United States
Department of
Agriculture

National
Agricultural
Statistics
Service

Research and
Applications
Division

SRB Research Report
Number SRB-89-08

March 1989

USE OF REINTERVIEW TECHNIQUES FOR QUALITY ASSURANCE: The Measurement of Response Bias in the Collection of December 1987 Quarterly Grain Stocks Data Using CATI

Brad Pafford

USE OF REINTERVIEW TECHNIQUES FOR QUALITY ASSURANCE: THE MEASUREMENT OF RESPONSE BIAS IN THE COLLECTION OF DECEMBER 1987 QUARTERLY GRAIN STOCKS DATA USING CATI. by Brad Pafford. Research and Applications Division, National Agricultural Statistics Service, U.S. Department of Agriculture, Washington, D.C. 20250. March 1989. NASS Research Report No. SRB 89-08.

ABSTRACT

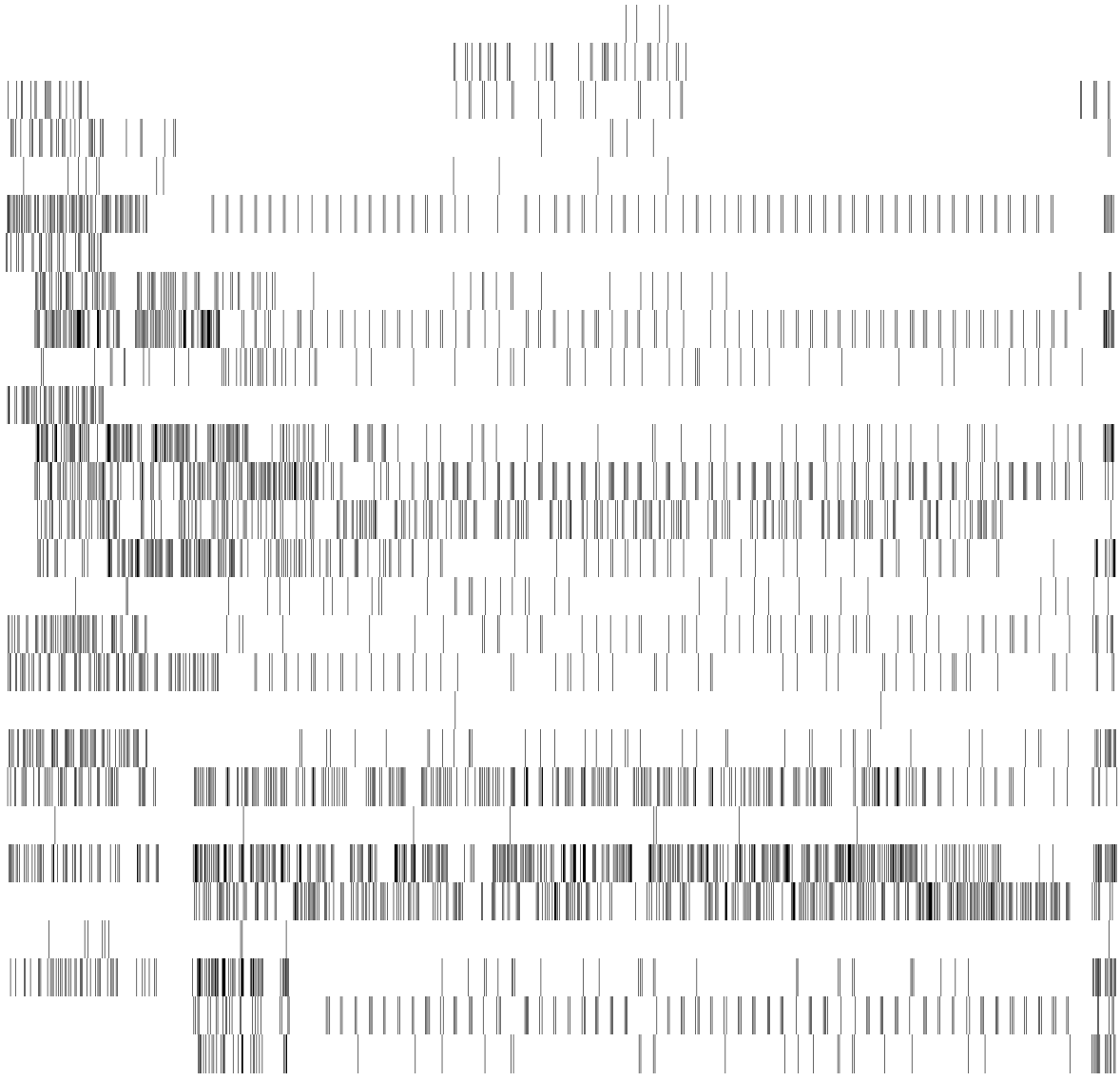
Reinterview survey procedures were used to recontact approximately 1,000 December 1987 Agricultural Survey respondents who were initially interviewed using the Computer Assisted Telephone Interviewing (CATI) system. The purpose of the reinterview program is to measure bias in the reporting of Agricultural Survey items such as grain stocks, hogs, major crops and operation screening questions, as well as to uncover why errors are occurring. This paper reports measures of bias in grain stocks and identifies reasons for biased reporting. Significant under-reporting was uncovered in almost all stocks items in the three reinterview states in the study - Minnesota, Indiana, and Ohio. A large portion of the bias, 46% in corn stocks reporting, could be classified as having occurred because of "definitional" errors, while a much smaller portion of the total bias was due to "estimating/rounding" reasons. The under-reporting of corn stocks on a per-sample basis was 34% larger when someone other than the operator, partner, or manager was interviewed on CATI. Finally, as the size of the farm operation increased so did corn stocks bias.

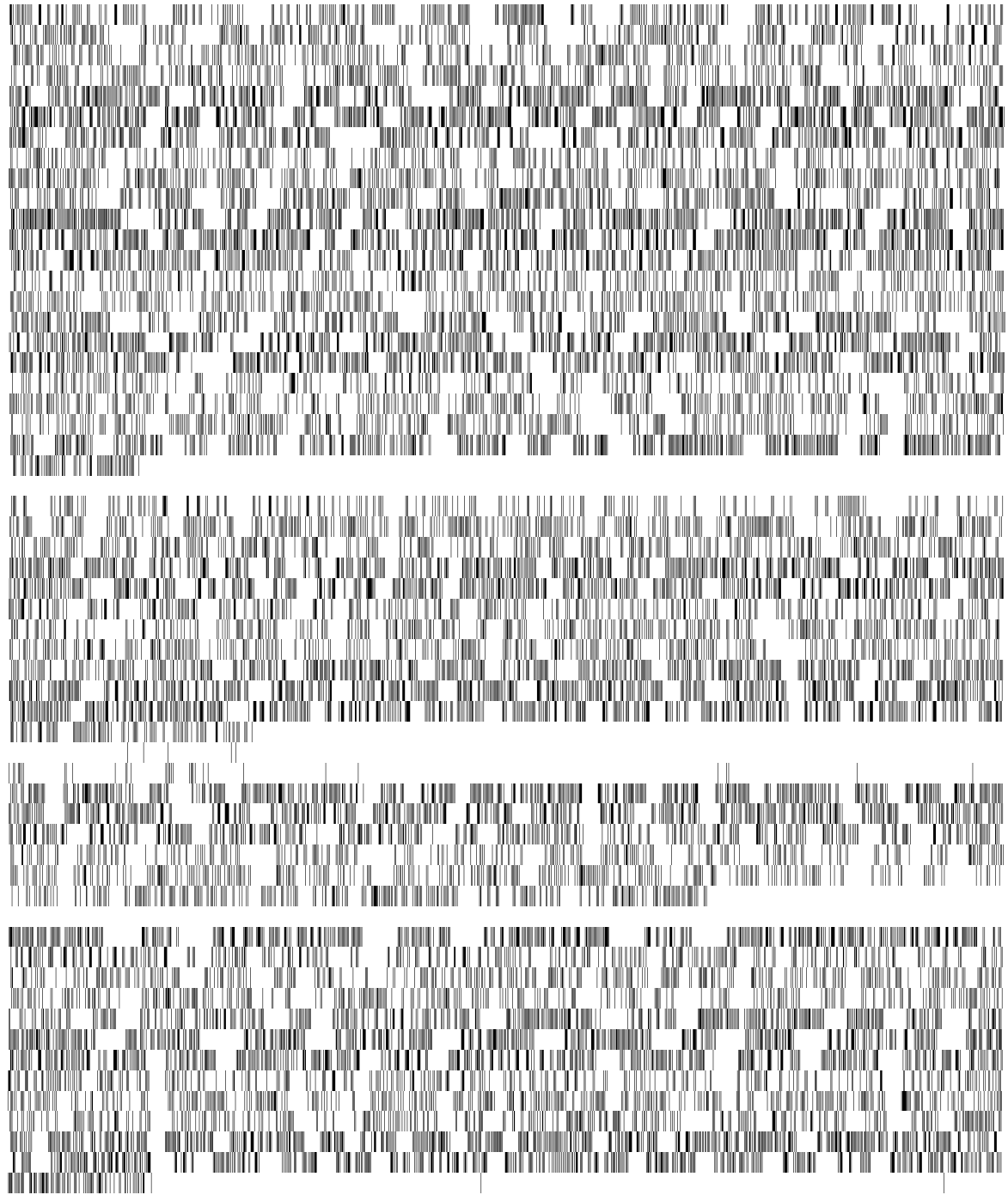
KEYWORDS: Bias, definitional errors, and respondent

****
* This paper was prepared for limited distribution to the research *
* community outside the U.S. Department of Agriculture. The views *
* expressed herein are not necessarily those of NASS or USDA. *
****

ACKNOWLEDGMENTS

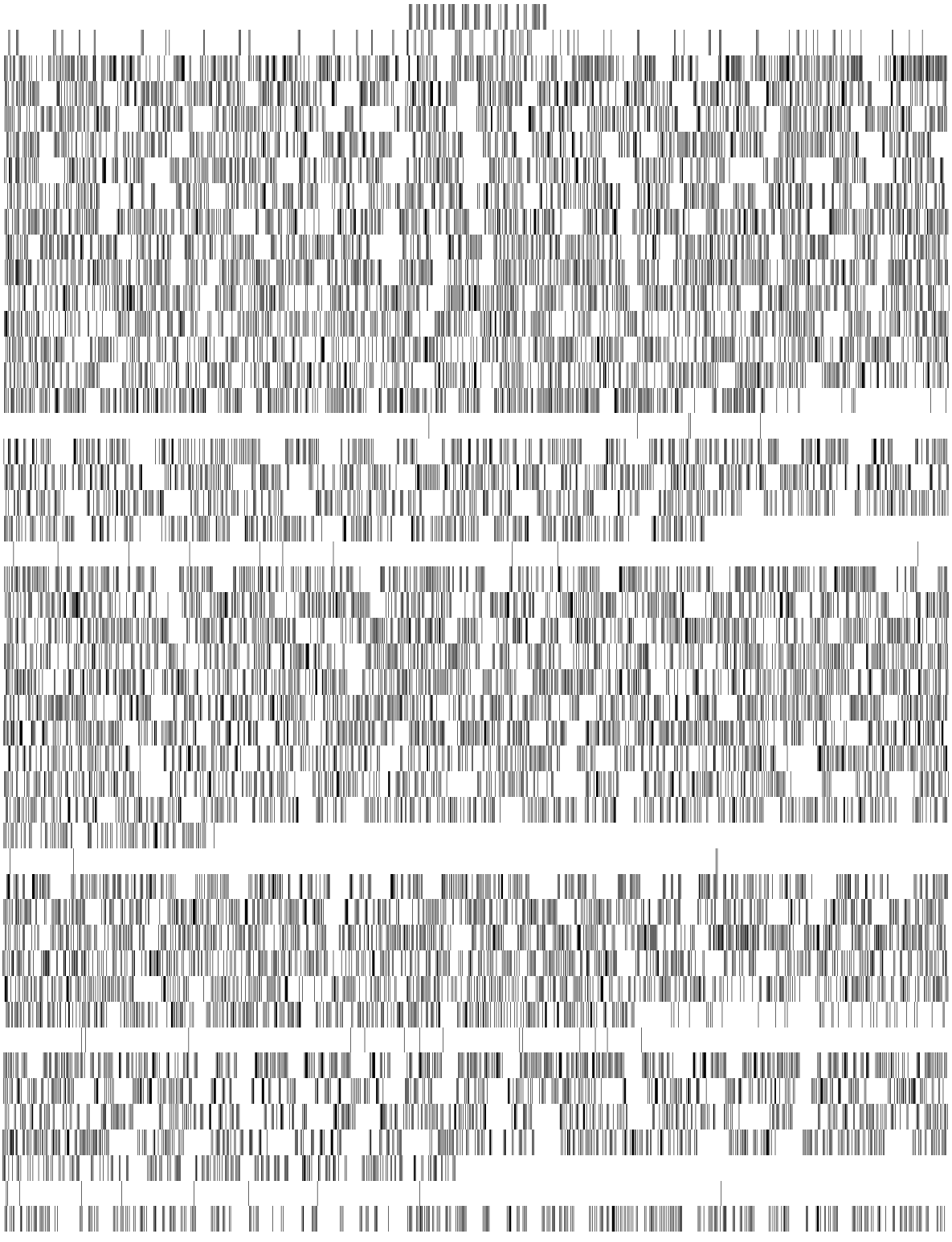
The author is deeply indebted to many people for the successful completion of this project. First, I would like to thank my immediate colleagues, Gretchen McClung and Vic Tolomeo for the many hours we prepared together for the survey, as well as in working with the data. Thanks also to Ron Fecso and George Hanuschak. There were many hands-in-the-pot from the Survey Management Branch. Thanks so much to this group who were extremely cooperative in getting this rather new type of data collection effort underway. Finally, special thanks go to the Minnesota, Indiana, and Ohio field offices and their enumerators who cooperated under otherwise heavy workloads.

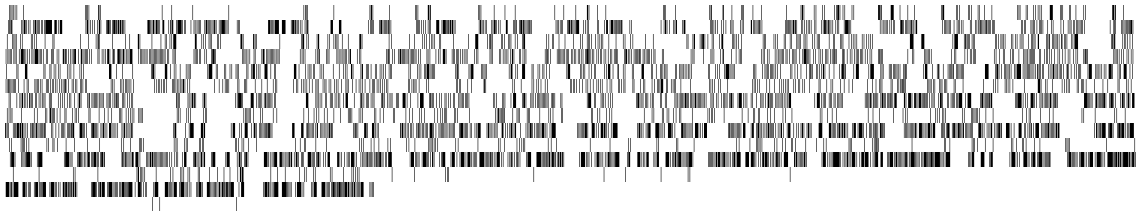


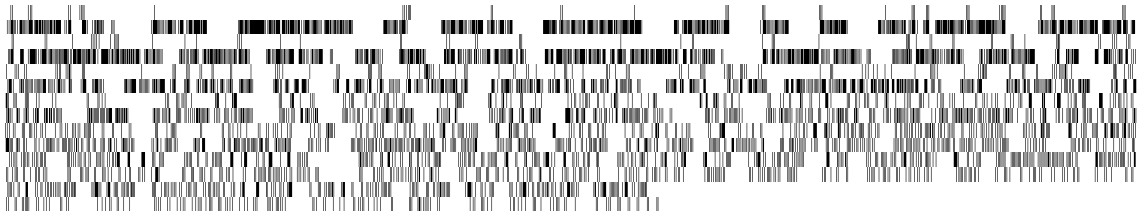


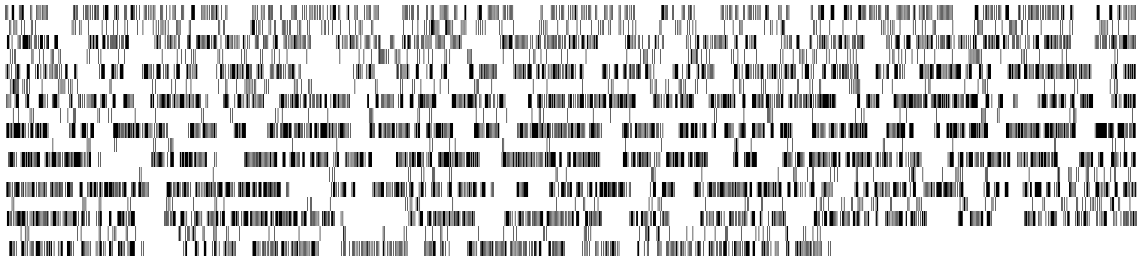
Finally, corn stocks bias was estimated by size of farm operation. The results showed that as the size of farm increases so does the corn stocks bias. The collection of stocks data from large farms needs to be done in the best possible manner as a few large farms can significantly bias stocks reporting. Interviews probably need to be face-to-face and contacts made with the farm operator(s). However, without measures of bias in face-to-face interviews a specific recommendation of this kind cannot be made.

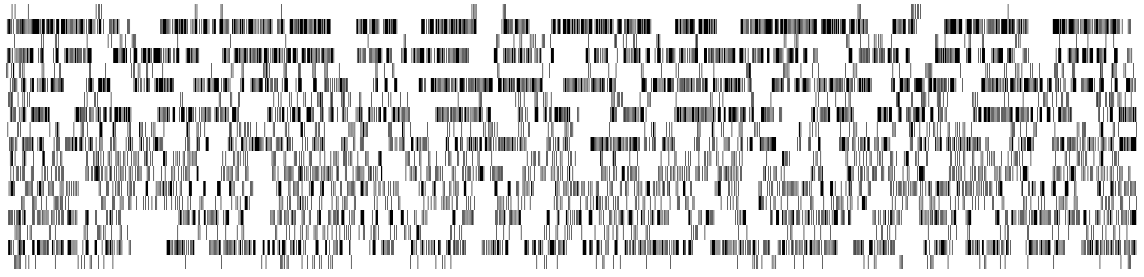
Recommendations call for expanding the reinterview study to a larger number of states, and to other modes of data collection such as non-CATI telephone and personal interviews. Then a data series on the bias can be established to provide the ASB with direct measures of quality of the AS data. A procedure is recommended on how the ASB might use the bias estimates over time. Finally, alternative CATI questionnaire designs and improved enumerator training should be evaluated to eliminate some of the major contributors to the bias.

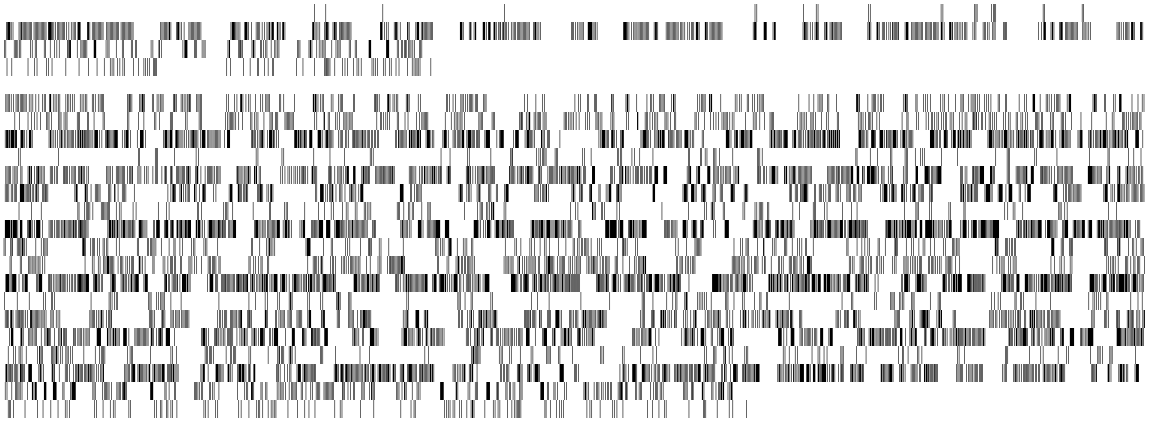


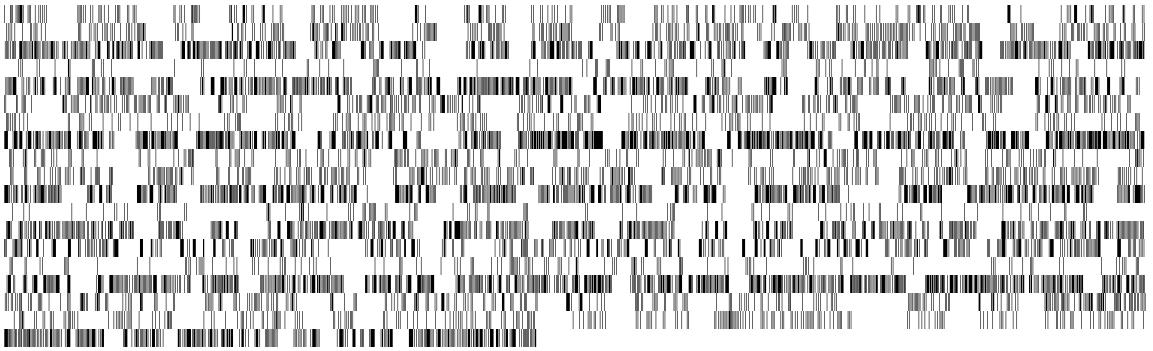


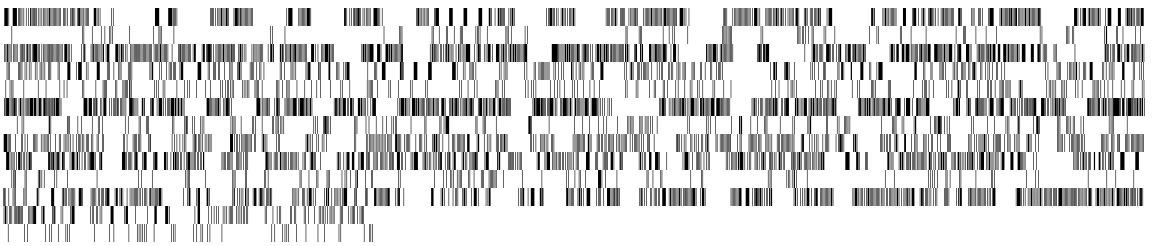


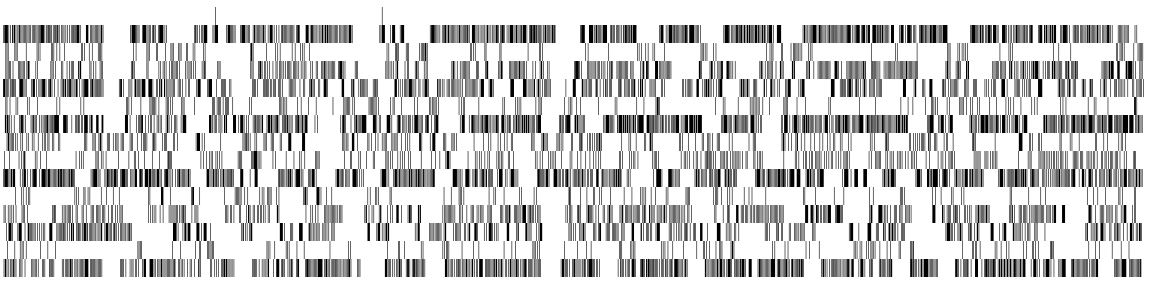


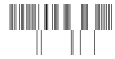


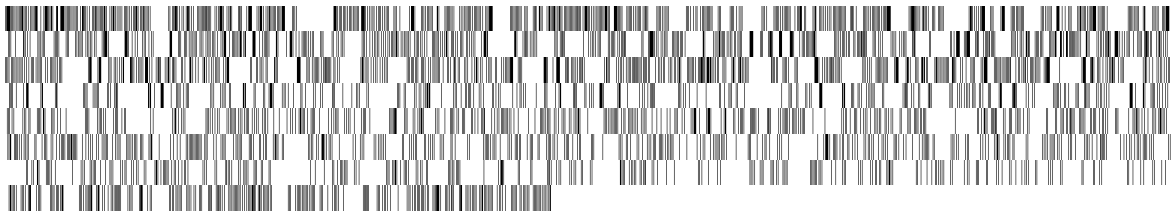


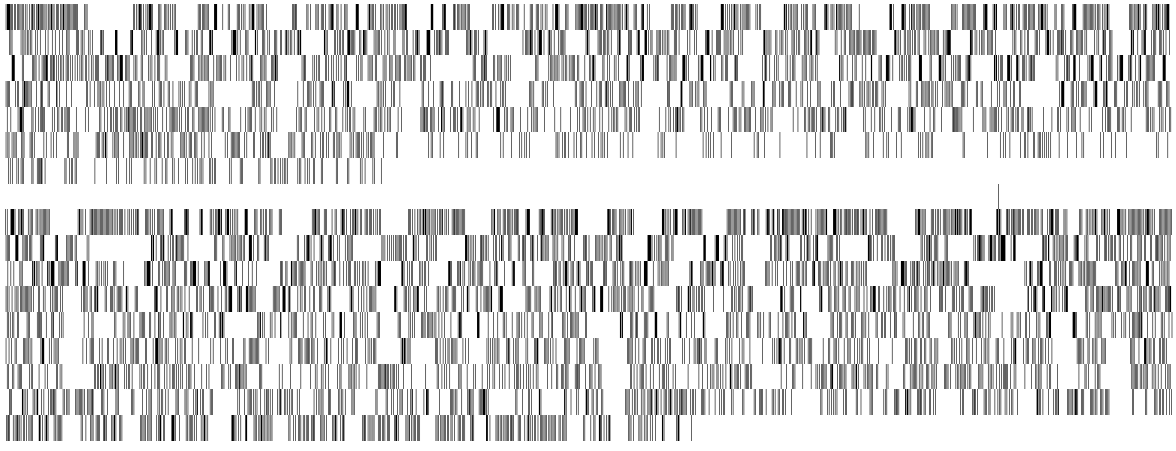


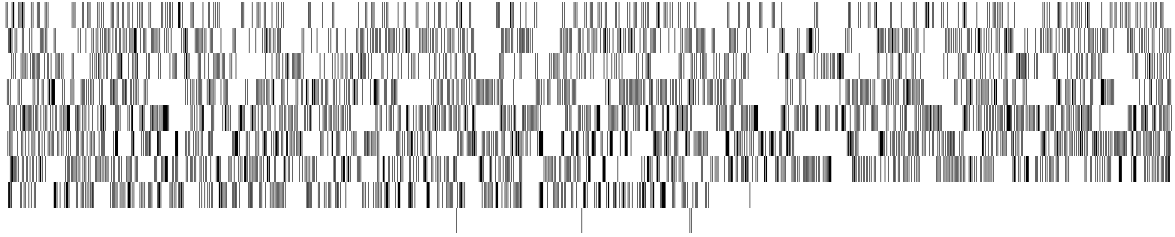


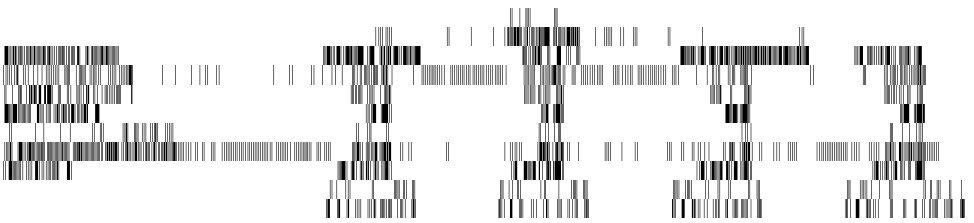


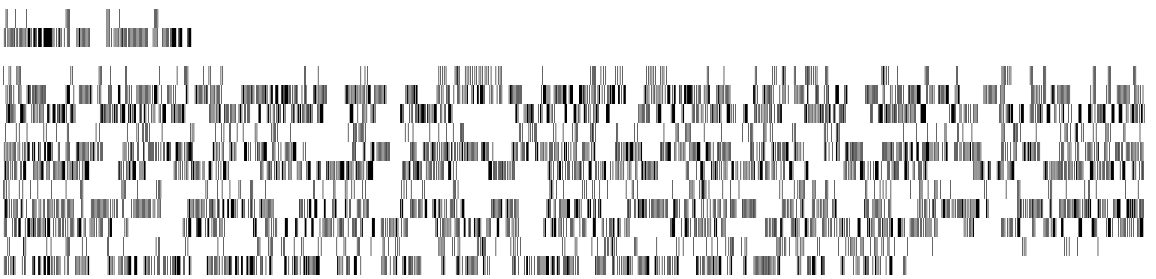


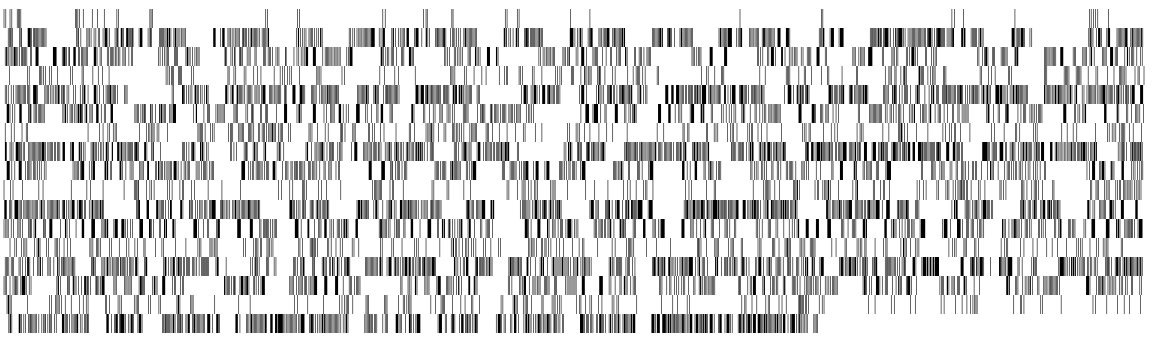


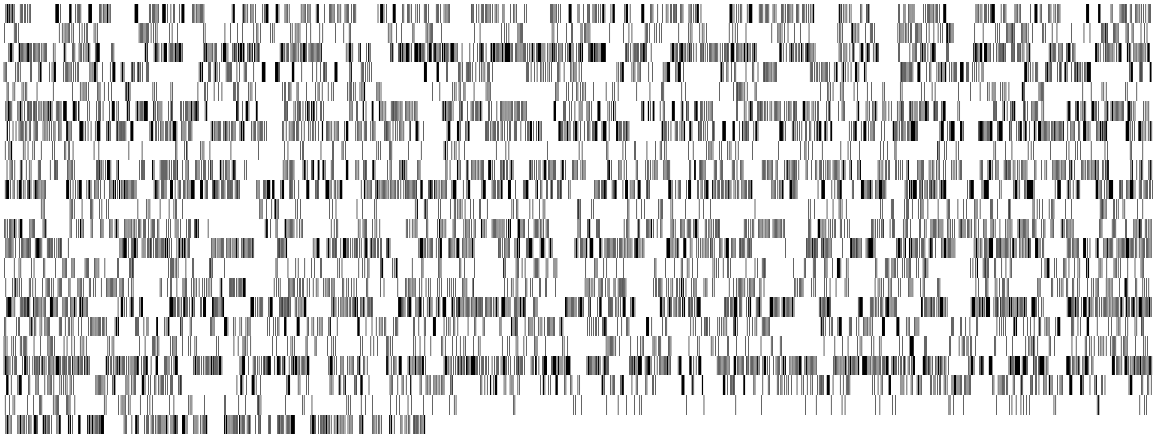


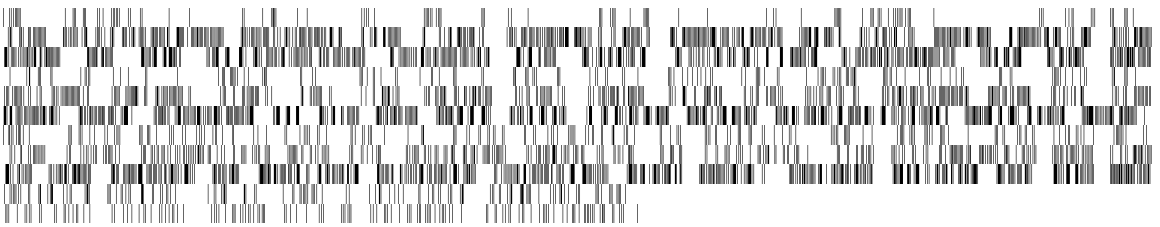












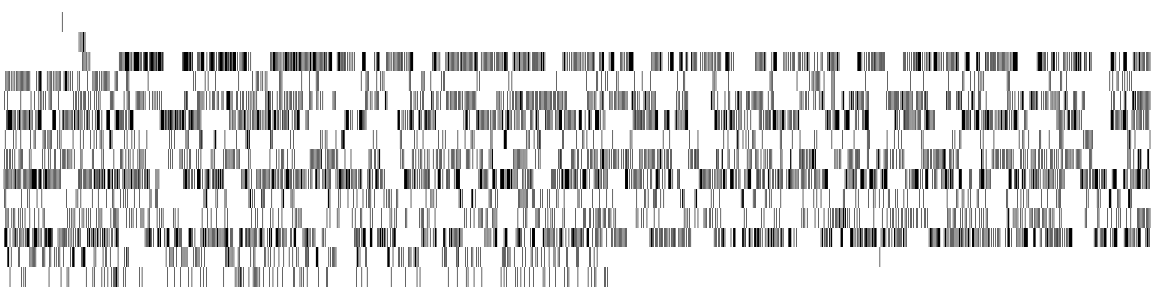


Table 2. Estimates of bias in the CATI collected data - December 1987 reinterview study.

Survey Item/State	(CATI - Reconciled) 000 bu.	% of CATI		(SAS - Reconciled) 000 bu.	% of SAS	
Corn Stocks						
Minnesota	-60,987	-10.5	*	-64,860	-11.4	*
Indiana	-48,289	-16.1	*	-46,974	-16.0	*
Ohio	-22,609	-12.0	*	-22,277	-11.8	*
Total	-131,885	-12.3	*	-134,111	-12.8	*
Soybean Stocks						
Minnesota	-13,082	-14.7	*	-13,082	-14.7	*
Indiana	-3,012	-5.9		-3,012	-5.9	
Ohio	-5,393	-13.6	*	-5,640	-14.2	*
Total	-21,487	-12.0	*	-21,734	-12.1	*
All Wheat Stocks						
Minnesota	-17,853	-24.9	*	-17,853	-22.7	*
Indiana	-2,513	-195.0		-2,513	-198.6	
Ohio	6	0.2		-253	-7.5	
Total	-20,360	-26.5	*	-20,619	-24.8	*
Storage Capacity						
Minnesota	64,186	4.7		71,495	5.5	
Indiana	55,379	9.1		-10,998	-2.0	
Ohio	-4,273	-1.1		-5,619	-1.4	
Total	115,292	4.9		54,878	2.4	

* - Indicates significance at the α=.05 level.

Reasons for Discrepancies

Appendix C, Tables 1-4, gives the distribution of reasons the reinterview respondents gave for differences in the CATI and reinterview responses. These differences are grouped into the categories of estimation/rounding, definitional, and other reasons. Estimation/rounding occurred when the respondent felt that one or more of the responses were estimated, or differences were too small to be bothered with, which included rounding problems. The second classification is termed definitional. Here, the respondent gave specific mention to reasons that are directly attributable to a lack of understanding of what should have been included and excluded in the stocks. These cover a large spectrum ranging from confusion with government storage to confusion with stocks sold but still on the operation. The final category was the "other" class. These responses were neither definitional nor estimation related. They included responses such as "doesn't give out information over the phone," "doesn't know why there is a difference," and "mistake in addition."

These categories help to draw attention to the percent of "definitional" differences that are perhaps fixable, compared with "other" and "estimating/rounding" differences that are perhaps not fixable. By fixable it is meant that a change in such things as survey procedures, questionnaire design, and training would decrease the error. The percentage of errors in corn stocks classified as definitional in nature ranged from 32.3% to 43.8% in the three states (Appendix C, Table 1). In soybean stocks, definitional problems occurred in a range of 15.8% to 36.7% in the three states (Appendix C, Table 2). Therefore, when differences existed there were often specific reasons for it, and not just rounding (however, I am sure rounding exists for every response to some extent). It will be shown in the next few sections that definitional reasons contribute extensively to the bias levels uncovered in this study.

A quick inspection of Tables 1-4 in Appendix C also reveals the sheer variety of reasons for definitional differences. This points to the difficulty NASS will have in designing a survey form, properly training enumerators, and/or developing survey procedures to accurately collect these data.

The other two classes of errors, estimating/rounding and other, account for approximately 60% to 70% of the reasons for discrepancies. Tables 1-4 in Appendix C reveal that the percentage of differences classified as estimating/rounding ranged between 30% and 50%, while the range was 17% to 35% for differences classified in the "other" category. These errors probably will be difficult to eliminate.

As mentioned, it is important to look at more than just these distributions. We must see which contribute most to the overall bias. This question will be addressed shortly.

Change in Respondent

Table 3 presents the distribution of respondents reporting for the CATI interview and the reinterview. Individual operator-individual operator, for example, means that there was only one operator for the farm who was the respondent in both the CATI interview and again in the reinterview. This occurred in approximately 74% of the interview-reinterviews. In general, the spouse reported for the operator in his/her absence in the reinterview.

Also shown in Table 3 are three catagories of respondent combinations. These are used in the next sections when bias is estimated for different subgroups of the reinterview sample. When the operator, manager or partner responded in the reinterview and were also the CATI respondent, they were placed in the "operator-operator" category. Estimated bias for this group represents the bias in operator reporting. When the operator, partner, or manager was the reinterview respondent but someone else was the CATI respondent they were placed in the "other-operator" group. Bias estimates for this group represent the bias in non-operator reporting. The "other combinations" group represents every other combination of reporting. Measures of bias from the "other combinations" group are not as clean as for the first two groups, because the operator(s) were not contacted in the reinterview. For this reason, the paper does not discuss this respondent category in detail.

Table 3. Frequency distribution of respondents reporting in the CATI interview and reinterview for all three reinterview states combined.

CATI-Reinterview	Freq.	% of Total
Operator-Operator	749 (80.5%)	
Indivdual operator-individual operator	684	73.6
Partner-same partner	51	5.5
Partner-different partner	12	1.3
Manager-manager	2	0.2
Other-Operator	59 (6.3%)	
Spouse-individual operator	49	5.3
Other-individual operator	10	1.1
Spouse-partner	0	0.0
Other-partner	0	0.0
Other Combinations	122 (13.2%)	
Individual operator-spouse	62	6.7
Individual operator-other	28	3.0
Spouse-spouse	20	2.2
Partner-spouse	2	0.2
Spouse-other	4	0.4
Other-same other	4	0.4
All Other Combinations	2	0.2
Total	930	100.0

Change in Respondent and Reason for Difference as they Affect Stocks Reporting

This section presents estimated stocks bias by reason for discrepancy and respondent. Also, discussed are relationships between respondent, reason for discrepancy, and stocks bias.

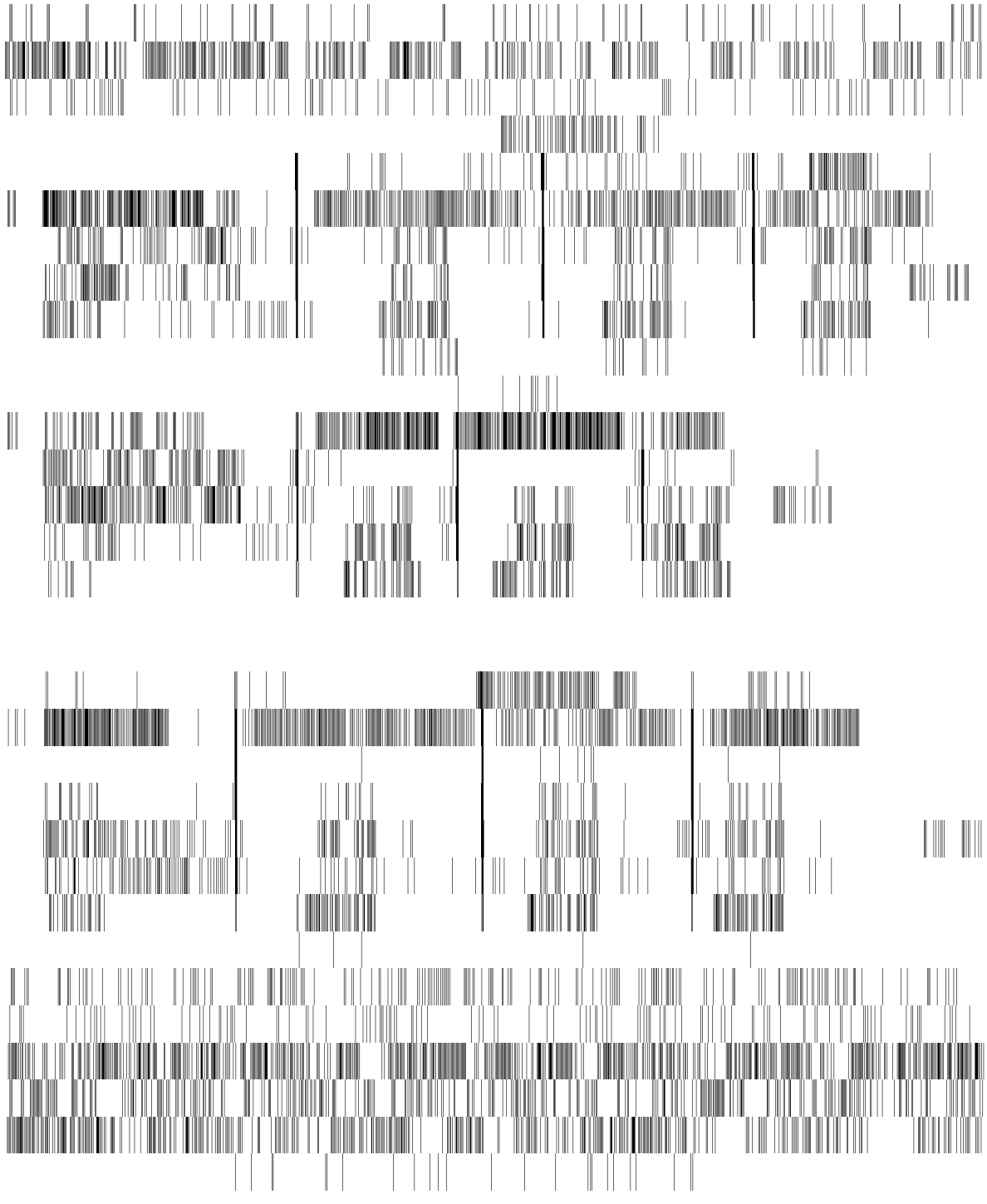
The analysis is given in two parts. First, Tables 4A thru 4C compare the distribution of responses in two corn stocks relative bias categories ("less than 25% change" and "more than 25% change) with the three respondent categories ("operator-operator," "other-operator," and "other combinations"), and with the three reasons for differences categories ("other," "definitional," or "estimating").

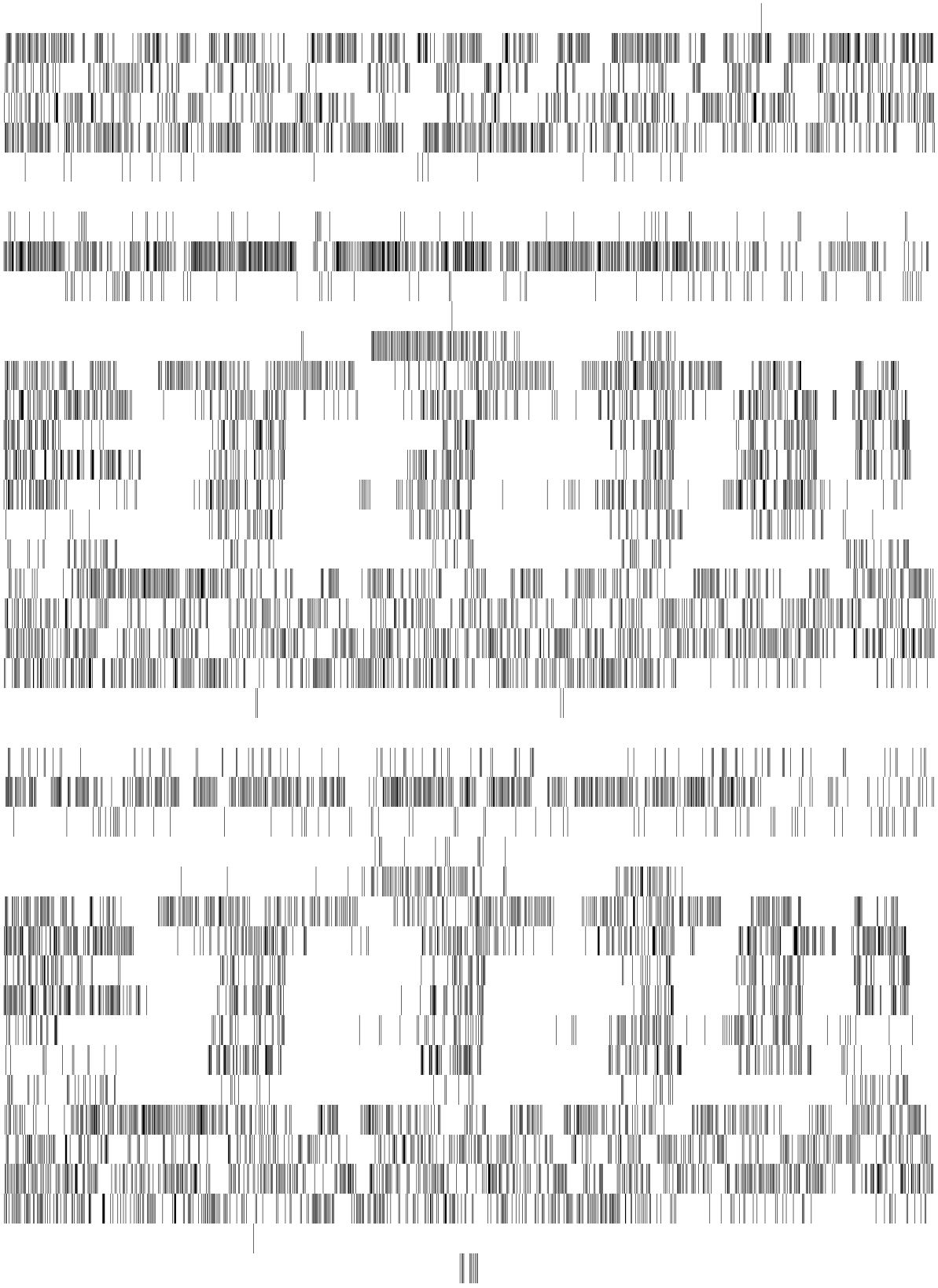
The second part of the analysis is to estimate the bias by expansion of the data for these same categories. With this second approach we can answer questions about what the impact is on the bias due to such things as definitional problems or respondents other than the operator reporting. These results appear in Tables 5 and 6.

First, in Table 4A we see that 66.5% of the "operator-operator" respondents had relative biases that were less than 25% in absolute value. Compare this with the "other-operator" group where there were only 38.9% with relative biases in this range. This unequal distribution is significant ($p=.03$). That is, we are more likely to see a larger relative bias when the CATI respondent is not a farm operator (i.e., the operator, manager, or partner).

In Table 4B we see that definitional problems are associated with large relative biases, while estimating problems are associated with small relative biases ($p<.01$). This can be seen by the fact that 59.1% of definitional problems are associated with relative biases of 25% or more in absolute value, compared with only 5.9% of estimating/rounding errors associated with relative biases of 25% or more in absolute value. Definitional problems are something to be concerned with because they probably contribute a lot to the overall bias.

Finally, in Table 4C we see that reasons for differences are reported in the same percentage for each respondent group ($p=.86$). That is, the "operator-operator, "other-operator," and "other combinations" groups report definitional, other and estimating problems in the same ratio (within sampling error).





This ratio was 45.4% for soybean stocks. "Other" reasons contribute the bulk of remaining bias (43.3% and 40.5%). Estimating reasons contribute a minimal 10.7% and 14.1% of the bias in corn and soybean stocks, respectively.

We also see from Tables 5 and 6 that operators and non-operators contribute to the under-reporting of stocks. For example, approximately 100 million additional bushels of corn stocks were uncovered in the reconciliation from operators (the "operator-operator" group), and approximately 10 million bushels from non-operators (the "other-operator" group).

Also interesting was that 50% of the "operator-operator" corn stocks bias was due to definitional problems (50,074 divided by 100,371) and only 8% to estimating problems, while the percentages were 31% and 63% for the "other-operator" group. In other words, definitional problems caused errors for the farm operators, while non-operator errors seemed to be mostly estimating in nature. The farm operators said the non-operators were estimating, while the operators said they themselves were confused on what to include or exclude in stocks reporting.

The final point of interest in tables 5 and 6 is that non-operators (the "other-operator" respondent group) contribute more bias on a unit or per respondent level than do operators (the "operator-operator" respondent group). This can be seen by simply dividing the total bias by the number of respondents in that category. For example, a bias in corn stocks of approximately 138,600 bushels per sample exists for the "operator-operator" group (100,371/724=138,600) compared with 185,800 bushels per sample for the "other-operator" respondent group. This is an increase of 34%. For soybean stocks, this percentage increase was 15%. The importance of this comparison is to draw further attention to problem of contacting anyone other than the farm operator(s).

Bias by Size of Farm Operation

Table 7 presents estimates of bias in corn stocks reporting by size of operation. Size of operation is defined by land-in-farm acreage.

Table 7. Corn stocks bias estimates by size of operation (land-in-farm acreage) for all states in the December 1987 Reinterview Study.

Land-in-Farm Classification (Acres)	Corn Stocks Bias (000 bu.)	Response (n)
0 - 99	-5,843	182
100 - 249	-5,842	262
250 - 499	-13,415	238
500 - 999	-43,575	138
1,000 +	-63,210	73
Total	-131,885	893

[redacted]

[redacted]

[redacted]

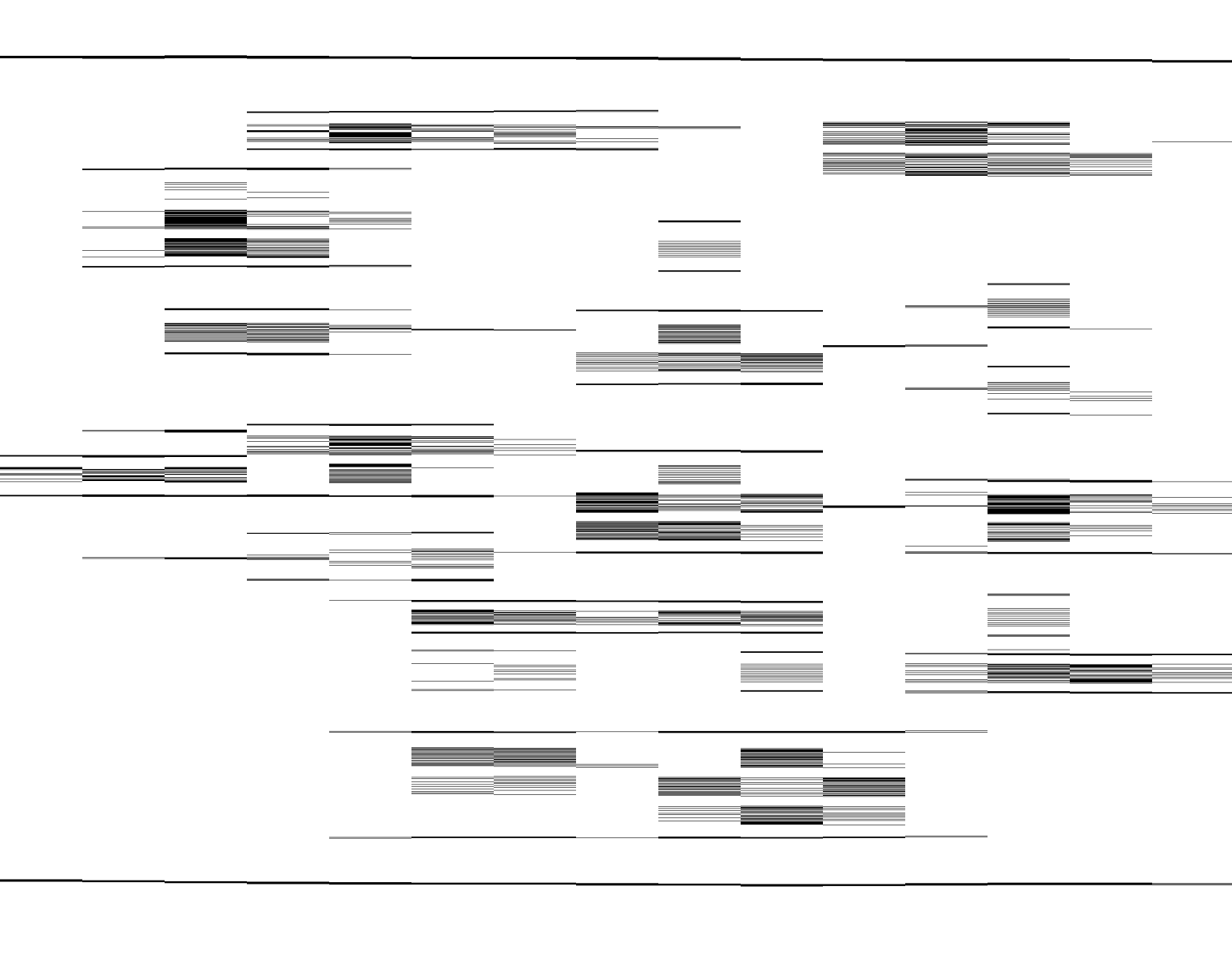
[redacted]

[redacted]

[redacted]

[redacted]

[redacted]



on the land a farmer operates. Continuing along with the flow diagram of figure 1 we see that once a bias is uncovered in year one there is a need to repeat the reinterview study to assess the new bias levels. If they do not exist, then the reinterview study could be discontinued for a period of time. If they do exist, one needs to ask whether the biases are constant or variable. If constant, periodic measures of bias would be needed to validate this constant bias. If the bias is variable, national measures of accuracy are needed on a yearly basis.

If a bias is not present in the first year, one follows down a different path of figure 1, however, similar concepts exist. Figure 1 provides answers then to the question, "How long do we need to do reinterviews?" It also provides a coherent and sound policy for the measures of accuracy in the AS program.

If we wish to consider eliminating the important bias producing factors in stocks reporting, the study suggests several areas to examine. The most controllable sources of bias creating agents come from the definitional type errors. These were found to contribute around 45% of the total bias. The ability to eliminate estimating/rounding errors will be minimal unless the respondent is restricted to the farm operators. Even then, they cannot be totally eliminated. The good news is that estimating/rounding errors contribute less than 15% of the bias. The "other" errors are also probably difficult to control. Some are due to the use of the telephone, such as "can't hear well on the phone" and "didn't take the time on the phone to add correctly." Some may be indirectly related to the use of the phone, such as "thought he/she reported this the first time," and "misunderstanding between enumerator and respondent." Finally, some "other" type errors will occur in any interview environment.

For the definitional type errors, a more detailed stocks section may need to be developed with CATI. Questions could appear on the CATI screens that specifically ask the respondent if, for example, he/she has included stocks on all land operated. Another CATI screen would ask if they included grain stored on his/her operation belonging to someone else. This could be continued until the most important reasons for biases are addressed with the respondent during the interview. The sheer number and variety of definitional problems would make including every reason for difference impossible.

A second procedure to remove the definitional type biases is through improved telephone training and careful screening of potential telephone enumerators (they must be qualified). It is reasonable to suspect that some definitional problems are the result of enumerators lack of understanding.

The respondent was also shown to be an important contributor to errors: in this case bias. NASS's operational procedures do not specifically say the operator(s) must be contacted, however, over three-quarters of the time they are. It is the remaining non-operator contacts that are more of a problem. The fluctuation in the level of other-than-operator respondents from survey to survey can be critical. For example, a change in the other-than-operator respondents from 25% to 30% may increase the level of bias by a factor of 34% (using the corn stocks bias per sample estimate) for the additional 5% of the farm operations that were not operator reported. NASS should begin to consider

research into modeling this source of error into its survey indications. This factor could be used to adjust survey indications conditioned on the percent of non-operator reporting.

Finally, the results of this study show that stocks reporting bias increases as the size of the farm operation increases. Because of this and because relatively few operations make up this large operation group, there is probably a need to improve the way these data are collected. One approach may be to conduct face-to-face interviews of farm operator(s) for these large stocks operations, similar to the way livestock extreme operators (EO's) are handled. However, without measures of bias in face-to-face interviews this cannot be a specific recommendation.

RECOMMENDATIONS

Based on the results of this study the following are recommended:

1. Continue reinterviews for the December AS to develop a data series on the bias. The number of states should be increased above three, and include those states which have the majority of the U.S. on-farm stocks.

2. Reinterviews should also be conducted for other modes of data collection, such as for mail, personal and non-CATI telephone interviews. The level of bias and response errors may not be the same in these modes of data collection as that found with CATI. However, if national CATI is implemented in the next few years, the use of other modes of data collection will diminish for the Agricultural Survey. If this occurs, then the resources required to test other modes of data collection may be better spent elsewhere.

3. Design a new stocks section for the Agricultural Survey questionnaire to eliminate bias caused by definitional problems. This should begin immediately. One alternative is to choose two or three of the largest definitional errors and include these questions in the stocks section. The largest definitional errors are defined as those with the largest bias.

4. CATI enumerator training procedures should be evaluated to see if errors can be reduced through improved training.

5. The Agricultural Statistics Board should use the data series on the bias in stocks reporting in formulating its official statistics. Use of a bias decision tree such as the one in figure 1 should also be used by the Board for recommending future data collection efforts of this type.

[illegible]

[illegible], William A., [illegible] A. [illegible], Michael W. [illegible], and [illegible] M. [illegible]. [illegible]. "[illegible] [illegible] [illegible] [

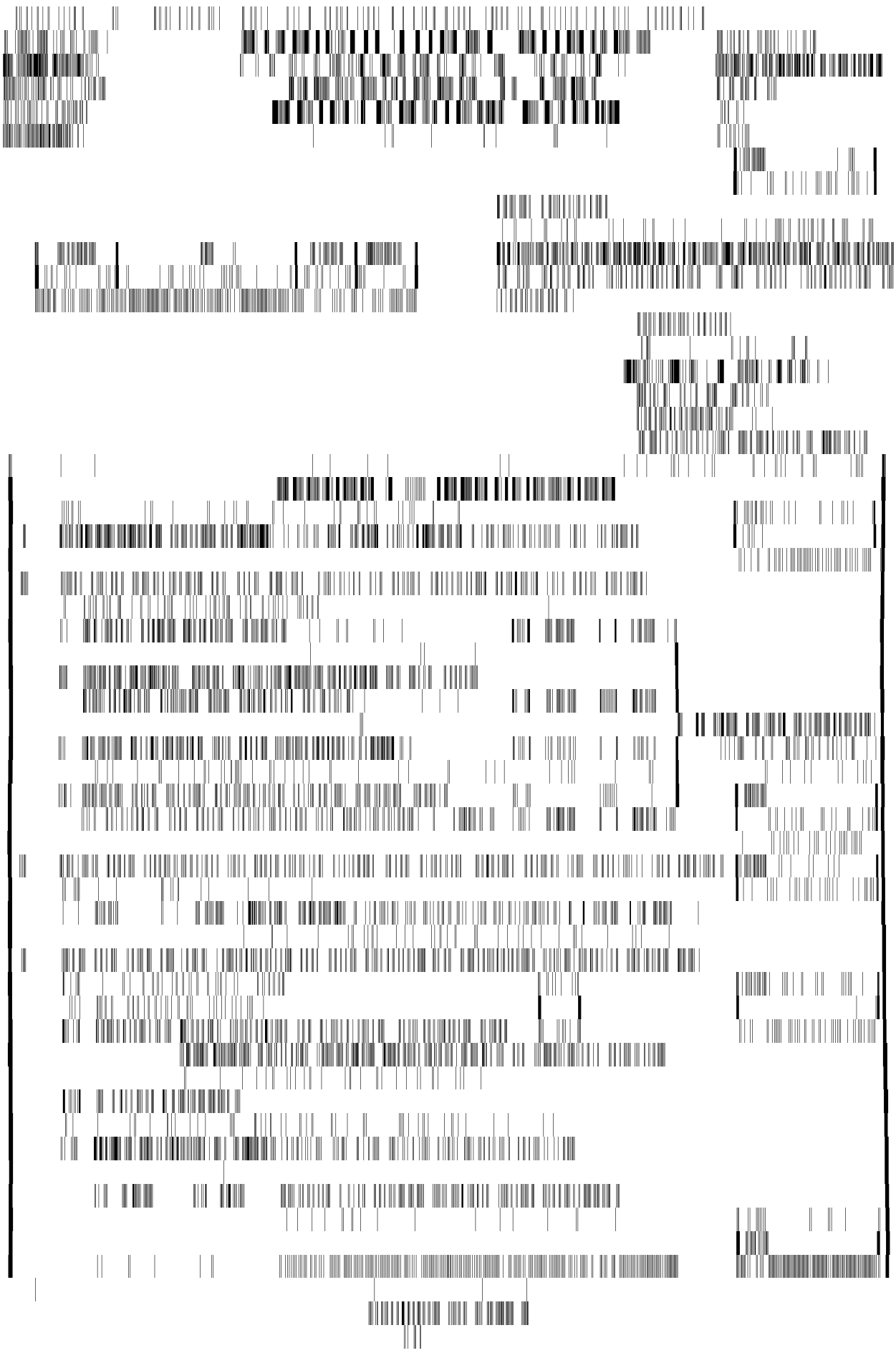
(SOS) Survey report recommendations was written by the Statistical Standards Staff (SSS), NASS, on March 8. Available from the author, or the SSS.

Steiner, Mike (1980) "Study of Hog Balance Sheet and Alternative Ways to Ask Farrowing Questions," Working Paper, National Agricultural Statistics Service, U.S. Department of Agriculture.

Tortora, Robert D. (1987) "Quantifying Nonsampling Errors and Bias," Journal of Official Statistics, Vol. 3, No. 4, pp 339-342.

Tortora, Robert D. (1988) "Statistical Process Quality Control Systems in NASS," Minutes of the September 1988 Program Planning Committee Meeting, National Agricultural Statistics Service, Washington, D.C.

U.S. Census Bureau, (1987) Reinterview Manual, Available from Brad Pafford, Survey Research Branch, National Agricultural Statistics Service, Washington, D.C.



SECTION 2 — ACRES OPERATED

1

Acres

1. How many **total acres of land** were in this operation on **December 1**? 900

Include: Farmstead, all cropland, woodland, pastureland, wasteland, government program land, all land owned, rented or managed.

Exclude: Land rented to others and all grazing land used on an AUM *(fee per head)* basis.

2. Of the total acres in this operation, how many acres would be considered cropland *(Include land in hay and cropland in government programs)*? 802

Continue On Next Page

SECTION 5 — CROPS

The purpose of the next questions is to collect information about crops for the 1987 crop year.

How to complete this section:
- Report for all the land you operate, including land rented from others.
- If harvest is not complete, make your best estimate of acres and total production.
- Production is equal to acres harvested and to be harvested times average yield per acre.
- Total production should include the landlord's share.

1. The following information is needed for crops harvested during 1987.

CORN *(exclude popcorn and sweet corn)*

Acres planted for all purposes	[illegible]	ac.
Acres harvested and to be harvested for grain and seed	[illegible]	ac.
Total grain and seed production	[illegible]	bu.

SOYBEANS

Acres planted for all purposes	[illegible]	ac.
Acres harvested and to be harvested for beans	[illegible]	ac.
Total production	[illegible]	bu.

SUGARBEETS

Acres planted	[illegible]	ac.

HAY CROPS

Count each acre only once regardless of number of cuttings or different uses.
Dry weight for any dehydrated hay should be included.

ALFALFA HAY AND ALFALFA MIXTURES

Acres cut for dry hay *(exclude haylage and green chop)*	[illegible]	ac.

ALL OTHER HAY *(include small grain cut for dry hay, clover, timothy, clover and grass mixtures, lespedeza, peanut, creme, coastal bermuda, sudan, sudan crosses, millet, other (tame and wild hay.)*

Acres cut for dry hay *(exclude haylage and green chop)*	[illegible]	ac.

TOBACCO

Acres harvested	[illegible]	ac.

2. For the following small grains, please report seedings for the 1988 crop year.

WINTER WHEAT acres seeded and to be seeded for all purposes	[illegible]	ac.

Continue on Next Page

Crops Incomp.	1 : HAS 2 : OPR 3 : OR	[illegible]
Comp.	4	

SECTION 4 — GRAINS AND SOYBEANS IN STORAGE

2

Please account for whole grains and soybeans on hand or stored **December 1** on the total acres operated, whether for feed, seed, or sale. They may have belonged to you or someone else, or been stored under a government program *(loan, farmer owned reserved, or CCC).*

	NO	YES		
1. On December 1, was any **whole grain corn** on hand or stored on the total acres operated?	☐	☐	How many bu?.....	121
Were any **soybeans** on hand or stored on these acres?	☐	☐	How many bu?	125
What about **wheat**, including all types? *(winter, durum and spring)*	☐	☐	How many bu?.....	126

UNHARVESTED CORN AND SOYBEANS

2. On December 1, did you have any corn or soybeans still in the field that you **intend to harvest** for grain or beans?

☐ **YES** ☐ **NO** - *Go to Item 4 below.*

3. Was this **unharvested production** included with corn and soybeans in storage?

☐ **YES = 1** - *Enter code in Code Box 460, then go to Item 4 below.*

☐ **NO = 3** - *Enter code in Code Box 460, then continue.*

.....**enter code** 460

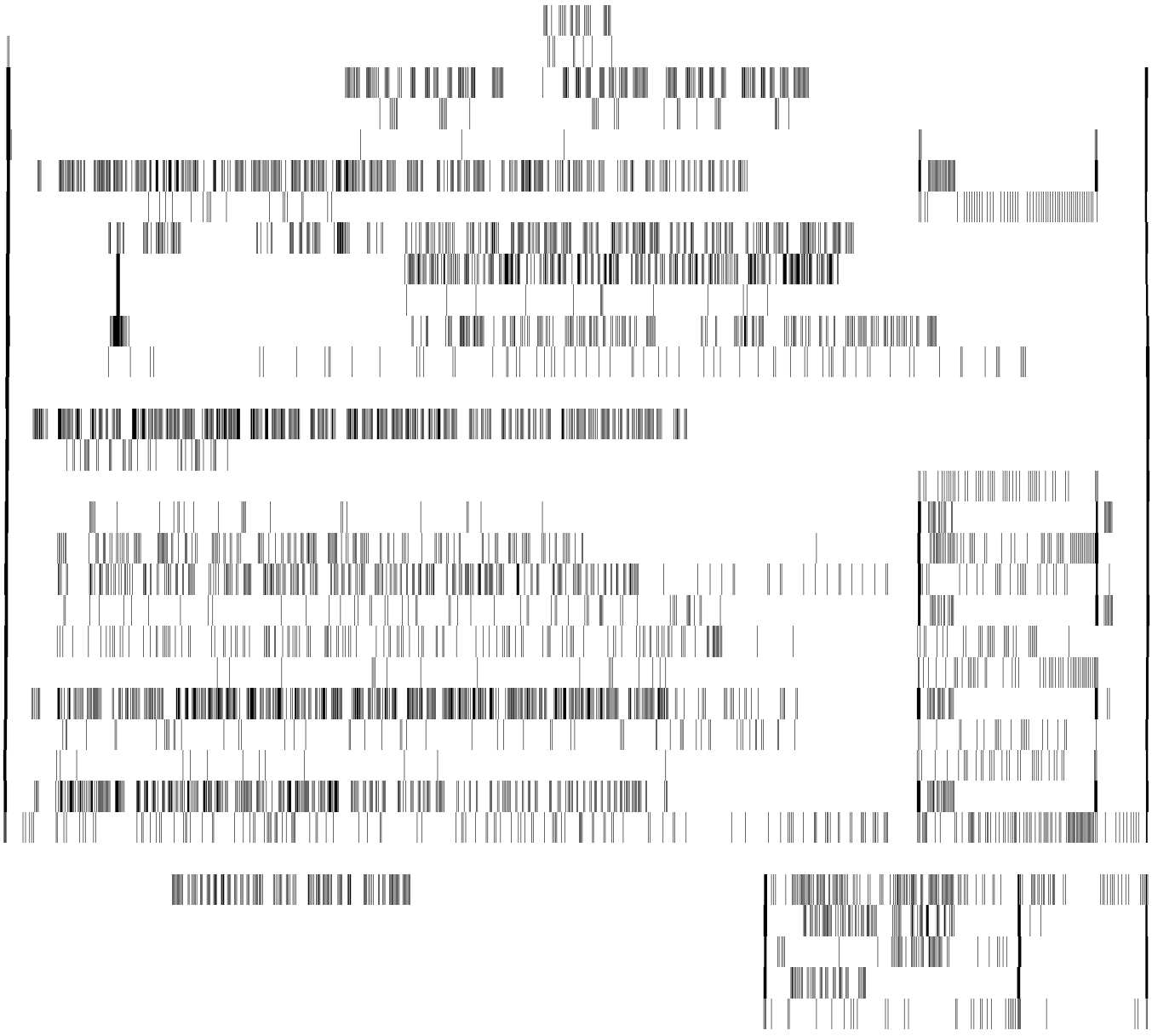
Corn	Acres remaining to be harvested.....ac	559
	AND	
	Expected yield per acre..... bu/ac	573
Soybeans	Acres remaining to be harvested.....ac	575
	AND	
	Expected yield per acre..... bu/ac	594

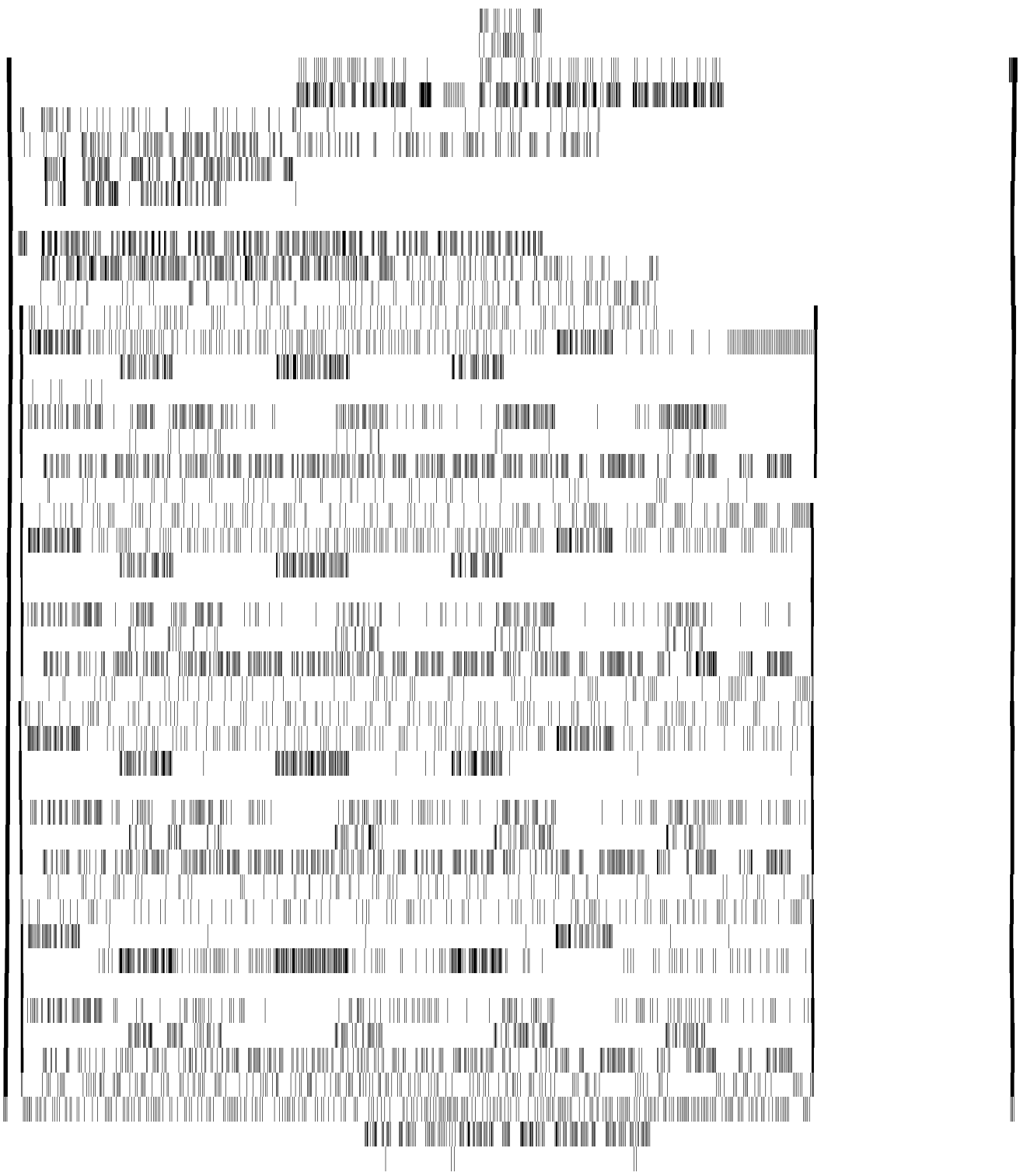
GRAIN STORAGE CAPACITY

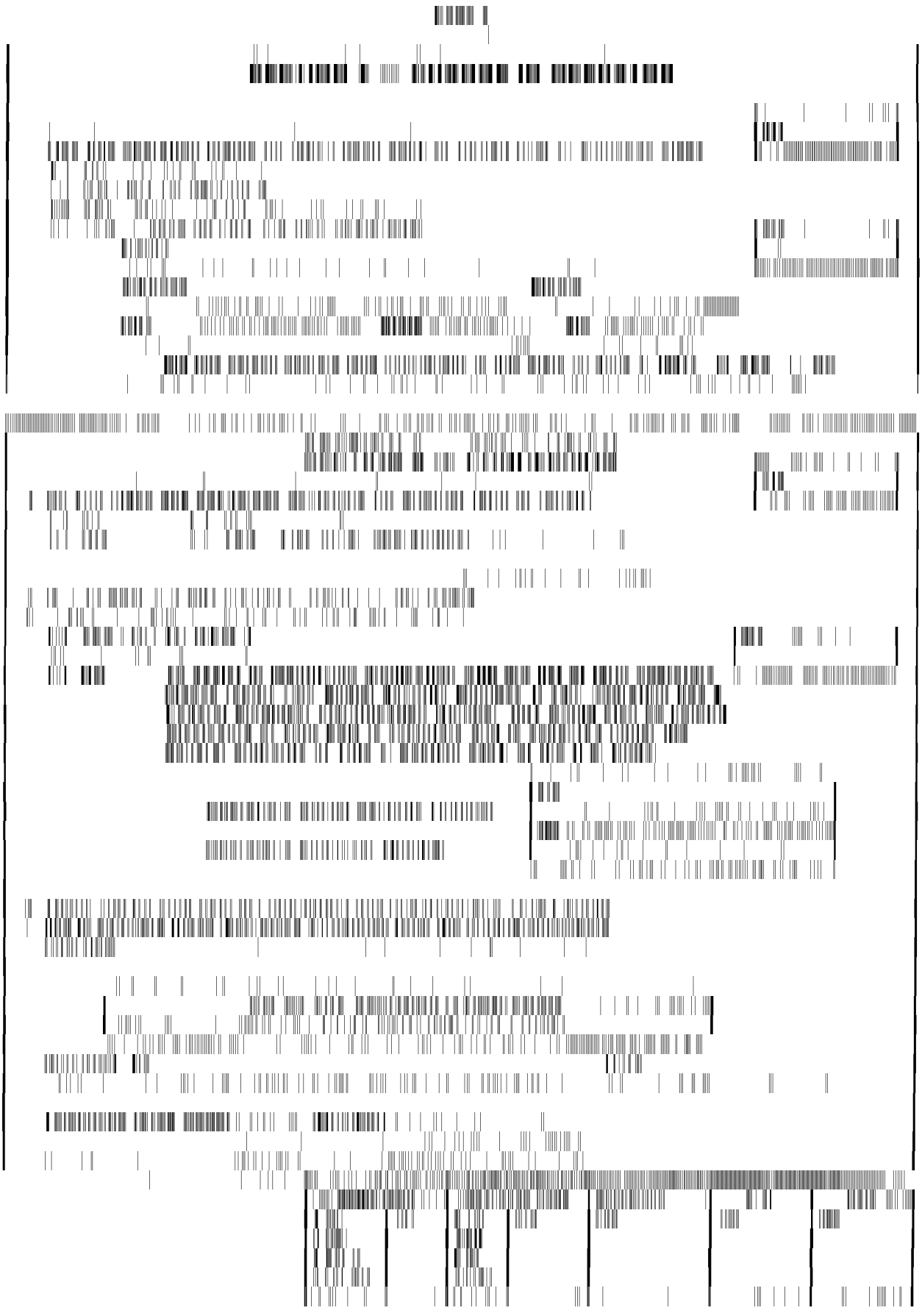
4. On December 1, what was the **TOTAL STORAGE CAPACITY** of all the bins, cribs, sheds, and other structures normally used to store whole grains or oilseeds on the total acres operated? .....bushels 808

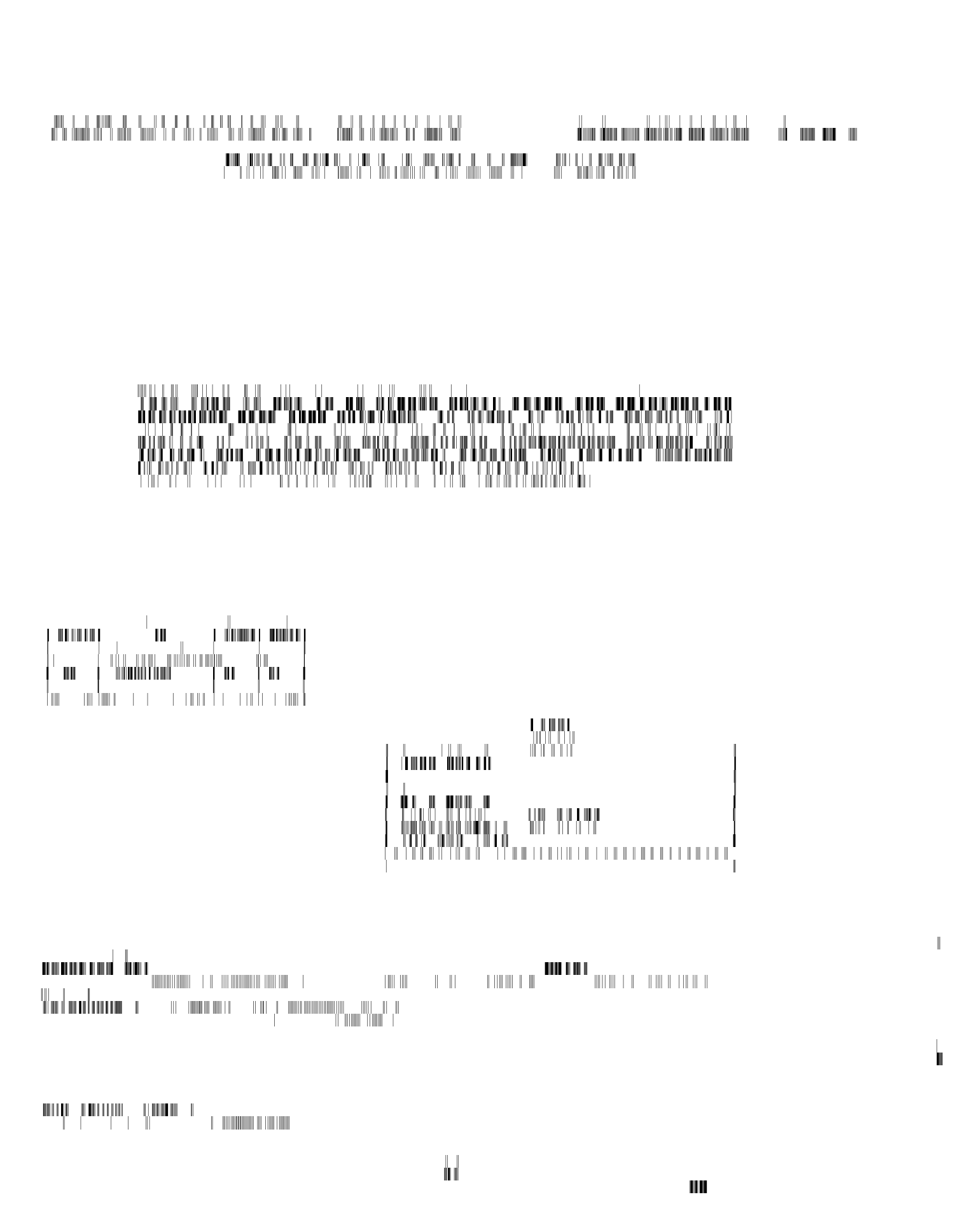
Continue On Next Page

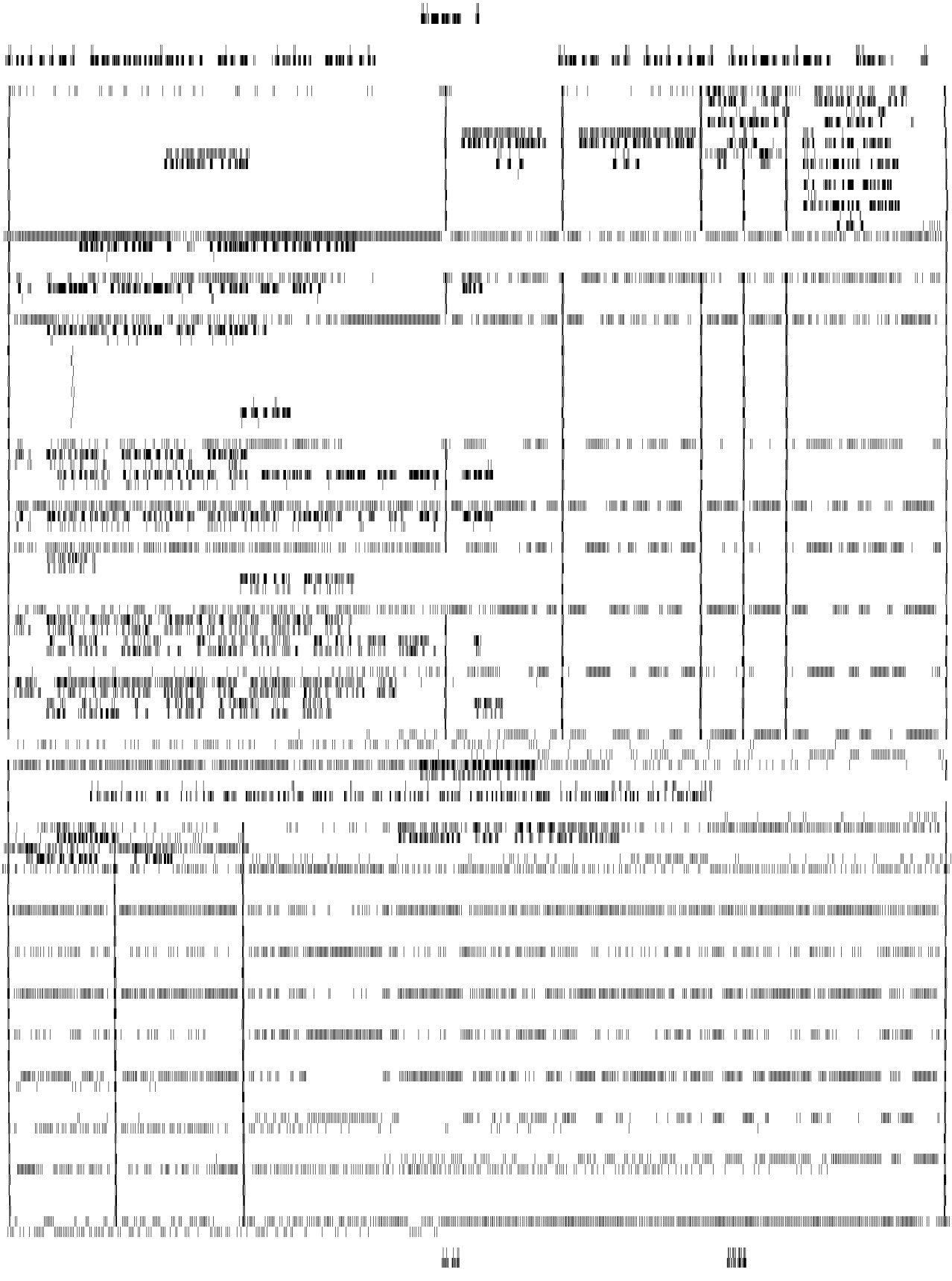
Stocks: Incomp.	1 - Has 2 - Unk 3 - No	141
Comp.	4	

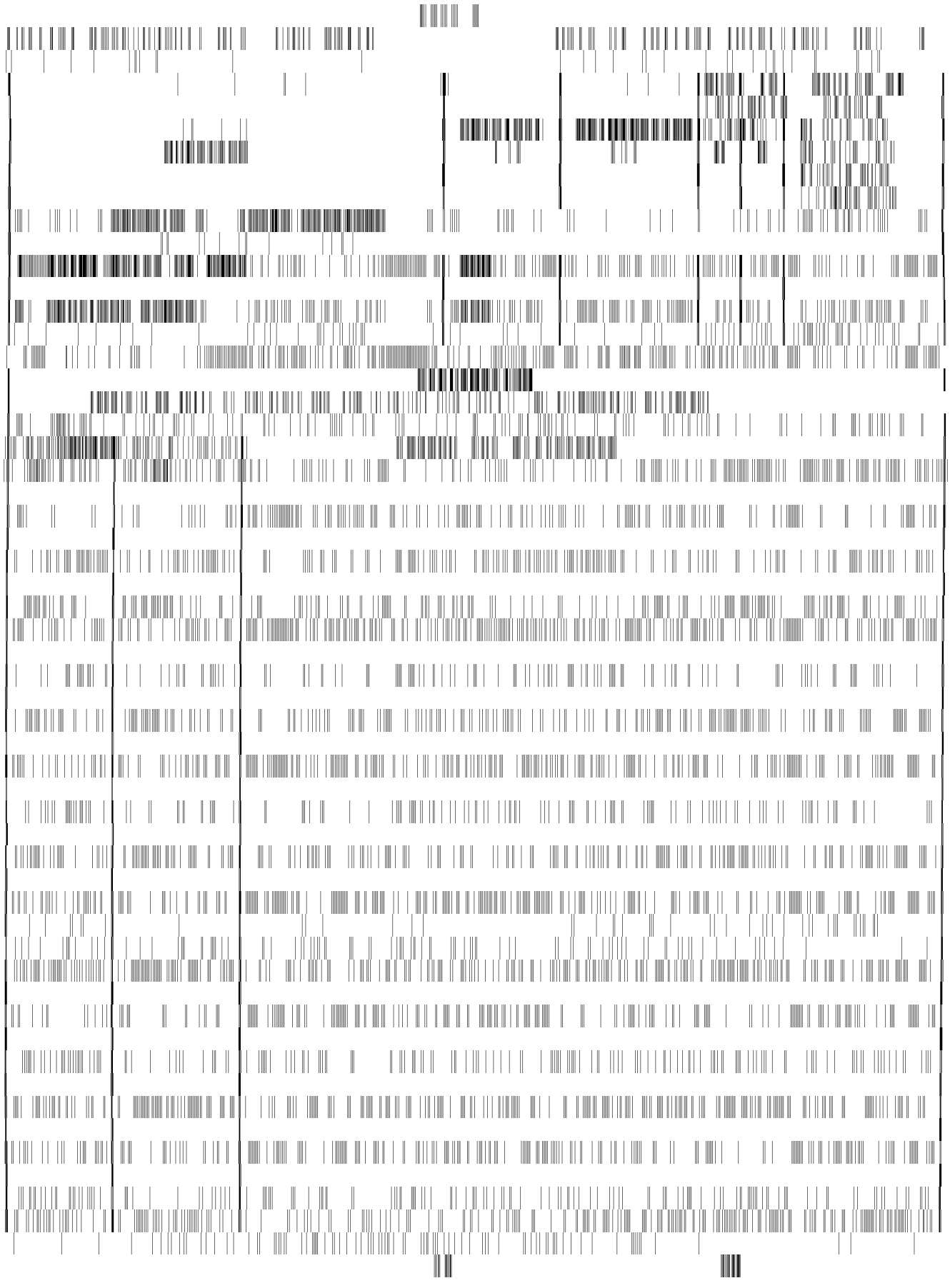


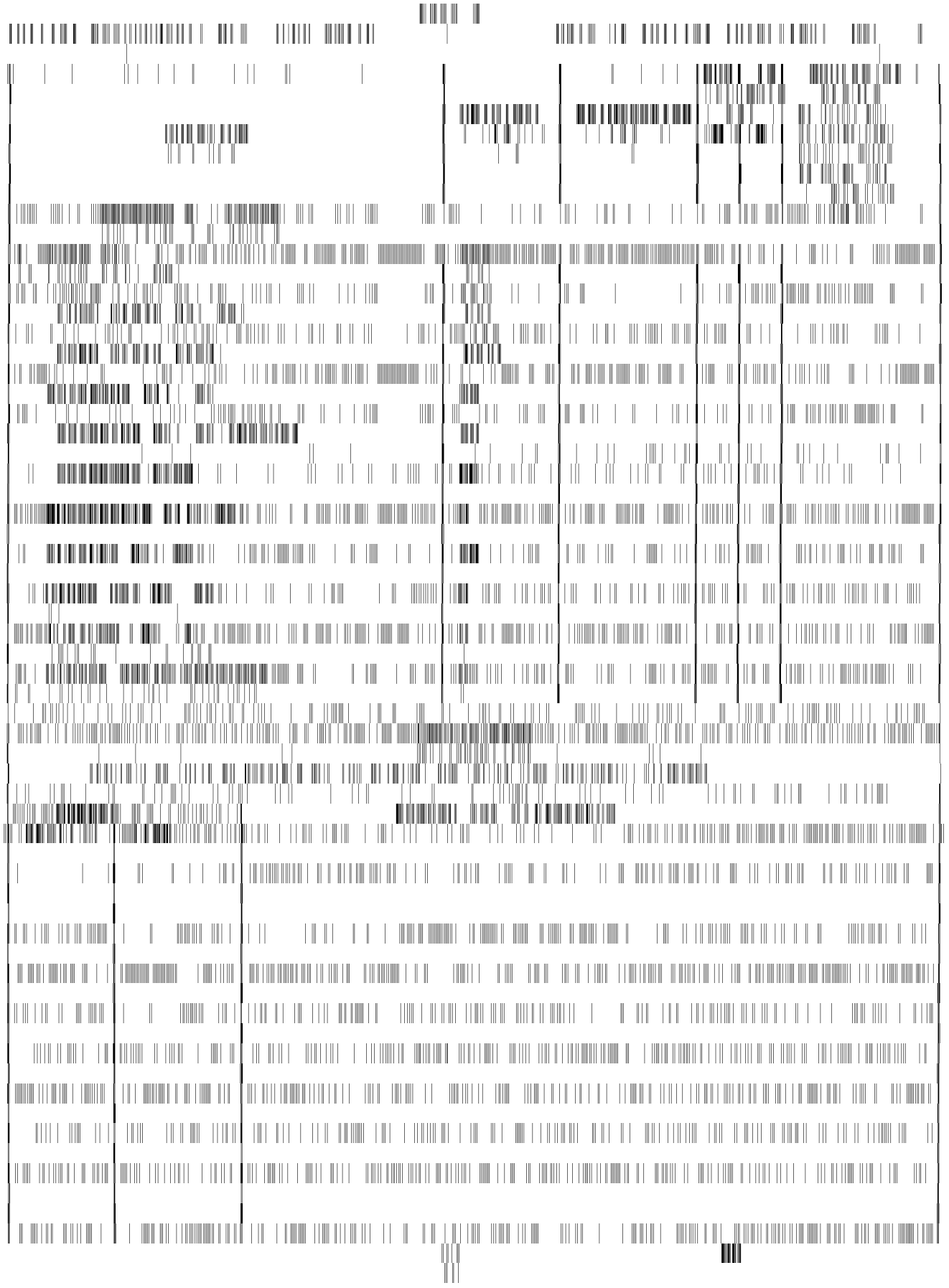


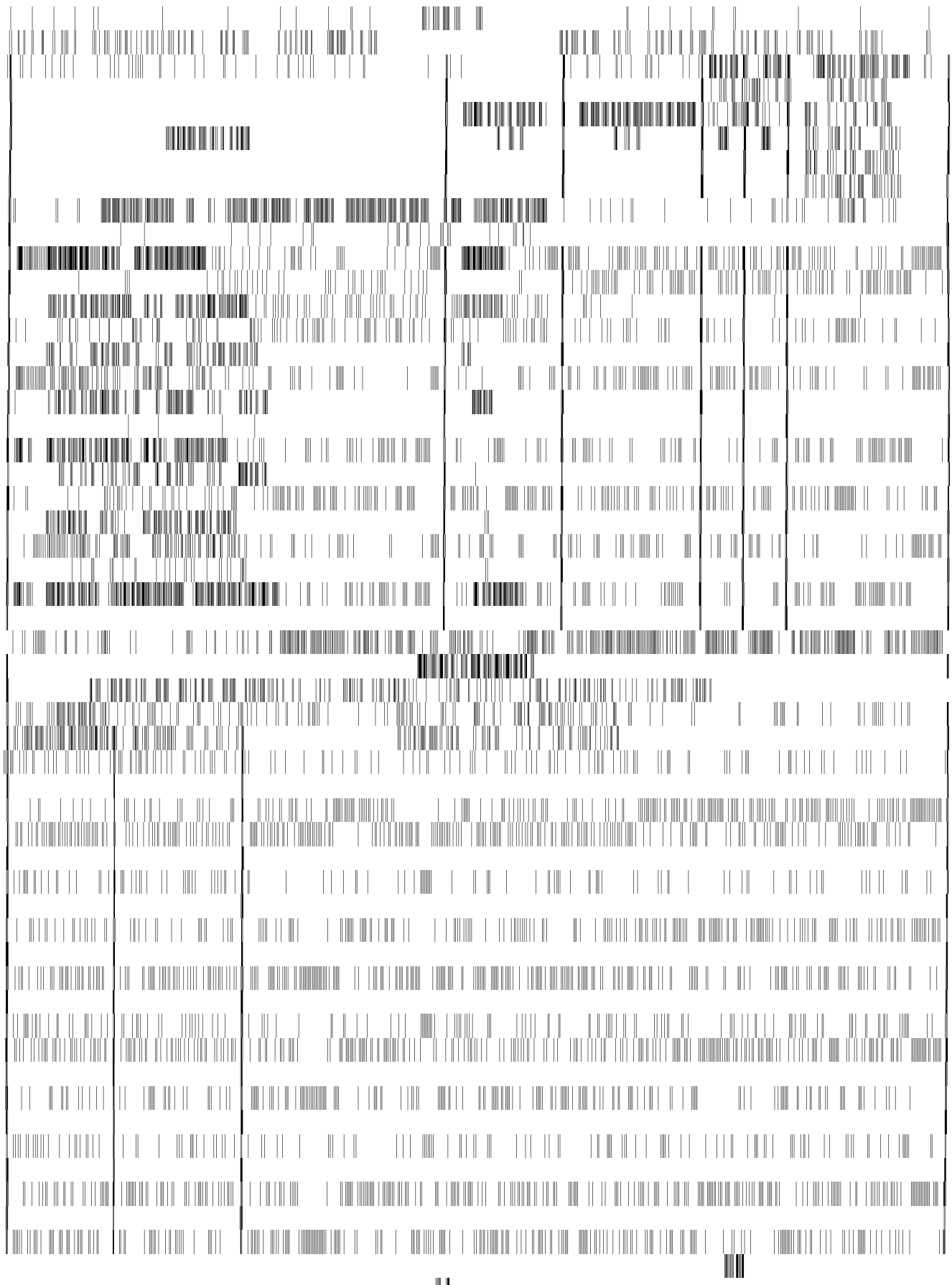


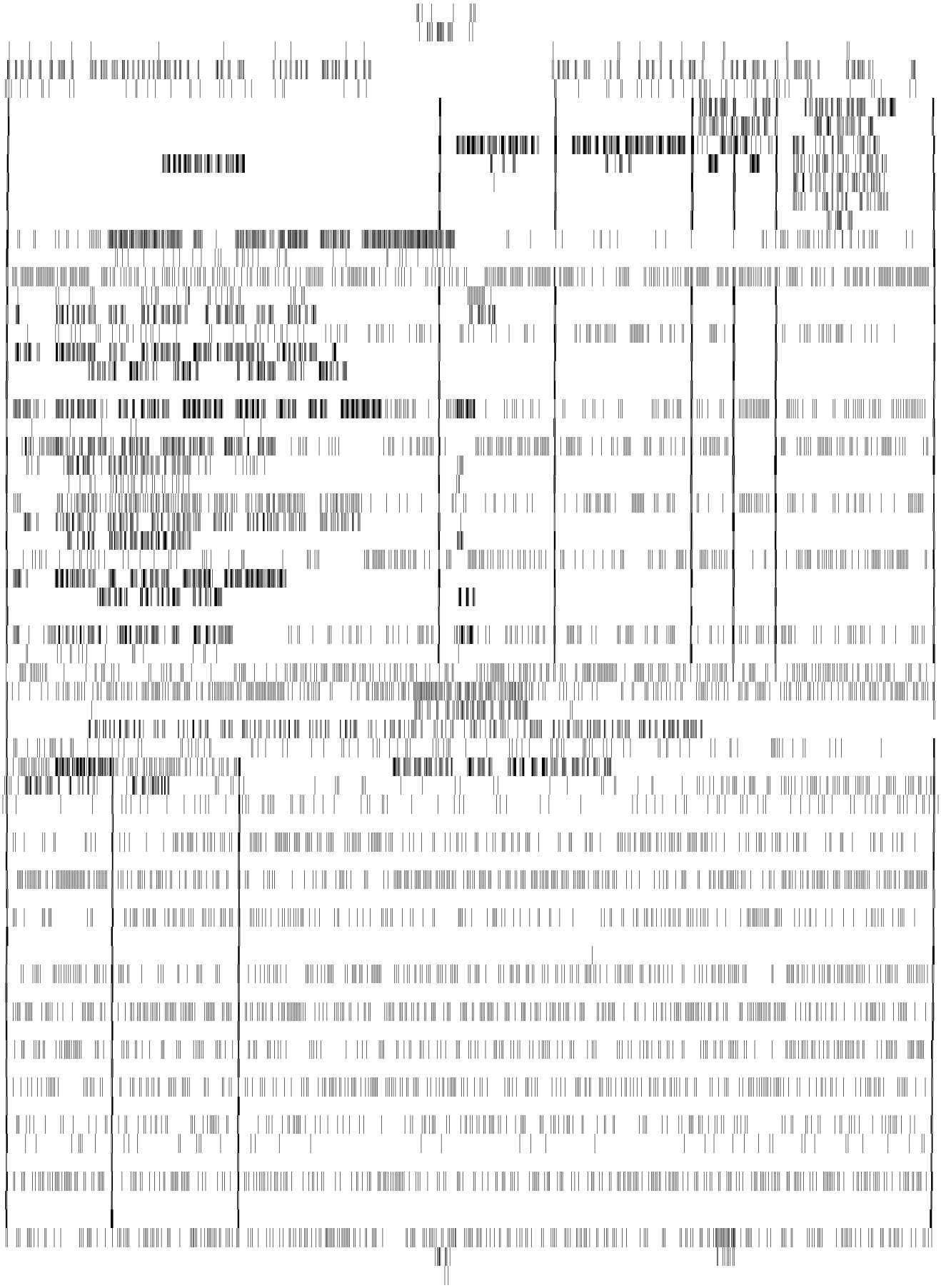


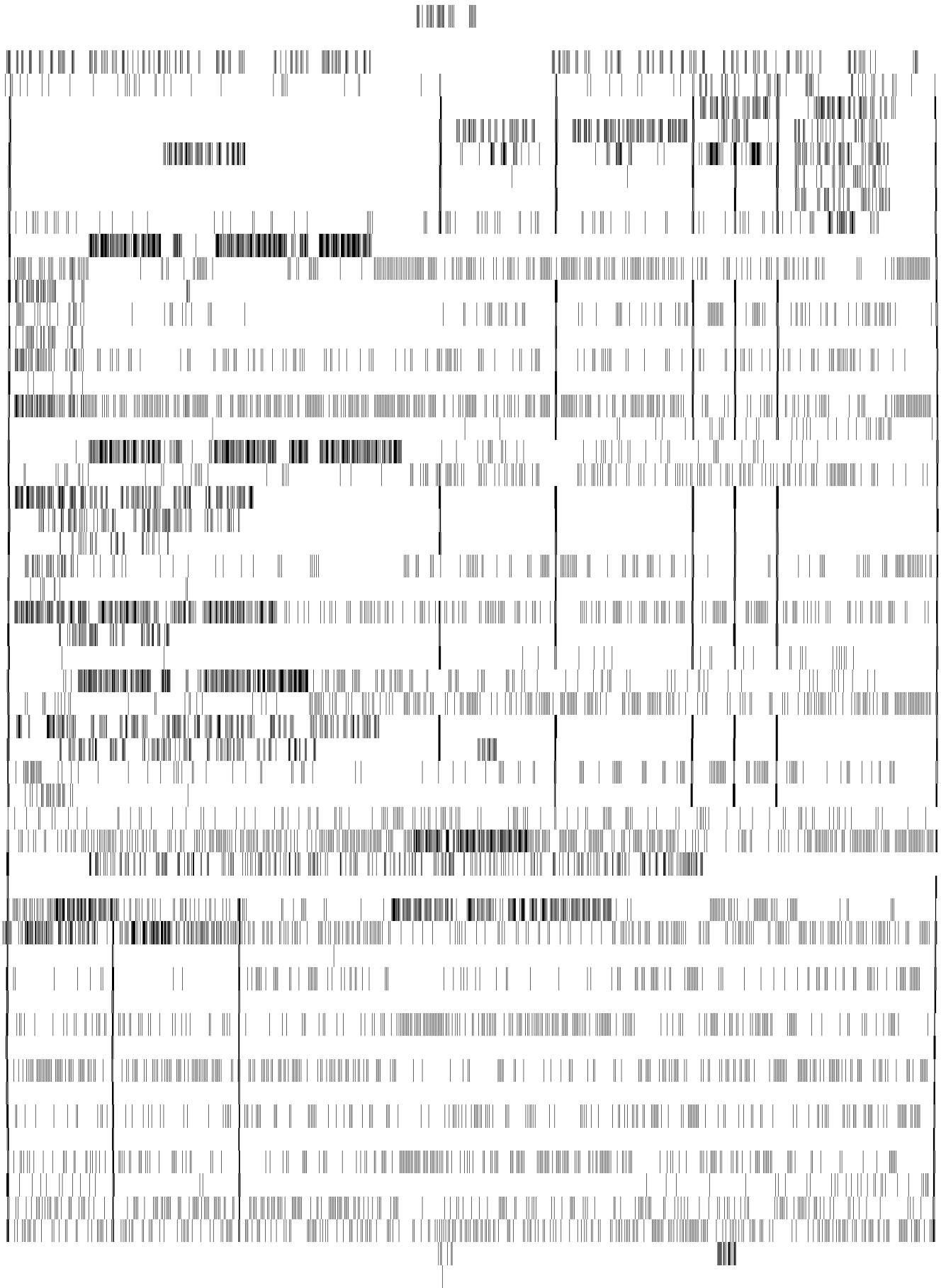


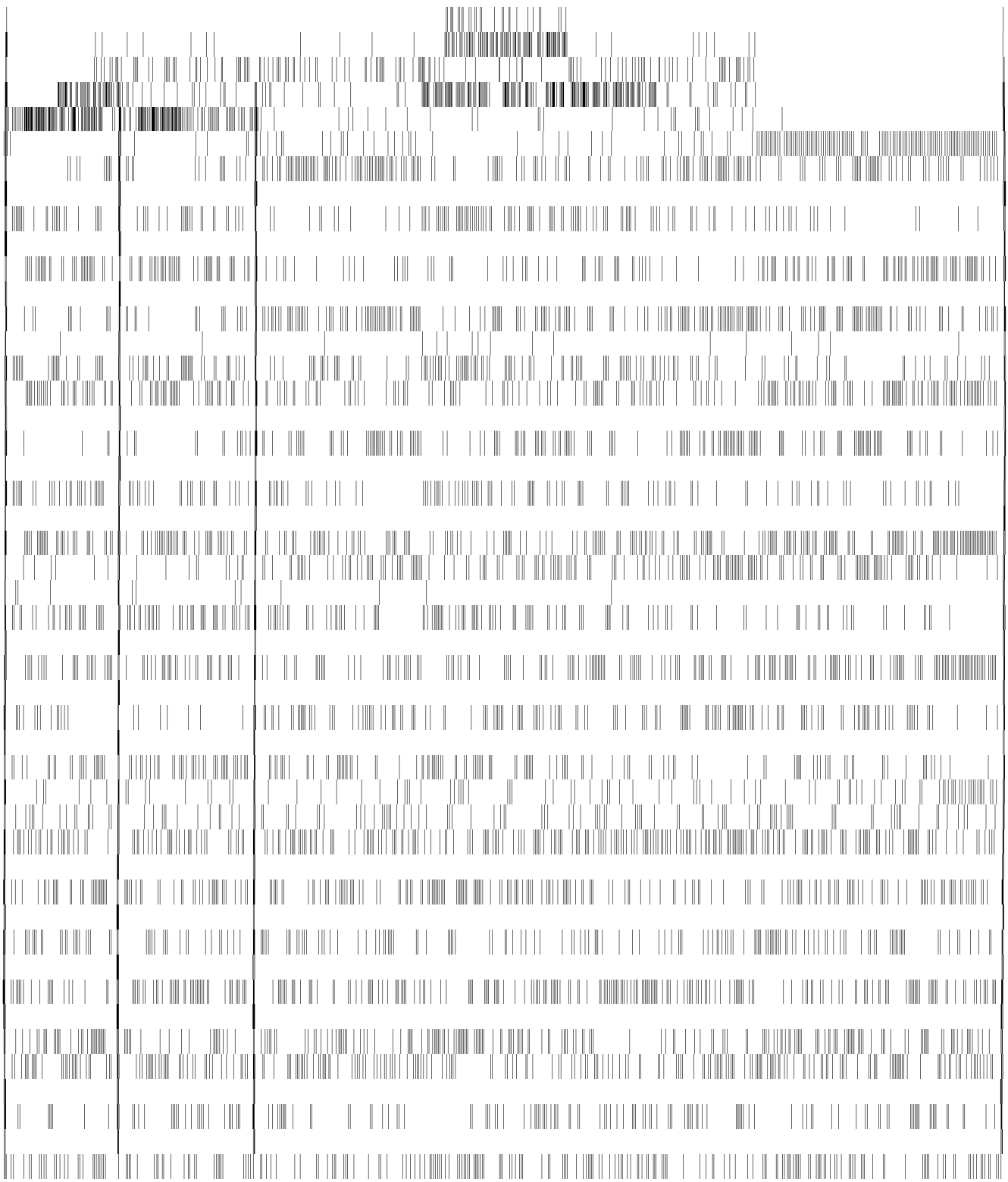












APPENDIX B

Estimation of Bias - CATI Minus Reconciled Response

Let, Y_{ij} be the ij-th observation for the CATI sample, where i = 1, 2, ...,L; j = 1, 2, ..., n_i; i = the stratum identification, and j is the element within stratum identification.

Let, V_{ij} be the ij-th observation for the reinterview sample, where i = 1, 2, ...,L; j = 1, 2, ..., n_i; i = the stratum identification, and j is the element within stratum identification.

Then, let T_{ij} be the truth measure,

$$T_{ij} = o_{ij}\ p_{ij}\ Y_{ij} + q_{ij}\ r_{ij}\ V_{ij} + C_{ij},$$

where,

$$o_{ij} = \begin{cases} 1 & \text{if operate in the CATI interview} \\ 0 & \text{otherwise (out-of-business),} \end{cases}$$

$$q_{ij} = \begin{cases} 1 & \text{if operate in the Reinterview} \\ 0 & \text{otherwise (out-of-business),} \end{cases}$$

$$p_{ij} = \begin{cases} 1 & \text{if the CATI response was correct,} \\ 0.5 & \text{if said "both estimates - unsure which is more correct"} \\ 0 & \text{otherwise} \end{cases}$$

$$r_{ij} = \begin{cases} 1 & \text{if the reinterview response was correct, and not equal to the CATI response,} \\ 0.5 & \text{if said "both estimates - unsure which is more correct"} \\ 0 & \text{otherwise} \end{cases}$$

$$C_{ij} = \begin{cases} \text{compromise value} & \text{if one was reported} \\ 0 & \text{otherwise} \end{cases}$$

Then, the bias is, $B_{ij} = Y_{ij} - T_{ij}$, and estimated by

$$\hat{B} = \sum_{i=1}^{L} \sum_{j=1}^{n_i} w_{ij}\ z_{ij}\ B_{ij},$$

where, w_{ij} is the weight for the ij-th observation (one divided by the sampling fraction, adjusted for subsampling), and z_{ij} is the list adjustment factor. More specifically:

$$z_{ij} = \begin{cases} 1 & \text{if LAF=1 in the SAS file and no difference uncovered with the operation description in the reinterview,} \\ 1 & \text{if LAF not equal 1 in the SAS file, yet information found in the reinterview that would require LAF=1,} \\ 0<z_{ij}<1 & \text{if 0<LAF<1 in the SAS file and no difference uncovered with the operation description in the reinterview,} \\ 0<z_{ij}<1 & \text{if a difference uncovered with the operation description found in reinterview which would require the LAF to be between 0 and 1,} \\ 0 & \text{if LAF=0 in the SAS file and no difference uncovered with the operation description in the reinterview, or} \\ 0 & \text{if LAF not equal to zero in SAS, yet the operation is out of business or other information available to set LAF=0.} \end{cases}$$

The estimated variance of $\hat{B}$ is,

$$\hat{V}(\hat{B}) = \sum_{i=1}^{L} (n_i - 1)^{-1} n_i (1 - f_i) \sum_{j=1}^{n_i} (d_{ij} - d_{i.})^2 ,$$

where

$$d_{ij} = w_{ij} z_{ij} B_{ij} ,$$

$$d_{i.} = n_i^{-1} \sum_{j=1}^{n_i} d_{ij} ,$$

f_i is the sampling rate for the i-th stratum, n_i is the number of responses in stratum i, and z_{ij} and w_{ij} defined as before.

Estimation of Bias - SAS Edited Minus Reconciled Response

Let, Y_{ij} be the ij-th observation for the SAS edited response, and V_{ij} be the ij-th observation for reinterview response, where i = 1, 2, ...,L; j = 1, 2, ..., n_i; i = the stratum identification, and j is the element within stratum identification.

As before, let T_{ij} be the truth measure,

$$T_{ij} = o_{ij}\ p_{ij}\ Y_{ij} + q_{ij}\ r_{ij}\ V_{ij} + C_{ij} ,$$

where, o_{ij}, q_{ij}, p_{ij}, r_{ij}, C_{ij} are defined as before.

Again, the bias is, $B_{ij} = Y_{ij} - T_{ij}$, and estimated by $\hat{B}$, with variance estimated by $V(\hat{B})$.

Domain Estimates of the Bias - CATI Minus Reconciled Response

Let, Y_{ij}, T_{ij}, B_{ij}, z_{ij}, and w_{ij} be given as above. Now, let,

$$B_{ij}' = \begin{cases} B_{ij} & \text{if the unit is in the k-th domain} \\ 0 & \text{otherwise.} \end{cases}$$

Then the population total for the k-th domain of the B_{ij}' is estimated by

$$\hat{B}_k = \sum_{i=1}^{L} \sum_{j=1}^{n_i} w_{ij}\ z_{ij}\ B_{ij}'$$

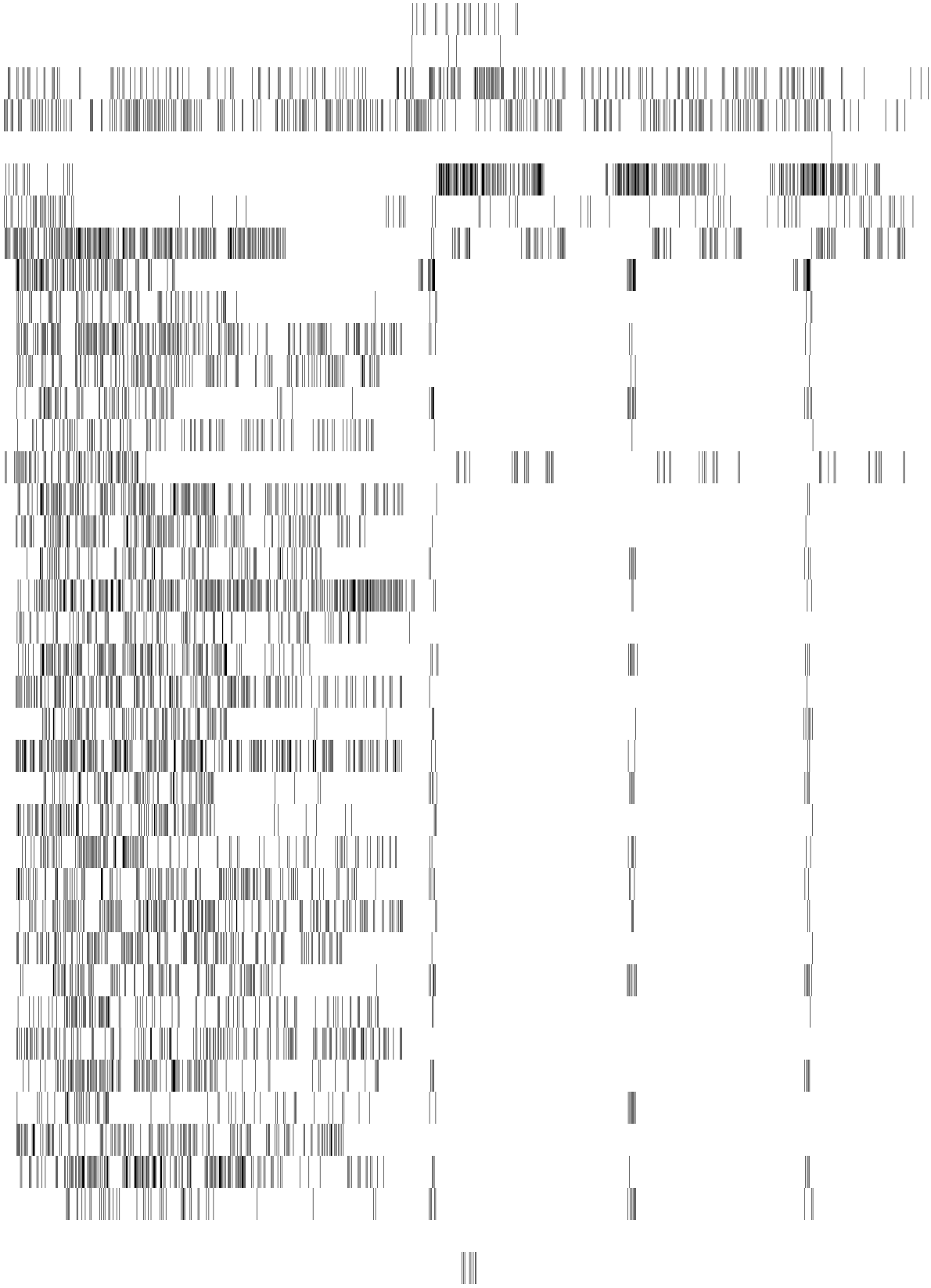
The estimated variance of $\hat{B}_k$ is,

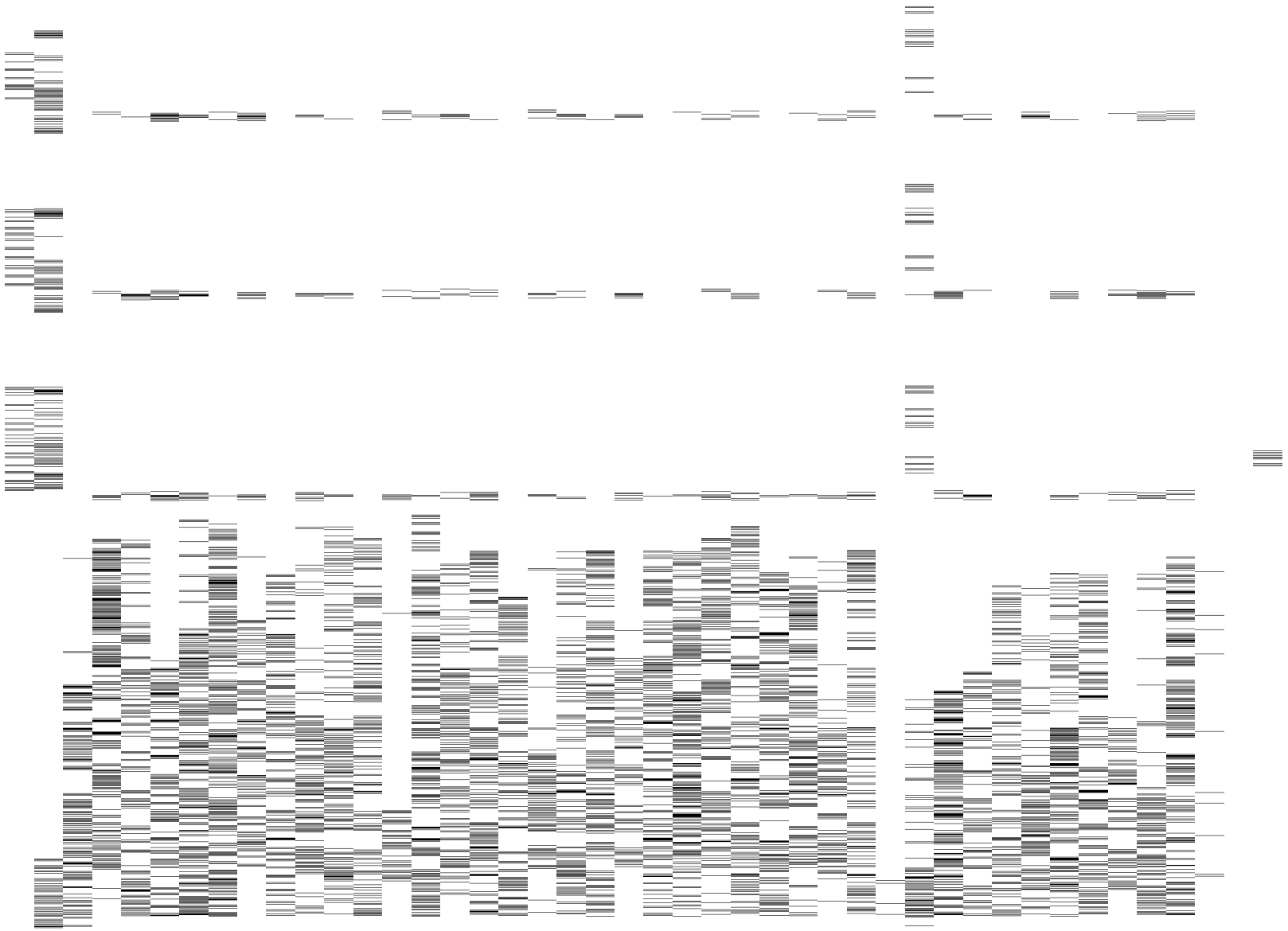
$$V(\hat{B}_k) = \sum_{i=1}^{L} (n_i - 1)^{-1}\ n_i\ (1 - f_i) \sum_{j=1}^{n_i} (d_{ij}' - d_{i.}')^2 ,$$

where

$$d_{ij}' = w_{ij}\ z_{ij}\ B_{ij}' ,$$

$$d_{i.}' = n_i^{-1} \sum_{j=1}^{n_i} d_{ij}'$$





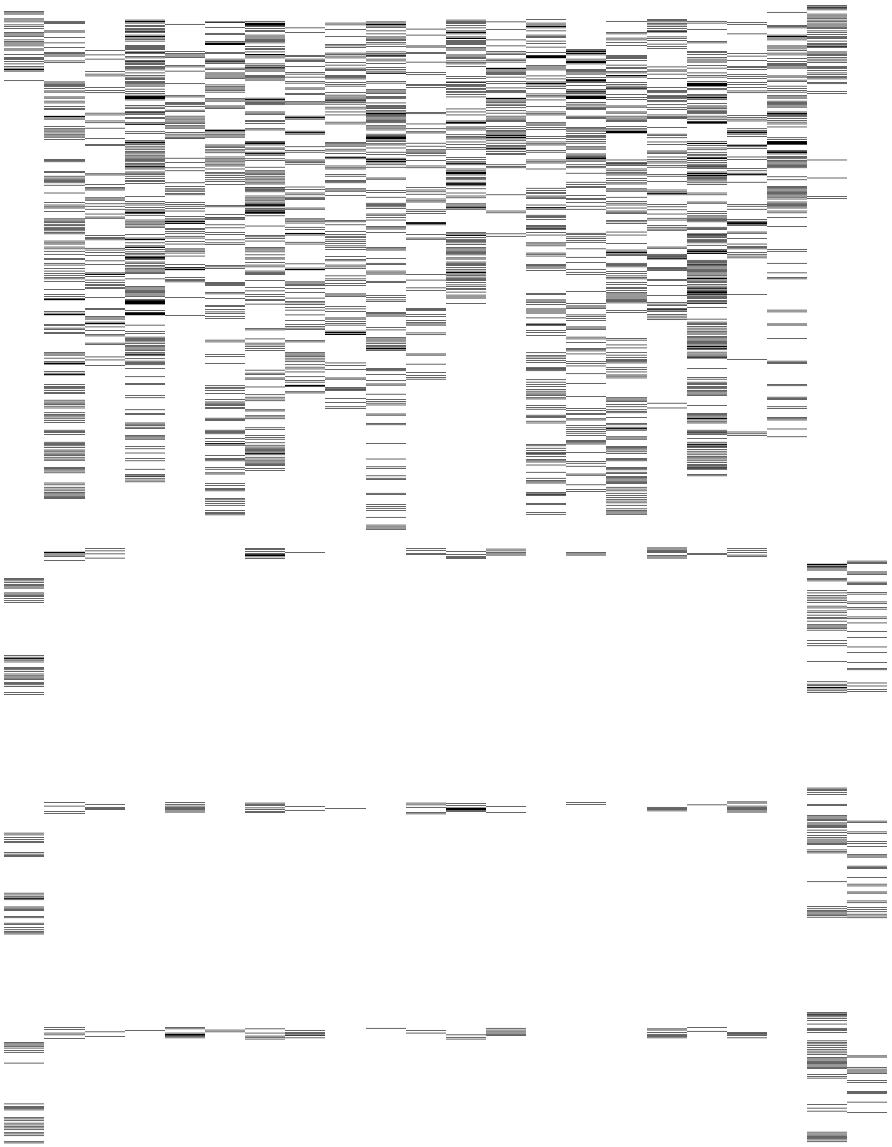
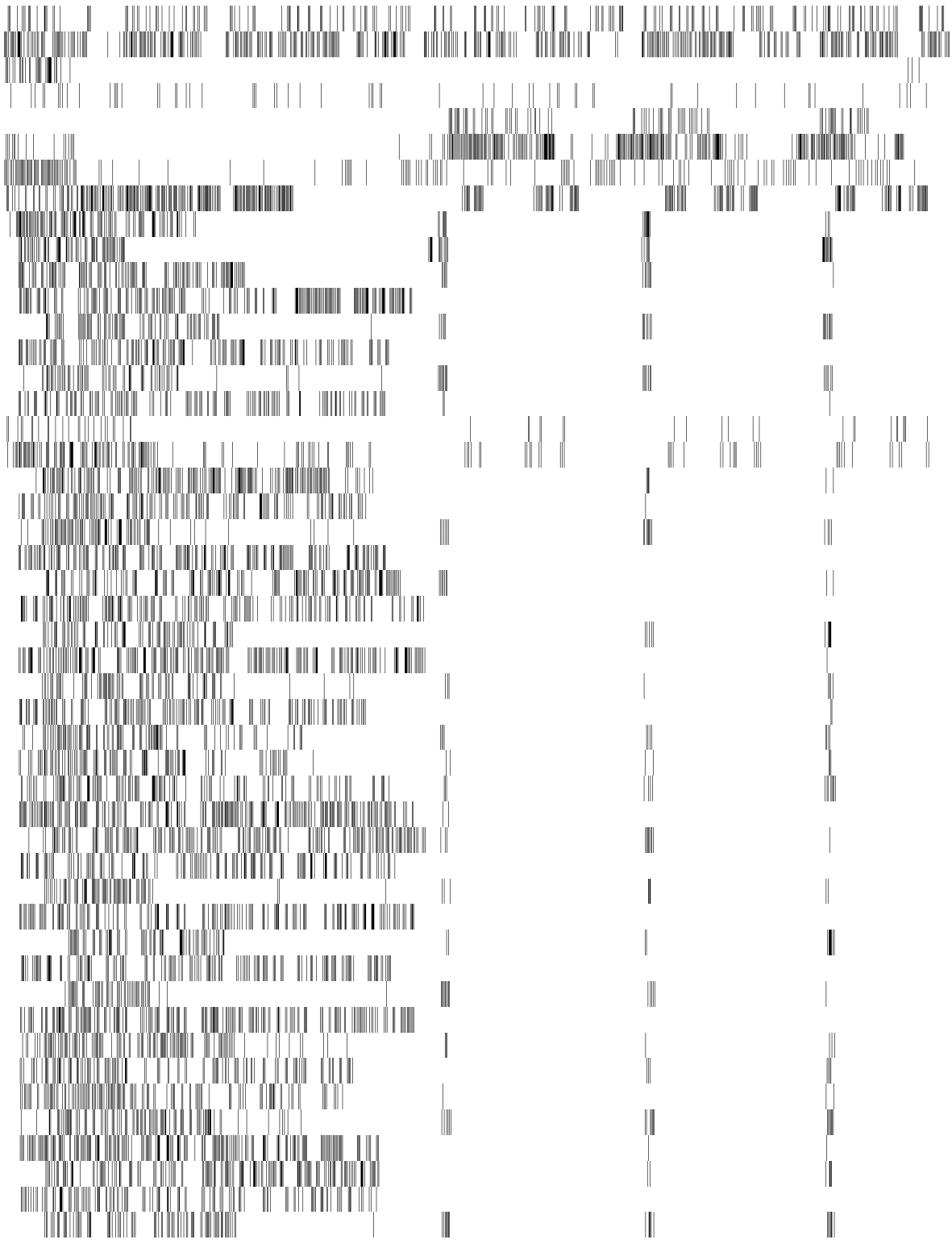


Table 2. Reasons for differences in the CATI and reinterview responses for soybean stocks (responses with positive bias) - December 1987 reinterview study.

Reason	Minnesota Freq.	Minnesota %	Indiana Freq.	Indiana %	Ohio Freq.	Ohio %
ESTIMATED/ROUNDING ERROR	21	42.9	14	36.8	19	50.0
Rounding error	2		1		2	
Estimated	9		5		8	
From actual records	5		3		-	
Both estimated - don't know which is more correct	3		-		3	
Both estimated- 2nd response is more correct	1		4		6	
Difference of only small amount	1		1		-	
DEFINITIONAL	18	36.7	12	31.5	6	15.8
Original respondent reported indiv. land not partnership land	1		-		-	
Confusion with storage on acres oper. vs. where he lives	2		-		1	
Original interviewer asked for Dales land, not for farm	-		1		-	
Included some Indiv. land with Partnership land	-		1		-	
Forgot to include	2		-		-	
Failed to report reserve corn	-		1		-	
Failed to include grain belonging to someone else	1		1		-	
Confused with reporting that that stored in town	4		4		1	
Confusion with reporting landlord grain on farm	-		1		1	
This operation originally coded out of business	2		-		-	
Confusion with reporting rented land	-		1		-	
Didn't ask/forgot to include last year's crop	2		1		-	
Forgot to include grain purchased for feed/seed	1		-		3	
Didn't include storage used by another	1		1		-	
Reference date problem	2		-		-	
OTHER	10	20.4	12	31.6	13	34.2
Doesn't give out information over the phone	-		1		-	
Spouse didn't know	2		-		1	

Reason	Minnesota Freq.	%	Indiana Freq.	%	Ohio Freq.	%
OTHER (cont'd)						
Partner that reported only familiar with livestock	-		-		1	
Don't know why there's a diff.	1		1		1	
Thought he/she reported this the first time	2		-		2	
Can't hear well on phone	-		3		-	
Misunderstanding between enum. and respondent	-		1		-	
Fatigued/tired/hurried when called on phone	-		1		-	
No explanation given	3		1		2	
Forgot to ask	-		-		2	
Didn't take the time on telephone to add correctly	-		2		1	
Compromise values given	-		2		3	
Not able to reconcile difference	2		-		-	
TOTAL	49	100	38	100	38	100



Reason	Minnesota Freq.	%	Indiana Freq.	%	Ohio Freq.	%
DEFINITIONAL (cont'd)						
Forgot to about oats storage	4		-		-	
Didn't think about bins being filled to the top	-		1		-	
Confusion with including bins & cribs no longer used	-		-		6	
Do not have any storage	-		1		-	
Confusion with listed capacity versus usable capacity	-		1		-	
Failed to include grain belonging to someone else	1		1		-	
Confused with reporting that that stored in town	-		1		-	
Confusion with reporting rented land	1		1		-	
Included rented storage on other farms not his	1		-		-	
Not all bins included	11		7		3	
Rented bins not included	-		-		1	
Storage is silage not whole grain	-		-		1	
Didn't ask about ear corn/ only reported shelled corn	-		1		-	
Confusion with reporting high moisture corn	2		1		-	
Forgot to include grain purchased for feed/seed	-		1		-	
Didn't include ear corn	-		1		-	
Didn't include storage used by another	1		-		-	
Included silos used to store silage	1		-		-	
Only included stocks of one partner-the one interviewed	1		-		-	
Reported that in use, not total capacity	2		-		-	
Confusion with storage on acres operated versus where lives	1		1		-	
This operation originally coded as out-of-business	1		-		-	
Included cribs no longer usable	1		-		-	
OTHER	18	20.0	27	32.1	18	29.5
No explanation given	2		4		3	
Wrong bushels given	-		-		1	
Misunderstanding between enum. and respondent	1		-			
Mistake in addition	2		3		2	
Figured the correct value after the telephoner called	1		-		-	
Don't know why there's a diff.	5		2		2	

[illegible] [illegible] [illegible] [