RESULTS FROM A FEASIBILITY TEST OF AN ALTERNATIVE AUTOMATED DATA CAPTURE METHODOLOGY DURING THE 1992 CENSUS OF AGRICULTURE

Magdalena Ramos, Bruce Hughes, and Betty McKay, U.S. Bureau of the Census
Magdalena Ramos, Bureau of the Census, ESMPD, Room 3112-4, Washington D.C. 20233

Key Words: Data Entry, Keying, Mode Effects, Mark Sense Reader

INTRODUCTION--The Census Bureau conducts the Census of Agriculture every five years, in years ending in 2 and 7. The purpose of the census of agriculture is to collect and publish information on land in farms, operator characteristics, and agricultural production and sales by farms in the United States. A census farm is defined as any place from which $1,000 or more of agricultural products were sold (or had the potential to be sold) during the census year. For the 1992 Census of Agriculture approximately 3.5 million addresses were mailed a census form, 2.9 million addressees returned a completed form and approximately 1.8 million forms were data keyed. Because of the large volume of records needing keying, data entry is one of the more expensive and labor intensive processing operations for the census of agriculture.

Regular production keying for the 1992 Census of Agriculture consisted of the conventional data keying procedures from a paper form to a computer terminal. During the 1992 Census of Agriculture, the Census Bureau tested an alternative to the conventional keying procedures. The new technology tested in 1992 employed a Census Bureau designed optical mark-sense reader system called Film Optical Sensing Device for Input to Computers (FOSDIC), in combination with key entry using a Microfilm Access Device (MAD). We refer to this system as either the FOSDIC/MAD or FOSDIC/Key system.

The FOSDIC technology is not new. In fact, the traditional FOSDIC system has been used as the primary data capture method for the last four decennial censuses. The system works well when the answer can be given in terms of a range or other categorical answers because FOSDIC recognizes a mark in a given answer position. It does not recognize a character. However, the application tested in 1992 intended to use FOSDIC to recognize and record the presence of written answers throughout the form.

To use the FOSDIC/MAD entry system, the respondent forms were microfilmed and the microfilm was processed through a FOSDIC scanner. The FOSDIC/MAD system used the information acquired from the FOSDIC scan to automatically record "yes/no" and other check box information. The new system also used the FOSDIC scan information as input to the MAD to identify the fields needing keying. It also used the scan information to guide the microfilm reader to stop on the pages with entries and, at the same time, advance the cursor of the connected computer terminal to the corresponding data entry field so the keyers could key the written responses. Hence, the new system eliminated the handling of the 12 page form by the keyer.

The objective of the FOSDIC test was to determine if the FOSDIC/Key data capture system was more timely and cost effective, while maintaining similar quality standards, compared to the data keying system used for the 1992 census production. In particular, we expected that the FOSDIC/Key system would improve timeliness and data quality as a result of the keyer not having to handle the 12 page paper form. We also expected that FOSDIC would reduce the chance for keyer omissions, reduce the processing time as a result of a reduction in keystrokes, and consequently, reduce cost. In addition, we envisioned that the microfilm could be used to enhance other processing activities for the census, such as edit resolution and table review. This paper presents the results of the evaluation comparing the mode effects of these two data capture systems.

TEST DESCRIPTION-- The feasibility test was conducted in conjunction with the 1992 Census of Agriculture since this would reflect actual production conditions, provide a large sample at no additional respondent burden and because the FOSDIC form could be designed similar to the census form, making it transparent to the respondent. The test was also designed to only include sample report forms in census Region 2, excluding the largest cases, which require special follow-ups--must cases.

The design of the FOSDIC form was limited by the decision to conduct the test concurrently with the census, which implied that the FOSDIC form should look as much like the regular census sample form as possible. This meant that the agriculture FOSDIC form had to exclude the conventional FOSDIC index marks. These are black printed squares that are used as reference marks or bench marks by FOSDIC to locate the answer positions. Instead, we converted the existing horizontal dotted lines on the current agriculture form into solid horizontal arrows and FOSDIC used them as reference marks. To use a FOSDIC form without conventional indexes required major modifications to the FOSDIC operating system and to the scan program so the horizontal lines could be used by the system. In addition, the FOSDIC form also contained special FOSDIC markings on the border of the report form to identify the questionnaire pages.

For ease of processing, the division decided that the sample forms for a given Region 2 state would be either FOSDIC or non-FOSDIC, except Iowa which was split 50-50. Table 1 shows the FOSDIC and non-FOSDIC processing counts of forms mailed out, received and keyed.

Table 1. FOSDIC Test Processing Counts

	Mailout (12 / 92)	Receipts	Keyed FOSDIC/MAD System	Keyed Regular System
FOSDIC	96,833	81,104	49,919[1]	51,936[1]
Non-FOSDIC	64,295	53,364	NA	NA

1. Different keyed totals resulted from the scan dropping cases it could not read or from microfilming rejects.

We can see from Table 1 that a total of 96,833 agriculture census addresses were mailed census sample FOSDIC forms (in ½Iowa, Ohio, Illinois and Nebraska) and 64,295 cases (in ½Iowa, Indiana and Kansas) received non-FOSDIC sample report forms. A total of 81,104 FOSDIC and 53,364 non-FOSDIC report forms were returned. Approximately 61.5 percent, 49,919 of the 81,104 returned FOSDIC forms, required keying.

The FOSDIC forms were mailed as part of the census mailout in December 1992 and were subjected to the same follow-up mailings as other census cases. Test forms were batched into groups of 95 for microfilming. Four groups (or batches) were microfilmed on one roll. The microfilm was scanned by FOSDIC to produce a data record for each form. These microfilms were used during the keying operation.

The states to which the non-FOSDIC forms were mailed were control states used to compare overall response rates. Data for all FOSDIC report forms mailed to Iowa, Ohio, Illinois and Nebraska were double-keyed, that is, the data were keyed using both the census regular "Paper" keying system (where the keyer keys from a paper report form onto a regular computer terminal) and the FOSDIC/Microfilm Access Device (FOSDIC/MAD) keying system. Measures from the two keying systems were used for all the tests on cost, time, quality and response rates. We intended for the two keying sections to operate concurrently to ensure that they would have similar working conditions. However, contrary to the original plans, the FOSDIC/MAD data keying section started 5 months after the FOSDIC/PAPER data keying section, at which time the paper section had already completed operations.

Still, to ensure that we had keyers with similar ability and background, we had requested the transfer of keyers with experience in the Economic Census and Survey area keying to work in both test keying units. Because the FOSDIC/MAD operation started later, the keyers in that section had more experience in the economic area keying (15.2 weeks versus 6 weeks for the paper keyers). They also had somewhat higher error rates (.42 percent versus .38 percent for the paper keyers), but the production rates were very similar--106 versus 102 percent MAD to paper. We did not have full control over these factors. The Data Preparation Division supervisors tried to match the rates as close as possible.

RESULTS--

Response Rates-- To test the assumption that the changes made to the report form did not adversely affect response, we compared the response rates between the two Iowa panels--the half of Iowa that received the regular sample forms versus the half of Iowa that received the FOSDIC compatible sample form at different points in the processing--approximately one week after the second, third, fourth, and fifth follow-up mailings, and also for the final response rates. We conducted one-tailed z tests and did not find any significant differences (at the .05 level of significance) between the response rates for the FOSDIC and non-FOSDIC panels of Iowa.

Table 2. 1992 Iowa FOSDIC and Non-FOSDIC Response Rates (Sample Non-Must Cases Only)

	Response Rates (Approximately 1 week after indicated follow-up mailing)				
	2nd follow-up mailing	3rd follow-up mailing	4th follow-up mailing	5th follow-up mailing	Final (No Further Change)
FOSDIC	60.46%	71.27%	80.02%	84.76%	88.14%
Non-FOSDIC	60.60%	71.49%	80.08%	85.12%	88.14%

1. Response rate is computed as the number of receipts over the total mailout minus those returned as undeliverable as addressed.

These results lead us to validate our original assumption that the changes made in the form to accommodate the use of FOSDIC in the processing did not result in any significant decreases in the response rates. The final response rate was 88.14 percent for both Iowa panels.

We also performed tests to compare the response rates from 1987 to 1992 for each of the states involved in the study. We observed a statistically significant decrease on response rates from 1987 to 1992 for all states in the study--both those that received the FOSDIC compatible sample forms, as well as those that received the standard (non-FOSDIC) sample forms. Table 3 shows the decrease for each state.

Table 3. Response Rate Decrease for Region 2 States for Sample Non-Must Universe from 1987 to 1992

State	Form Type	Response Rate Decrease from 1987 to 1992
Ohio	FOSDIC	0.88%
Illinois	FOSDIC	3.87%
Nebraska	FOSDIC	0.90%
Indiana	Non-FOSDIC	3.72%
Kansas	Non-FOSDIC	2.77%
Iowa	Both FOSDIC and Non-FOSDIC	3.09%

We believe the observed response rate decrease could be due to an apparent trend toward a decrease in overall response rates from 1987 to 1992. (As of August 20, 1994, the U.S. response rate remained at 84.5 percent. This is 1.3 percentage points lower than the final 1987 census response rate.) Since these issues are beyond the intended scope of this study, and since the decrease was evident in both the FOSDIC (.88 to 3.88 decrease) and non-FOSDIC (2.77 to 3.72 decrease) states, we did not pursue the analysis of this trend further.

Keying Processing Time and Number of Keystrokes-- To test for mode effects on time, we compared the averages of time and keystrokes per document for each system. We evaluated the *keying processing time* by comparing, for both the keyer and verifier, the average time per document required to key (or verify) the data for each system. The evaluation of the *number of keystrokes* was done by comparing the average number of keystrokes per document that the keyer and verifier keyed for each system. In addition, we made a comparison of *processing time* for the FOSDIC versus non-FOSDIC keyers that included in the FOSDIC time estimate the average time for microfilming and running the FOSDIC scan. In these comparisons we used estimates of average microfilming and scanning times provided by the Data Preparation Division (DPD) microfilming and scanning supervisors. The average microfilming time per batch is approximately 17 minutes and the average scanning time is approximately 47 seconds. Since the processing time and keystrokes were not recorded at the record level, we conducted one-tailed t tests using a paired comparison methodology where each keying batch was considered an observation.

Results from all test states combined for average time required per document indicate that the average time required per document for keying or verifying the data is significantly lower when using the FOSDIC keying method than when using the conventional keying method from the paper questionnaire. The observed average time savings is 20.7 percent for the keying operation and 23.5 percent for the verifying operation. However, when we compared the time for the two systems while adjusting the FOSDIC time to include microfilming and scanning time, the observed keyer average time savings was reduced from 20.7 percent to 13 percent.

Results from all test states combined for average keystrokes required per document show that the FOSDIC keying system resulted in a significant reduction of 37 percent in the average number of keystrokes required per document compared to the paper keying system. It is this reduction in the average number of keystrokes per document that is the major factor in the time savings mentioned above.

Please note that the specific results of these tests apply to the Region 2 sample non-must universe which was used in the study. Caution should be used when generalizing the observed average time per document and keystrokes per document to the census of agriculture universe. Due to the large differences in the number of valid keycodes for each region and the extra questions on the sample versus nonsample questionnaires, these estimates cannot be generalized to the census universe. However, we should be able to generalize the observed ***percentages*** of keystrokes and time savings from the test universe to the other states and types of agriculture questionnaires.

Cost Comparisons-- A direct salary keying cost comparison between the two systems could not be made due to the differing starting dates and consequently different salary rates for the two test keying units. Instead, we compared the two systems by evaluating the overall number of keystrokes and "charged hours" for the two systems (Table 4) and then used this data to project the keying savings that could be expected for an entire agriculture census (Table 5).

The data results presented in Table 4 show that even though 557 batches were keyed by each of the two systems, the number of documents keyed for each system was different with 49,919 documents keyed using the FOSDIC system versus 51,936 documents keyed using the Paper system. This difference occurred because the FOSDIC scanning program dropped records if FOSDIC could not read the microfilm. Also, during the pre-FOSDIC screening program, records were rejected if a complete report form (six frames) could not be identified due to microfilming problems. These situations would not have affected the Paper keying system since microfilming was not necessary. For the agriculture census, we believe that these problems could be resolved by re-microfilming problem report forms.

Table 4. Statistical Data for the FOSDIC/MAD and Paper Keying Systems

	FOSDIC/MAD System	Paper System	Percent Difference with FOSDIC
Batches	557	557	-
Documents	49,919	51,936	NA
Keystrokes	10,895,542	17,924,675	NA
Keystrokes/Document	218	345	-36.8
Keyer Hours	1,742	2,270	NA
Adjusted Hours	1,812	2,270	-20.2
Documents/Hour	28.7	22.9	25.3

NA=not applicable due to differing workloads

For the documents keyed, a total of 10,895,542 keystrokes were charged to FOSDIC keyers, compared to 17,924,675 keystrokes charged to the Paper keyers. This difference is a 36.8 percent reduction in overall keyer keystrokes per document for the FOSDIC system.

It took the FOSDIC keyers a total of 1,742 hours to key the 49,919 documents. This compared to 2,270 hours for the Paper keyers to key 51,936 documents. We adjusted the FOSDIC system keying hours by the ratio of Paper system to FOSDIC system records keyed to be able to compare the hours based on the same workload. The results indicated a 20.2 percent reduction in keying hours for the FOSDIC system. This translates into a 25.3 percent increase in documents keyed per hour with the FOSDIC system.

Projected Cost Savings--As shown in Table 5, the keying cost for the regular Paper keying system in 1992 was $1,950,000. Based on a 20.2 percent reduction in overall hours required for the FOSDIC/MAD keying operation, we estimate the FOSDIC/MAD system keying cost to be $1,556,000, resulting in a keying cost savings of approximately $394,000. However, this reduction is offset by the microfilming ($269,000), scanning ($16,000) and MAD units' reconditioning ($100,000) costs required for the FOSDIC system. As a result, the projected total data entry savings for the census of agriculture is approximately $9,000 if we were to use the FOSDIC/MAD data capture system in the census.

Table 5. Estimated Keying Costs for the Census

Activities	FOSDIC/MAD Keying Cost	Paper Keying Cost
Microfilming Forms	$ 221,000	$ 0
Duplication of Microfilm	$ 48,000	$ 0
FOSDIC Scanning	$ 16,000	$ 0
Recondition 50 MAD Units	$ 100,000	$ 0
Keying Cost	$ 1,556,000	$ 1,950,000
Total Cost	$ 1,941,000	$ 1,950,000

FOSDIC keying cost computed at a 20.2% cost reduction

Even though such relatively small savings might not warrant changing the current data entry keying system from the standard Paper keying system, we should consider that if the FOSDIC/MAD system was determined to be a reliable alternative with no negative mode effects and, was used in the future, the indexed microfilm would be available for use in edit resolution and table review. This could significantly reduce the edit and table review processing cost as well as eliminate a good portion of the cost for filing and retrieving report forms. Also, the forms could be stored using much less space on microfilm after the census processing.

Note that the costs discussed in this section do not include development cost for a revised FOSDIC compatible form or the associated changes to the FOSDIC/key system prompted by such a change to the form. It does not include either the costs associated with addressing certain constraints and limitations put on the test because it was conducted as part of the 1992 Census of Agriculture, and to prevent confounding by testing too many changes. Those changes would have to be incorporated before any future implementation.

Data Quality Tests--We evaluated the data entry quality by first matching the 49,395 records which were successfully keyed by both systems. This process identified all entries which did not match between the two data entry systems for each record. Table 6 shows the results from this initial match.

Table 6. Distributions of FOSDIC Test Records With/Without Discrepancies and Matched/Non-Matched Keycode Entries

Total Keyed CFNs (Records)	49,395	%
CFNs with no Discrepancies	17,489	35.4
CFNs with one or more Discrepancies	31,906	64.6
Total Keyed Entries	2,826,487	%
Matched Entries	2,752,247	97.4
Nonmatched Entries	74,240	2.6

From Table 6 we can see that although 64.6 percent of the records contained some nonmatched entries, only 2.6 percent of the total keyed entries were nonmatched.

Our next step in evaluating the data quality was to review the nonmatched entries and resolve the discrepancy by making a determination as to which data entry system was in error and the type of error. Due to the large number of records needing resolution, 31,906 records, we systematically selected a random 20% sample of the cases containing nonmatched entries. After deleting certain records that underwent extra processing in one of the data entry systems, the sampling process resulted in 6,292 records with 14,683 nonmatched entries.

We reviewed the microfilm for these cases and made a determination as to which system was in error and the type of error. We then computed correct rates for the discrepancies and overall error rates for each system. During the process of resolving the discrepancies, we identified and excluded from the evaluation certain recurring discrepancies that we believed were related to forms design and were therefore not mode effects. As a result, the final distribution of matched and nonmatched entries for the 6,292 sampled records was 12,807 nonmatched entries to 353,444 matched entries.

Table 7 shows how these 12,807 nonmatched entries were distributed between FOSDIC/Key correct, Paper/Key correct, and Neither system correct for both known and unknown causes.

Table 7. Discrepancy Resolution Distribution by Cause

Cause	All Discrepancies	FOSDIC Correct		Paper Correct		Neither Correct	
	n	n	%	n	%	n	%
TOTAL	12,807	4,275	33	7,951	62	581	5
Cause = Unknown	9,437	3,500	37	5,726	61	211	2
Cause = Known (Total)	3,370	775	23	2,225	66	370	11
Cause = R	1,978	537	27	1,186	60	255	13
Cause = S	702	227	32	383	55	92	13
Cause = O	352	11	3	318	90	23	1
Cause = P	261	0	-	261	100	0	-
Cause = C	53	0	-	53	100	0	-
Cause = X	24	0	-	24	100	0	-

Cause=R READABILITY
The data on the microfilm was difficult to read either due to the respondent's handwriting or due to the entry appearing light or faint on the microfilm.

Cause = S SHIFT
The keyer keyed an entry which would be correct with the exception that they keyed it for the wrong keycode.

Cause = O OUTSIDE
The respondent wrote his/her answer outside of the space provided.

Cause = P PAGE
The microfilming process resulted in either a page missing or an unreadable double exposure frame of film.

Cause = C CORRESPONDENCE
The respondent attached correspondence which obscured the data on the microfilm and was not removed prior to microfilming.

Cause = X X'd OUT
The respondent crossed out and scratched out sections of the form and the form was not removed/rejected prior to keying.

From Table 7 we can see that the FOSDIC/Key system was correct for 33 percent of the nonmatched entries, while the paper system was correct for 62 percent of the nonmatched entries. We can also see that while we were able to identify the cause of the error for 3,370 or 26 percent of the nonmatched entries, we were unable to identify the cause for 9,437 or 74 percent of the nonmatched entries.

Table 8 shows the final correct and error rates for the 12,807 nonmatched entries in the final sample.

Table 8. Tabulation of Correct and Error Rates for Each System

	FOSDIC/Key	Paper/Key
Total Entries Keyed	366,251	366,251
Matched Entries	353,444	353,444
Correct Discrepancies[1]	4,275	7,951
Total Correct Entries	357,719	361,395
Correct Rate	**97.6%**	**98.7%**
Total Incorrect Entries[2]	8,190	4,515
Error Rate	**2.2%**	**1.2%**

1. Excludes 581 discrepancies where the correct entry could not be determined or neither the FOSDIC/Key system nor the Paper/Key system was correct.

2. Incudes 239 discrepancies where neither the FOSDIC/Key system nor the Paper/Key system was correct and excludes 342 discrepancies where the correct entry could not be determined.

As can be seen from Table 8, the error rate is 2.2 percent for the FOSDIC/Key system and 1.2 percent for the Paper/Key system. Our original expectation was that the FOSDIC/Key system would result in better data entry quality and, therefore, a lower error rate. Given this expectation, we intended to conduct a one-tail t test to identify any significant reduction in the error rate when using the FOSDIC/Key system. However, the FOSDIC/Key system resulted in a higher error rate. Therefore, a one-tail t test to determine if the FOSDIC/Key system had a significantly lower error rate would not have shown significance. As a result of these findings, we failed to reject the null hypothesis that the FOSDIC/Key system had an error rate equal or higher than the error rate of the Paper/Key system.

However, if our test had been to detect any difference between the error rates, a two-tailed test of this data would have shown a statistically significant difference in the error rates at the .05 level of significance. That result would have implied that the error rate for the FOSDIC/Key system was significantly higher than that of the Paper/Key system.

CONCLUSIONS/RECOMMENDATIONS-- We did not find clear advantages in using the FOSDIC/Key system for keying future censuses of agriculture. Our results indicate that there are essentially no expected cost reductions and no data entry quality improvements that result from the proposed data entry system. Furthermore, although the limited changes made to the questionnaire in order to conduct the test did not negatively affect response rates, we found some recurring discrepancies that were associated with the lack of questionnaire development to a FOSDIC specific form. If such a form were developed, it would increase the cost for the use of the FOSDIC/Key system and might require further testing to ensure no negative effects on response rates.

Finally, we found the FOSDIC/Key system resulted in an improvement over the Paper/Key system in the number of keystrokes per document and consequently, a reduction in the time per document. The reduction in keystrokes could have resulted in a cost reduction, but due to the added costs of microfilming, scanning, and reconditioning the equipment necessary to use this system for an entire census, the actual estimated cost reduction is relatively nil.

Consequently, the only area of improvement that could be expected by using the FOSDIC/Key system is a possible earlier publication date as a result of the time reduction. However, due to other census of agriculture processing constraints, it is likely that the observed time reduction would result instead in the hiring of fewer keyers for the census keying operation and completing the operation in approximately the same length of time. Hence, the time reduction becomes instead a staff reduction with no real impact on the publication timeframe.

In addition, it is likely that other data capture technologies such as electronic imaging will be available for use in the 2002 Census of Agriculture. This would mean that any resources invested in order to use a system such as the FOSDIC/Key system for the 1997 Census of Agriculture would likely be investments for a system used only once. Consequently, based on the results from this study and the rapid advances in data capture technology, our recommendation is not to use the FOSDIC/Key system for the 1997 Census of Agriculture.

ACKNOWLEDGEMENTS-- We would like to acknowledge Erne Wilkins, Ruth Ann Killion, George Pierce, Stephen Ash, Sandra Hairston, Ann Vacca, and Vicki Huggins for their involvement in the initial planning and review of the evaluation proposal, Ruth Detlefsen and Paul Lewis for their comments on the report of findings, and Ricardo Ruiz for his involvement in the programming aspects of the test. We would also like to thank the many divisions of the Census Bureau that contributed to the implementation and evaluation of this study.

This paper reports the general results of research undertaken by Census Bureau staff. The views expressed are attributable to the author(s) and do not necessarily reflect those of the Census Bureau.