IC770	11	43	5	48	22.92	5,1
IC120	10	18	6	24	41.67	1
IC122	10	195	1	196	5.10	2,1
IC215	10	25	8	33	30.30	1
IC578	10	44	1	45	22.22	5
IC625	10	58	0	58	17.24	7
IC630	10	45	5	50	20.00	11
IC727	10	16	1	17	58.82	7
IC762	10	36	9	45	22.22	1
IC772	10	62	5	67	14.93	1
IC790	10	60	1	61	16.39	5,1

1/   Positive count is the number of questionnaires with summarized positive values.
2/   Zero count is the number of questionnaires where an item was edited to zero.
3/   Total Count = Positive Count + Zero Count
4/   Code(s) of the reason(s) for at least half of the EMI's.
*   At least half of the EMI's were clerical in nature.