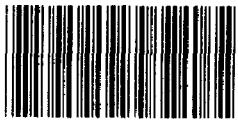


~~SHH~~  
129876

UNITED STATES GENERAL ACCOUNTING OFFICE  
Washington, D.C. 20548

FOR RELEASE ON DELIVERY  
EXPECTED AT 10:00 A.M.  
THURSDAY, MAY 15, 1986

STATEMENT OF  
GENE L. DODARO  
ASSOCIATE DIRECTOR, GENERAL GOVERNMENT DIVISION  
BEFORE THE  
SUBCOMMITTEE ON CENSUS AND POPULATION  
COMMITTEE ON POST OFFICE AND CIVIL SERVICE  
HOUSE OF REPRESENTATIVES  
ON  
THE CENSUS BUREAU'S PREPARATIONS FOR THE  
1990 DECAENNIAL CENSUS



129876

129876 - 036416

Mr. Chairman and Members of the Subcommittee:

I am pleased to appear today to discuss preparations for the 1990 Decennial Census in two interrelated areas, questionnaire development and data capture technology. Improvements in these areas could greatly contribute to controlling costs and enhancing the quality of the census. With me is Jerry Donoghue, who is responsible for our work on the questionnaire, and Johnnie Butts, who is responsible for our census automation work.

The Bureau's planning and preparations, including its tests and decisions to date, have led us to believe that the 1990 census will not be as cost efficient as it could be. The Bureau's reluctance to test a shorter short form, its questionable approach to procuring optical mark reader (OMR) equipment, and the questions raised by its recent decision on data capture and processing office configuration all point to missed opportunities to significantly improve upon the 1980 census.

ADVANTAGES OF A SHORTER

SHORT FORM

As we mentioned in our previous testimonies before your subcommittee in June 1984 and April and July 1985, we have reservations about the size and content of the short form sent to about 81 percent of the households in 1980. We continue to believe that the short form should be limited to the basic

questions needed to obtain an accurate population count--that is, questions oriented towards population characteristics and those used to improve the count. Housing questions--such as plumbing, value, and rent of housing units--increase the questionnaire's complexity and consequently discourage response. We believe that housing data obtained from the long form (sample questionnaire) may meet federal needs. This sample form contained not only the short form questions but more detailed questions on population and housing as well. One housing unit in six was asked to complete the long form in 1980 except for communities under 2500 people, where one half of the housing units were sampled.

Our more recent work has raised questions regarding the federal need for housing data from 100 percent of the nation's households. For example, we found that some federal data users were actually using sample data even though 100 percent data or data collected from all of the households was available, and that some users had requested 100 percent housing data for geographical levels for which data are also estimated from sample questionnaires. The need for housing data from all households should be more closely weighed against associated collection costs and respondent burden.

The content and design of the questionnaire is a major factor affecting response rates, quality of response, response

burden, and data processing requirements. Since about 70 percent of the \$1.1 billion in 1980 Decennial Census costs were incurred with the collection, preparation, and processing of the data, efforts to reduce the short form questionnaire, both in size and questions, could be cost beneficial. This is particularly important considering that the 1990 Decennial Census will record information from about 106 million households, 18 million more than for 1980. If the Bureau does not streamline and simplify the questionnaire form, it will be missing out on a potential cost savings opportunity, as well as a chance to improve the perceptions and receptiveness of the U.S. public to the census.

Cost could be reduced

The mail response rates in 1990 will have a direct impact on the nonresponse followup costs. Streamlining and simplifying the short form should improve the mail response rates for 1990, and might greatly reduce the high cost of sending out enumerators to followup with nonrespondents, particularly if better response rates are obtained from the hard-to ENUMERATE areas. For 1990, the Bureau estimates that each 1- percent increase in the mail response rate would save about \$6 million in followup costs.

That the size of the questionnaire form influences the response rates is particularly evident in inner city areas. For the 1980 census, the mail return rates for the short form was over 7 percent better than the long forms in these

hard-to-enumerate areas. Also, rates for the short form questionnaires have been consistently higher than for the long forms in the 1985 and 1986 Pretests, as shown below.

<u>Pretest</u>	<u>Response Rate</u>	
	<u>Short form</u>	<u>Long form</u>
Jersey City	39	31
Tampa	58	48
Los Angeles	36	28
East Central Mississippi	65	59

We do not know the extent to which a shorter short form will increase response rates. However, without a test no one will know. Considering that a short form will be sent to about 85 million households in 1990, and the pressures to hold down government spending, evaluation of a revised form seems worthwhile.

Some enumerator visits also could be eliminated with a shorter form because fewer respondents may return incomplete questionnaires which require enumerators to collect the missing data. Of the 64 million questionnaires that were returned by mail in the 1980 census, 13 percent of the short forms did not meet the Bureau's standards for completeness. On the other hand, 36 percent of the long forms did not meet these criteria.

A shorter, simplified short form questionnaire also could be processed at less cost. However, the Bureau must act on the size of the form to potentially benefit from data processing savings. Also, a reduced short form may provide the Bureau with cost-saving options for its data automation decisions.

Respondent burden reduced  
with fewer questions

Reducing the size of the 1980 short form will benefit respondents because the burden of completing the form will be reduced in terms of number of questions, time, and in the perception of difficulty.

With a short form oriented toward population and coverage only, about 85 million households would answer fewer questions. This could result in up to a one-third reduction in the time to complete the questionnaire.

The perception of the questionnaire is an important factor. It is a burden for the respondent if the individual perceives the form to be a difficult task, and this could affect the completion of the form. The Bureau has indications that perception of difficulty is a factor affecting form completion.

For example, a Bureau post-1980 Census study found that nonrespondents attributed not starting to fill out the form to the perceived difficulty of the task. The study showed that "the

easier to fill the form was perceived, the more likely it was to be started." The Bureau also found that, for the critical phase of finishing the form, nonrespondents attributed not completing the form to the "amount of work involved." In another study, the Bureau discovered that the difficulty--experienced or perceived--associated with completing a self-enumeration form can adversely affect response to subsequent items on the form.

Confirming these study findings, enumerator supervisors at the Jersey City Pretest site told us that some respondents commented after enumeration interviews that the short form appeared complicated. This concern, along with the perception that the form was too long, was also expressed by focus groups at the Los Angeles and Mississippi test sites.

PARTICIPANTS IN RECENT FOCUS GROUP

STUDIES WERE CONCERNED OVER SHORT

FORM LENGTH AND COMPLEXITY

The Bureau's recently conducted "focus group" studies at the Tampa, Los Angeles, and Mississippi pretest sites showed that lower income groups generally had questions on, or objections to, the housing questions on the short form questionnaire. In addition, concerns were expressed that the short form questionnaire was too long and complex at the Los Angeles and Mississippi sites.

Focus group studies were generally done to ascertain the motivational messages and themes that would encourage hard-to-enumerate populations to respond to the census and return their forms. The focus group approach attempts to develop qualitative insight and direction for further research. While these three focus groups studies were limited to a total of 95 participants and the results cannot be projected, they provide useful indications of individuals' views of the short form questionnaire.

All three studies showed the participants had concern over the housing questions. In a Bureau observation memorandum on the Tampa study, the plumbing question caused groups to discuss why any of the housing questions were needed for the census. In the report on the Tampa focus groups, views were reported that some of the information requested for the census--e.g., value of an individual's home--was "none of the government's business." At Mississippi, housing questions such as those concerning entrance to living quarters and value of home were considered too personal, and participants wanted to know "why do they need to know all those other questions" when the census is a count of population. At the Los Angeles site, housing questions on "entrance to living quarters," "value of property" and "rent" were objected to the most.

In a memorandum of observation on the Los Angeles study, the Bureau noted that those queried believed the questions on plumbing (which were thought to be "ridiculous"), value of property and rent were considered to be too personal. According to the Los Angeles report, most of the objectionable questions related to housing items. These items were "specifically perceived to be a means of tricking people into exposing additional unreported individuals who occupy the household." Also, the number of rooms question was considered by some to be outside the purpose of a head count. One of the Los Angeles report's conclusions states that "objections were raised to some of the questions contained in the Census short form (particularly the housing-related questions) to the effect that their perceived true purpose is surreptitious."

At the Los Angeles and Mississippi sites, the short form's length and complexity concerned the participants. In Los Angeles, initial reactions to the short form were "it's much too much, too long"; "it looks too big. What they're asking, you should be able to put it on a 3 x 5 index card." The Bureau's observation memorandum on Los Angeles noted that a first reaction to the short form was "wow ... too big ... overwhelming." The Bureau's observation memorandum on Mississippi noted that people said the "form looked complicated, too long ... too nosey." The memorandum further noted that people felt that it was not very complicated when they were walked through the

questionnaire. One representative quotation in the Mississippi consultant's report was "too much work to filling that out."

#### IMPROVED DATA ACCURACY

The degree of literacy skills affects data accuracy, and form completion. Respondents with marginal literacy skills should be better able to respond correctly to a simplified and reduced short form questionnaire. More people could respond, resulting in a more precise population count; and since less data would be collected through door-to-door enumeration, enumerator bias would be reduced.

The Bureau has not conducted any literacy tests since the 1970's when experiments were conducted to measure respondent literacy skills or the reading level required to complete the short form questionnaire. Literacy level is an important issue in developing a successful questionnaire. As such, the Bureau should have gained some insights into problems people have in understanding census questions by conducting studies with pretest respondents. While realizing that the impact of marginal literacy skills on accurate completion of the short form is unknown, it seems reasonable to assume that a simpler, reduced size, easier-to-read form could only help.

Moreover, since the collection and processing of data could be accomplished more efficiently and effectively using a

streamlined short form, the population counts could be tabulated earlier allowing more time to review the results. Additionally, a streamlined short form would allow the Bureau more flexibility in considering data capture equipment.

Now, I will discuss the subject of data capture equipment.

BUREAU'S DECISION NOT TO USE OMR

We believe the Bureau's decision to discontinue consideration of OMR equipment was influenced by its late start in detailed planning, reluctance to revise the questionnaire form, and a slow procurement process. Whether OMR equipment could have been adapted for use during the 1990 census may never be known. However, because of its actions, the Bureau has excluded an option for using new technology without fully exploring its potential.

The commercially available OMR equipment considered by the Bureau had both advantages and disadvantages for a census. The OMR equipment employs a one-step process to read and record the data. The Bureau's traditional film to tape data capture system called FACT required three sequential processes--filming, film development and scanning the film. Also, on the basis of our observation of the 1985 pretest, the OMR equipment was easy to operate, and training time is short. For some of the FACT

processes, a knowledge of chemistry and other technical subjects are needed.

On the other hand, the commercially available OMR was designed to process a single page form much smaller than the census short form questionnaire. In addition, the Bureau for the past several censuses has used a multipage booklet form for its long form questionnaire. The OMR was also designed for use in a controlled environment such as grading test answer sheets where the students are provided with #2 pencils. The Bureau's FACT equipment is generally not affected by these constraints.

Despite these known disadvantages, the Bureau decided to use existing OMR equipment in its 1985 pretest, and as a result the pretest proved very little. Most of the major problems that occurred with the OMR in the test had previously been known. The test simply helped to identify the magnitude of these problems.

The prospective vendor had made several proposals to overcome the limitations by designing a modified OMR but the procurement effort for a modified OMR was protracted and eventually terminated. Thus, no OMR is being used in the 1986 pretest, although testing data capture alternatives was a major objective of the test.

In November 1985, the Bureau decided to terminate its consideration of OMR technology as the primary data capture technology for the 1990 census. As a result of its decision, the Bureau ruled out the possibility of exploring the usefulness of this equipment, the new technology being considered for the 1990 census. Thus, the Bureau limited its options to the traditional FACT system and data keying.

From the beginning, the Bureau was aware that the commercially available OMR equipment did not satisfy all existing decennial census needs such as paper size and the use of a variety of marking instruments. Thus, testing the unmodified OMR equipment unless the Bureau was considering revising its requirements did not seem prudent. Absent this possibility, the Bureau should have formalized its requirements and initiated an effort to test a modified optical mark reader early in the decade.

#### CURRENT DECISIONS AND PLANS ON DATA CAPTURE

The Bureau's present plans on data capture and processing office configuration, as developed in late April, will result in a processing operation for most of the nation similar to that used in 1980. Our understanding is that the Bureau's FACT system will be used as the primary data capture technology. Filming of all questionnaires will be performed in 10 to 15 processing offices. The location of the film processing and

scanning has not yet been decided. Questionnaires from hard-to-enumerate areas (possibly 12 percent of the nation) will be sent directly to these processing offices to be captured and where they may be automatically edited.

For the remainder of the nation, questionnaires will be returned to district offices, where they will be manually edited, similar to the 1980 census. Only after the questionnaires are "perfected" will they be sent in batches to the processing offices for data capture. In 1980, questionnaires were not sent to data capture until all questionnaires from a district were reviewed, "perfected" and batched, and the office was closed.

We have several reservations about the Bureau's late April decisions on data capture and the data processing office configurations. The Bureau's current plan for data processing is a hybrid system which incorporates concurrent data processing for a small portion of the nation and a modified 1980 data capture for the remainder of the country. On the one hand, we are pleased that the Bureau has decided on these basic automation activities 5 months earlier than originally planned as we have previously advocated, albeit without evaluation information from the 1986 pretest. We are also pleased that the Bureau has decided against using data keying, a relatively slow, error prone and expensive technology as a primary data capture technology. On the other hand, we are concerned with the decision to forego

concurrent processing and revert to manual procedures for most of the nation because we believe the Bureau has thus foregone the benefits that could be derived from a more automated operation.

Sufficient details of the plans for the processing operations have not been developed to allow us to fully assess them. However, on the basis of the information obtained to date, we believe that the Bureau has compromised its goals for automation and for the census as a whole.

Because the majority of the questionnaires will not be captured until they are manually edited and "perfected," the important benefits of concurrent processing will not be obtained. Manual processes, particularly editing, will be used. Thus, some of the benefits from automation including speed and consistency and accuracy will not be obtained. Instead, a small army of temporary employees will probably be used. About 37,000 clerks were employed in the 1980 census to check returned questionnaires for complete and consistent entries. An early automated back up file will not be prepared. And premature destruction of the questionnaire forms as occurred in the 1980 census would remain a potential problem.

Maintaining controls over forms returned may be difficult if the questionnaires are sent to the processing office on a piecemeal basis. If information, such as population count, is

not manually recorded from the returned questionnaires in the district offices, it also may be difficult to perform some coverage improvement programs, such as local review or canvassing neighborhoods.

Additionally, we believe that the use of manual processes in most of the district offices and the use of 10-15 processing offices as compared to 3 in the 1980 census will help drive the per household cost, exclusive of inflation, of the 1990 census beyond the cost of the prior census.

We intend to monitor the continuing developments on data capture and processing office configuration because of the importance of the data processing operation in a decennial census.

This concludes my remarks, and I would be happy to respond to any questions.

UNITED STATES GENERAL ACCOUNTING OFFICE  
Washington, D.C. 20548

FOR RELEASE ON DELIVERY  
EXPECTED AT 10:00 A.M.  
THURSDAY, MAY 15, 1986

STATEMENT OF  
GENE L. DODARO  
ASSOCIATE DIRECTOR, GENERAL GOVERNMENT DIVISION  
BEFORE THE  
SUBCOMMITTEE ON CENSUS AND POPULATION  
COMMITTEE ON POST OFFICE AND CIVIL SERVICE  
HOUSE OF REPRESENTATIVES  
ON  
THE CENSUS BUREAU'S PREPARATIONS FOR THE  
1990 DECENTNIAL CENSUS

Mr. Chairman and Members of the Subcommittee:

I am pleased to appear today to discuss preparations for the 1990 Decennial Census in two interrelated areas, questionnaire development and data capture technology. Improvements in these areas could greatly contribute to controlling costs and enhancing the quality of the census. With me is Jerry Donoghue, who is responsible for our work on the questionnaire, and Johnnie Butts, who is responsible for our census automation work.

The Bureau's planning and preparations, including its tests and decisions to date, have led us to believe that the 1990 census will not be as cost efficient as it could be. The Bureau's reluctance to test a shorter short form, its questionable approach to procuring optical mark reader (OMR) equipment, and the questions raised by its recent decision on data capture and processing office configuration all point to missed opportunities to significantly improve upon the 1980 census.

ADVANTAGES OF A SHORTER

SHORT FORM

As we mentioned in our previous testimonies before your subcommittee in June 1984 and April and July 1985, we have reservations about the size and content of the short form sent to about 81 percent of the households in 1980. We continue to believe that the short form should be limited to the basic

questions needed to obtain an accurate population count--that is, questions oriented towards population characteristics and those used to improve the count. Housing questions--such as plumbing, value, and rent of housing units--increase the questionnaire's complexity and consequently discourage response. We believe that housing data obtained from the long form (sample questionnaire) may meet federal needs. This sample form contained not only the short form questions but more detailed questions on population and housing as well. One housing unit in six was asked to complete the long form in 1980 except for communities under 2500 people, where one half of the housing units were sampled.

Our more recent work has raised questions regarding the federal need for housing data from 100 percent of the nation's households. For example, we found that some federal data users were actually using sample data even though 100 percent data or data collected from all of the households was available, and that some users had requested 100 percent housing data for geographical levels for which data are also estimated from sample questionnaires. The need for housing data from all households should be more closely weighed against associated collection costs and respondent burden.

The content and design of the questionnaire is a major factor affecting response rates, quality of response, response

burden, and data processing requirements. Since about 70 percent of the \$1.1 billion in 1980 Decennial Census costs were incurred with the collection, preparation, and processing of the data, efforts to reduce the short form questionnaire, both in size and questions, could be cost beneficial. This is particularly important considering that the 1990 Decennial Census will record information from about 106 million households, 18 million more than for 1980. If the Bureau does not streamline and simplify the questionnaire form, it will be missing out on a potential cost savings opportunity, as well as a chance to improve the perceptions and receptiveness of the U.S. public to the census.

Cost could be reduced

The mail response rates in 1990 will have a direct impact on the nonresponse followup costs. Streamlining and simplifying the short form should improve the mail response rates for 1990, and might greatly reduce the high cost of sending out enumerators to followup with nonrespondents, particularly if better response rates are obtained from the hard-to-enumerate areas. For 1990, the Bureau estimates that each 1- percent increase in the mail response rate would save about \$6 million in followup costs.

That the size of the questionnaire form influences the response rates is particularly evident in inner city areas. For the 1980 census, the mail return rates for the short form was over 7 percent better than the long forms in these

hard-to-enumerate areas. Also, rates for the short form questionnaires have been consistently higher than for the long forms in the 1985 and 1986 Pretests, as shown below.

<u>Pretest</u>	<u>Response Rate</u>	
	<u>Short form</u>	<u>Long form</u>
Jersey City	39	31
Tampa	58	48
Los Angeles	36	28
East Central Mississippi	65	59

We do not know the extent to which a shorter short form will increase response rates. However, without a test no one will know. Considering that a short form will be sent to about 85 million households in 1990, and the pressures to hold down government spending, evaluation of a revised form seems worthwhile.

Some enumerator visits also could be eliminated with a shorter form because fewer respondents may return incomplete questionnaires which require enumerators to collect the missing data. Of the 64 million questionnaires that were returned by mail in the 1980 census, 13 percent of the short forms did not meet the Bureau's standards for completeness. On the other hand, 36 percent of the long forms did not meet these criteria.

A shorter, simplified short form questionnaire also could be processed at less cost. However, the Bureau must act on the size of the form to potentially benefit from data processing savings. Also, a reduced short form may provide the Bureau with cost-saving options for its data automation decisions.

Respondent burden reduced  
with fewer questions

Reducing the size of the 1980 short form will benefit respondents because the burden of completing the form will be reduced in terms of number of questions, time, and in the perception of difficulty.

With a short form oriented toward population and coverage only, about 85 million households would answer fewer questions. This could result in up to a one-third reduction in the time to complete the questionnaire.

The perception of the questionnaire is an important factor. It is a burden for the respondent if the individual perceives the form to be a difficult task, and this could affect the completion of the form. The Bureau has indications that perception of difficulty is a factor affecting form completion.

For example, a Bureau post-1980 Census study found that nonrespondents attributed not starting to fill out the form to the perceived difficulty of the task. The study showed that "the

easier to fill the form was perceived, the more likely it was to be started." The Bureau also found that, for the critical phase of finishing the form, nonrespondents attributed not completing the form to the "amount of work involved." In another study, the Bureau discovered that the difficulty--experienced or perceived--associated with completing a self-enumeration form can adversely affect response to subsequent items on the form.

Confirming these study findings, enumerator supervisors at the Jersey City Pretest site told us that some respondents commented after enumeration interviews that the short form appeared complicated. This concern, along with the perception that the form was too long, was also expressed by focus groups at the Los Angeles and Mississippi test sites.

PARTICIPANTS IN RECENT FOCUS GROUP

STUDIES WERE CONCERNED OVER SHORT

FORM LENGTH AND COMPLEXITY

The Bureau's recently conducted "focus group" studies at the Tampa, Los Angeles, and Mississippi pretest sites showed that lower income groups generally had questions on, or objections to, the housing questions on the short form questionnaire. In addition, concerns were expressed that the short form questionnaire was too long and complex at the Los Angeles and Mississippi sites.

Focus group studies were generally done to ascertain the motivational messages and themes that would encourage hard-to Enumerate populations to respond to the census and return their forms. The focus group approach attempts to develop qualitative insight and direction for further research. While these three focus groups studies were limited to a total of 95 participants and the results cannot be projected, they provide useful indications of individuals' views of the short form questionnaire.

All three studies showed the participants had concern over the housing questions. In a Bureau observation memorandum on the Tampa study, the plumbing question caused groups to discuss why any of the housing questions were needed for the census. In the report on the Tampa focus groups, views were reported that some of the information requested for the census--e.g., value of an individual's home--was "none of the government's business." At Mississippi, housing questions such as those concerning entrance to living quarters and value of home were considered too personal, and participants wanted to know "why do they need to know all those other questions" when the census is a count of population. At the Los Angeles site, housing questions on "entrance to living quarters," "value of property" and "rent" were objected to the most.

In a memorandum of observation on the Los Angeles study, the Bureau noted that those queried believed the questions on plumbing (which were thought to be "ridiculous"), value of property and rent were considered to be too personal. According to the Los Angeles report, most of the objectionable questions related to housing items. These items were "specifically perceived to be a means of tricking people into exposing additional unreported individuals who occupy the household." Also, the number of rooms question was considered by some to be outside the purpose of a head count. One of the Los Angeles report's conclusions states that "objections were raised to some of the questions contained in the Census short form (particularly the housing-related questions) to the effect that their perceived true purpose is surreptitious."

At the Los Angeles and Mississippi sites, the short form's length and complexity concerned the participants. In Los Angeles, initial reactions to the short form were "it's much too much, too long"; "it looks too big. What they're asking, you should be able to put it on a 3 x 5 index card." The Bureau's observation memorandum on Los Angeles noted that a first reaction to the short form was "wow ... too big ... overwhelming." The Bureau's observation memorandum on Mississippi noted that people said the "form looked complicated, too long ... too nosey." The memorandum further noted that people felt that it was not very complicated when they were walked through the

questionnaire. One representative quotation in the Mississippi consultant's report was "too much work to filling that out."

#### IMPROVED DATA ACCURACY

The degree of literacy skills effects data accuracy, and form completion. Respondents with marginal literacy skills should be better able to respond correctly to a simplified and reduced short form questionnaire. More people could respond, resulting in a more precise population count; and since less data would be collected through door-to-door enumeration, enumerator bias would be reduced.

The Bureau has not conducted any literacy tests since the 1970's when experiments were conducted to measure respondent literacy skills or the reading level required to complete the short form questionnaire. Literacy level is an important issue in developing a successful questionnaire. As such, the Bureau should have gained some insights into problems people have in understanding census questions by conducting studies with pretest respondents. While realizing that the impact of marginal literacy skills on accurate completion of the short form is unknown, it seems reasonable to assume that a simpler, reduced size, easier-to-read form could only help.

Moreover, since the collection and processing of data could be accomplished more efficiently and effectively using a

streamlined short form, the population counts could be tabulated earlier allowing more time to review the results. Additionally, a streamlined short form would allow the Bureau more flexibility in considering data capture equipment.

Now, I will discuss the subject of data capture equipment.

#### BUREAU'S DECISION NOT TO USE OMR

We believe the Bureau's decision to discontinue consideration of OMR equipment was influenced by its late start in detailed planning, reluctance to revise the questionnaire form, and a slow procurement process. Whether OMR equipment could have been adapted for use during the 1990 census may never be known. However, because of its actions, the Bureau has excluded an option for using new technology without fully exploring its potential.

The commercially available OMR equipment considered by the Bureau had both advantages and disadvantages for a census. The OMR equipment employs a one-step process to read and record the data. The Bureau's traditional film to tape data capture system called FACT required three sequential processes--filming, film development and scanning the film. Also, on the basis of our observation of the 1985 pretest, the OMR equipment was easy to operate, and training time is short. For some of the FACT

processes, a knowledge of chemistry and other technical subjects are needed.

On the other hand, the commercially available OMR was designed to process a single page form much smaller than the census short form questionnaire. In addition, the Bureau for the past several censuses has used a multipage booklet form for its long form questionnaire. The OMR was also designed for use in a controlled environment such as grading test answer sheets where the students are provided with #2 pencils. The Bureau's FACT equipment is generally not affected by these constraints.

Despite these known disadvantages, the Bureau decided to use existing OMR equipment in its 1985 pretest, and as a result the pretest proved very little. Most of the major problems that occurred with the OMR in the test had previously been known. The test simply helped to identify the magnitude of these problems.

The prospective vendor had made several proposals to overcome the limitations by designing a modified OMR but the procurement effort for a modified OMR was protracted and eventually terminated. Thus, no OMR is being used in the 1986 pretest, although testing data capture alternatives was a major objective of the test.

In November 1985, the Bureau decided to terminate its consideration of OMR technology as the primary data capture technology for the 1990 census. As a result of its decision, the Bureau ruled out the possibility of exploring the usefulness of this equipment, the new technology being considered for the 1990 census. Thus, the Bureau limited its options to the traditional FACT system and data keying.

From the beginning, the Bureau was aware that the commercially available OMR equipment did not satisfy all existing decennial census needs such as paper size and the use of a variety of marking instruments. Thus, testing the unmodified OMR equipment unless the Bureau was considering revising its requirements did not seem prudent. Absent this possibility, the Bureau should have formalized its requirements and initiated an effort to test a modified optical mark reader early in the decade.

#### CURRENT DECISIONS AND PLANS ON DATA CAPTURE

The Bureau's present plans on data capture and processing office configuration, as developed in late April, will result in a processing operation for most of the nation similar to that used in 1980. Our understanding is that the Bureau's FACT system will be used as the primary data capture technology. Filming of all questionnaires will be performed in 10 to 15 processing offices. The location of the film processing and

scanning has not yet been decided. Questionnaires from hard-to-enumerate areas (possibly 12 percent of the nation) will be sent directly to these processing offices to be captured and where they may be automatically edited.

For the remainder of the nation, questionnaires will be returned to district offices, where they will be manually edited, similar to the 1980 census. Only after the questionnaires are "perfected" will they be sent in batches to the processing offices for data capture. In 1980, questionnaires were not sent to data capture until all questionnaires from a district were reviewed, "perfected" and batched, and the office was closed.

We have several reservations about the Bureau's late April decisions on data capture and the data processing office configurations. The Bureau's current plan for data processing is a hybrid system which incorporates concurrent data processing for a small portion of the nation and a modified 1980 data capture for the remainder of the country. On the one hand, we are pleased that the Bureau has decided on these basic automation activities 5 months earlier than originally planned as we have previously advocated, albeit without evaluation information from the 1986 pretest. We are also pleased that the Bureau has decided against using data keying, a relatively slow, error prone and expensive technology as a primary data capture technology. On the other hand, we are concerned with the decision to forego

concurrent processing and revert to manual procedures for most of the nation because we believe the Bureau has thus foregone the benefits that could be derived from a more automated operation.

Sufficient details of the plans for the processing operations have not been developed to allow us to fully assess them. However, on the basis of the information obtained to date, we believe that the Bureau has compromised its goals for automation and for the census as a whole.

Because the majority of the questionnaires will not be captured until they are manually edited and "perfected," the important benefits of concurrent processing will not be obtained. Manual processes, particularly editing, will be used. Thus, some of the benefits from automation including speed and consistency and accuracy will not be obtained. Instead, a small army of temporary employees will probably be used. About 37,000 clerks were employed in the 1980 census to check returned questionnaires for complete and consistent entries. An early automated back up file will not be prepared. And premature destruction of the questionnaire forms as occurred in the 1980 census would remain a potential problem.

Maintaining controls over forms returned may be difficult if the questionnaires are sent to the processing office on a piecemeal basis. If information, such as population count, is

not manually recorded from the returned questionnaires in the district offices, it also may be difficult to perform some coverage improvement programs, such as local review or recanvassing neighborhoods.

Additionally, we believe that the use of manual processes in most of the district offices and the use of 10-15 processing offices as compared to 3 in the 1980 census will help drive the per household cost, exclusive of inflation, of the 1990 census beyond the cost of the prior census.

We intend to monitor the continuing developments on data capture and processing office configuration because of the importance of the data processing operation in a decennial census.

This concludes my remarks, and I would be happy to respond to any questions.