BUREAU OF THE CENSUS

STATISTICAL RESEARCH DIVISION REPORT SERIES

SRD Research Report Number: CENSUS/SRD/RR-84/12

RESEARCH PLAN ON ADJUSTMENT

by

Undercount Research Staff
Statistical Research Division
U.S. Bureau of the Census

Room 3582, F.O.B. #3
Washington, D.C. 20233

(301)763-5926

Recommended by: Kirk M. Wolter

Report completed: July 25, 1984

Report issued: August 1, 1984

This research plan was developed by the Undercount Research Staff, Statistical Research Division, under the supervision of Howard Hogan, Chief. Important contributions were made by Kirk Wolter, Chief of Statistical Research Division, Gregg Diffendal of the Undercount Research Staff, and Charles Cowan, now of the International Statistics Program Center. Many other people both from within the Bureau of the Census and from outside groups have made important suggestions which were incorporated into the final draft.

This paper was presented by Howard Hogan at the 1984 American Statistical Association Annual Meeting, August 13-16, 1984, Philadelphia, Pennsylvania.

Research Plan on Adjustment
for the 1990 Decennial Census
(Prepared: July 1984)

I. RESEARCH AREA:

Should the Bureau use statistical estimating techniques to adjust any of the data obtained in the 1990 census by the more traditional counting and self-enumeration techniques? If so, what characteristics of population and housing should be adjusted, and what geographic level should be adjusted?

In order to resolve these global issues, one must break them into separate sub-issues which are capable of being researched. For the purpose of this paper seven groupings will be used. These are:

A. How would adjustment affect critical uses of census data?

B. What is the legal and policy context for adjustment?

C. How can census coverage best be measured?

D. How can local area estimates of coverage best be made?

E. How should adjustment be implemented as part of the census process?

F. How should the adjusted figures be published and used?

G. What are the other implications of census adjustment?

The first part of this document discusses the background, meaning, and context of census adjustment. The reader who is already conversant with the problem is invited to skip to Section V: Detailed Discussion of Plans.

II. DEFINITION AND BACKGROUND:

Since the first census, there have been problems in accurately counting every person living in the United States. The resulting undercount is not new. What is new is the pressure to adjust the census data to correct for the undercount. The pressure results from the use of census data to distribute Federal funds, the economic pressures on local governments to obtain those funds, the use of census data for political redistricting, and the increasing emphasis on equity for minority groups. There is concern that direct census methods may never be successful at reducing the gap to negligible levels.

Ironically, the impetus for adjusting census data in part comes from the Census Bureau's evaluations of the completeness of coverage. Since the 1950 census, the Bureau has published estimates for the national undercount rate—that is the percentage of the

population that is not counted by the census. For the 1950 and 1960 census, the undercount for the nation was estimated to be about 3 percent. For 1970, the rate was near 2 percent. A preliminary estimate for 1980 is that less than one percent of the legally resident population was not counted.

However, these estimates do not reflect the differential undercount rates for different population groups. The Bureau has estimated that in 1970 blacks were undercounted at nearly an 8 percent rate compared to less than 2 percent rate for whites. The evidence for the 1980 census suggests an undercount rate of legal residents on the order of 5 percent for blacks and virtually zero for whites.

There are also disparities between the sexes, different age groups, and different geographical areas (regions, states, cities). Adjustment may be the only way to correct for these differential undercount rates.

Another reason for adjustment is to reduce cost. A simple enumeration can be conducted at reasonable cost, but will result in significant omissions. Efforts were made in 1980 to increase the proportion enumerated by introducing special procedures. These procedures added people, but were enormously expensive. It might be possible to achieve the same level of accuracy, but at a lower cost, by statistical adjustment. The issue becomes how best to allocate fixed resources to achieve the greatest accuracy.

III. LEVELS OF CENSUS ADJUSTMENT:

Adjusting the census can mean many different things, depending upon what is adjusted, when adjustment is made, at what geographic level adjustment is made, which characteristics are adjusted, and for what purposes the adjusted results are used. The section below on critical issues discusses the variations in some detail. However, it might help the reader to mention three of the principal uses and levels of adjustment.

A. Adjustment for Revenue Sharing
It is possible to adjust only the figures used for federal revenue sharing and other funds distribution. Estimates of total corrected population would be made for each revenue sharing area. These estimates would be clearly differentiated from the published census data and may not be a part of the same publication series. Separation of these estimates from the federal apportionment counts would ease time pressures on the adjustment process.

B. Adjustment for Apportionment and Redistricting
The official census counts provided to the President and to the States would be adjusted for census undercount. The requirements of this adjustment are much more stringent than for adjustment for revenue sharing. By law, apportionment figures are due by December 31, 1990 and redistricting figures by April 1, 1991. This means that the adjustment process, (which would probably include interviewing, matching, follow-up, estimation, tabulation, analysis, and local area estimation) would have to be completed in a matter of months. For redistricting, estimates would be required at the block level, and separate estimates probably would be needed by race and ethnic group. Any delay in the census evaluation also would delay the official counts. The unadjusted census counts also would be published.

C. Complete Census Adjustment
In a complete census adjustment, all figures would be corrected for undercount. This option requires that one build a model that predicts not only how many people

were missed but also predicts their characteristics. Further, the prediction would be done for all geographic areas and for all tables. Only limited results of the basic enumeration would be published, perhaps in a methodological appendix. The adjusted figures would be the census.

These three options do not exhaust the possibilities. Some would argue that one or the other is not really census adjustment. Others will argue that adjustment should correct for all known errors in census results, replacing census data with improved estimates whenever possible. However, reference to these options will help the reader to appreciate fully the critical issues involved.

IV. CRITICAL ISSUES:

This section raises the critical issues that will need work in order to make an informed decision on adjustment. By investigating the legal and policy issues and by testing and developing the measurement and statistical issues, the Census Bureau can understand how adjustment fits into 1990 Decennial Census plans.

A. How would adjustment affect critical uses of census data?

1. What is the effect on data used for apportionment and redistricting?

2. What are the data needs for Program Implementation and Funds Distribution (PIFD)?

   a) How are the census counts and other census variables used for PIFD?

      1) In what programs and at what geographic levels do other federal government agencies (e.g., HHS) use census data?

      2) Do states and cities use census data for PIFD? How? With what other variables?

      3) What other official uses of census data are there for PIFD?

   b) How would adjustment affect different funding formulas?

      1) Competitive Grants

      2) Threshold Programs

      3) Population Change Formula.

3. What are the other mandated uses of census data and how would these be affected by adjustment of census counts or characteristics?

4. What is the effect on other important uses of census data such as EEO and Affirmative Action?

B. What is the legal and policy context for adjustment?

1. Are adjusted counts acceptable/required for congressional reapportionment?

2. Are adjusted counts acceptable/required for state redistricting for the U.S. Congress or for drawing other election and political districts?

3. What constitutes "a reasonable estimate of the number of resident individuals not counted" for the purposes of Title 31 Section 6713 ("The Moynihan amendment")?

4. Under what circumstances would the Bureau consider adjusting the census?

5. How will the decision be made?

6. What measures of errors would be helpful in making these decisions?

Discussion:
Many of these issues will not be resolved by the Bureau, but instead will be the result of litigation and legislation. However, the Bureau needs to insure that any decision takes adequate account of limitations of data accuracy, timing etc.

C. **How can census coverage best be measured?**
Four major techniques can be used to derive estimates of population coverage. These are:

- Post enumeration surveys (PES)
- Reverse record check (RRC)
- Administrative record matches (ARM) (including composite lists)
- Demographic analysis (DA)

Research projects must seek not only to compare the techniques in general, but also to address the question of how different designs compare for different population groups. Therefore, the following questions must be asked for each method.

1. Is it possible to design the method so as to largely eliminate bias due to

   a) matching errors

   b) treatment of the nonresponse and other unresolved cases

   c) curbstoning in either the undercount or overcount samples

   d) correlation bias

   e) bias due to misclassification errors in the post-stratification variables

   f) bias due to spurious reporting of individuals, names, and characteristics in either the undercount or overcount sample

   g) bias due to incorrect reporting of April 1, 1990, addresses?

2. How do the residual biases vary by race, sex, ethnicity, age, and location?

3. What are the requirements for trained programmers, supervisors, and field staff?

4. What are the costs?

5. When can the results be available?

6. How does one best model nonresponse?

7. How can matching error be modeled?

8. Can one design interpenetration or other studies to measure the degree of matching error?

9. Can one design studies to measure the degree of other error?

10. What measures of error are helpful in evaluating these methods?

11. How can one design the sample so as to permit easy adjustment?

12. How can estimates best be combined?

13. What is the most effective combination of methods for particular demographic groups and for particular geographic areas?

14. What methods should be used in estimating gross overcoverage? How can these be combined with estimates of gross undercount to form estimates of net undercount?

15. How can the studies be designed to assist census planning by pinpointing the causes of undercount and suggesting alternative techniques to reduce the overall and differential coverage errors?

Discussion:

- Post Enumeration Surveys

A post enumeration survey uses a sample survey to re-enumerate the population. The Bureau has more experience with the post enumeration surveys than with any other method. Post enumeration surveys were conducted in 1950 and 1960. The 1970 CPS-Census match was essentially a PES, but was limited in scope. As part of the 1980 census work, the Bureau conducted PES pretests in 1977 and 1978 and conducted the 1980 Post Enumeration Program. Because of this experience, the Bureau knows more about the technical issues and inherent limitations of a PES than about the problems of the other methods.

Among the technical issues are the following:

Whether to use a block sample or a list sample.

Whether to conduct an independent evaluation survey or to utilize an on-going survey such as the CPS.

How to measure duplicate enumerations and curbstoned cases in both the census and the evaluation survey.

How to gain sufficient information at the first interview so that cases can be resolved without follow-up.

How to minimize nonresponse, ungeocodable, and other unresolved categories.

How to shorten the time requirements to complete the evaluation.

How to sample people who moved between the time of the census and time of the PES (i.e. PES-A vs PES-B).

How to define the extent of search and sufficient information to decide that a person was not enumerated.

How computerized matching can be used to increase the efficiency, speed, and accuracy of the method.

The inherent limitations of the method are that some groups are almost always missed by any household interview, including both the census and the Post Enumeration Survey. This lack of independence leads to a downward bias of the estimated population. Careful interviewing and thoughtful procedures can minimize this group, but a core of "invisible" people will always exist. Using a nonhousehold sample (presumably more independent) to estimate the undercount (e.g., administrative record match) may be better than a household survey in measuring the "invisible" people.

Reverse Record Check

The reverse record check (RRC) is a match from the previous census to the current census. The previous census is supplemented by a sample of births, immigrants, and persons known to be missed in the previous census. A reverse record check has several advantages. The previous census and supplementary records constitute a more complete frame than can be achieved through a sample survey, such as a post enumeration survey. The separation in time is thought to increase the independence between the census and the evaluation frame. The advantages are offset by the loss of both completeness and independence due to failure to trace sample people. This loss presumably increases with time.

A reverse record check utilizing the previous census is only a specific case of the idea of a longitudinal evaluation study. One could also conduct an independent survey sometime before the census and match it to the census or use an existing frame such as a CPS. This approach could be called a pre-census survey. A pre-census survey loses the advantage of having a more complete sampling frame. However, it would be available for early matching and the interval for tracing is flexible: one can choose an "optimal" time for the survey. Other than that, it has the same problems of the PES.

- Administrative Records

Administrative records can be used in any of these ways. They can be matched to the census singly and the aggregat results combined with other estimates according to their variances. Several lists can be unduplicated, combined into a composite list and then matched to the census. Several lists can be matched to each other and to the census to form a multiple system estimate. Each approach has its own theoretical, operational and public preception difficulties. Specific national, state and local lists must be identified and investigates with respect to their coverage and accuracy.

- Demographic Analysis

Demographic analysis relies on various types of demographic data and the constancy of certain demographic relationships and identities. Since demographic analysis uses births, deaths, migration and Medicare data which do not depend on the current census counts, the demographic estimates are a valuable adjunct to matching studies, providing survey control totals and parameters such as sex ratios which serve as powerful checks on the adequacy of survey estimates.

Limitations of the method of demographic analysis as a tool for estimating census coverage are as follows:

Illegal Aliens—Demographic Analysis was designed to produce estimates of the legally resident population. Estimates of illegal residents are needed to make comparisons with the census. The problem of including illegal aliens in the demographic estimates has not been completely solved.

Immigration and Emigration—Data sources to determine net legal movement have traditionally been weaker than the data sources for births and deaths. This problem includes not only net immigration by foreign nationals, but also in and out movement by U.S. nationals from Puerto Rico and other outlying areas. The quality of migration data is not under the Census Bureau's control.

Underreporting—The quality of data sources other than migration is high but needs to be checked periodically (e.g., birth registration completeness, Medicare enrollment completeness). Last check occurred some time ago and may need updating.

Bias and Variation—No method has been developed which can estimate or measure the mean squared error of the Demographic Analysis estimates.

- Overcount measurement

Estimates of gross overcoverage must be made in order to balance the estimates of net undercount. One issue is how to define and measure gross overcount so as to balance the gross omissions estimates. This means designing a sample to fit the omissions sample designs. One must also design a questionnaire and interview procedures so as to minimize nonresponse and unresolved cases in this sample as well. The procedure must be able to distinguish between a fictitious enumeration and one that refers to a person who is hard to find.

- Causes of Undercount
  The pressure to provide measures for adjustment must not be allowed to blind us to the need to learn more about the causes of undercount. The original and ongoing purpose of census coverage evaluation is to guide in the planning of the next census. A good evaluation program will increase knowledge of the causes of undercount and overcount, rather than merely supplying the data for adjustment.

- Housing Unit Coverage
  The completeness of housing unit coverage must also be measured. A post enumeration survey lends itself naturally to proving housing unit coverage estimates. Other techniques would have to be specially designed to provide these estimates, or supplemented by a special survey. Complete census adjustment would have to account for housing units as well as households and individuals.

D. **How can local area estimates of coverage best be made?**

1. Method (techniques) of adjustment

   There are at least four separate techniques that can be used to make adjustment for small areas:

   - Synthetic techniques
   - Regression techniques
   - Bayesian techniques
   - Demographic analysis estimates for states.

   These techniques need to be developed with regard to undercount estimation and the anticipated level of information available for development and assessment.

2. Design for the development of adjustment methods
   a) How should the studies listed elsewhere be designed so as to provide the information needed to study and develop adjustment methods?

   b) How should the research studies listed elsewhere be designed to enable the timely implementation of these techniques in a census situation?

3. Application of methods

   a) How would the techniques in 1. above be applied: to make estimates at a certain geographic level and such adjustments allocated downward (e.g., synthetic technique), or at the lowest geographic levels and summed up to county and state levels (regression technique)?

   b) Are there combinations of adjustment techniques that would be most effective?

   c) Are there different techniques that could be used in different areas or at different levels?

d) Are there other techniques not listed above that should be considered for use in adjustment of the census for undercount?

4. Evaluation of the evaluation

a) How can the local area estimates of coverage be evaluated?

b) How would the Bureau identify when the data would not support an adjustment, or more importantly a certain level of adjustment such as the adjustment for individual revenue sharing areas? Can reasonable error models be constructed to guide the choice?

c) What are appropriate measures of error (loss function or yardsticks) to compare alternative estimates? How do these measures relate to considerations of equity?

5. Block level adjustments

There is the technical issue of how to develop weighting algorithms or imputing algorithms for the generation of block estimates and adjustments. This issue is extremely important under variant C. in section III.

a) What variables are relevant to the weighting and imputation described elsewhere?

b) Would the weighting or imputation be used to make adjustments for both persons and housing units, or only persons? This depends on the variant of census adjustment in section III.

c) At the block level, how does one deal with biases due to small samples or small populations receiving an adjustment?

E. **How should adjustment be implemented as part of the census process?**

1. Integration of the adjustment into census processing

a) If an early adjustment is required, how would census processing have to change to guarantee estimates can be made in the time required?

b) How can census procedures be changed to accommodate estimation of the undercount without compromising the integrity of the census counts or the accuracy of the undercount estimates?

2. 1990 census district office processing

a) If an adjustment is required, how would the measurement of the undercount be incorporated into district office processing?

b) Would certain Census Bureau operations (e.g., local review) have to be structured around providing undercount information before or while the census data were being processed?

c) Would district office or regional census center operations be restructured to assist in the measurement of the undercount?

3. Treatment of overcounts

   Should adjustments allow the possibility of overcounts (and thus the diminution of a population count)?

4. Arithmetic consistency of estimates

   a) Must all adjusted numbers be arithmetically consistent (e.g. county estimates sum to state estimates)?

   b) Are there other types of arithmetic consistency that should be considered in the adjustment (e.g. maintenance of correlational structures)?

5. Other characteristics

   a) If an adjustment is made only for persons and not for other statistics, how will this be reconciled in the publications (tapes...)?

   b) If adjustments are made for housing units (HU's), will they be made for only occupied HU's, or other types of HU's as well?

   c) Can other statistics be adjusted?

   d) Will there be a separate adjustment for each variable?

   e) If there are separate adjustments for each variable (e.g., persons, income, race), how are these published and explained to users?

   f) How are rates and other derived measures affected?

Discussion:
The Bureau has always conducted the census as a head count. Census evaluation has not had the time pressures that accompanies other census operations. If the Bureau is to provide corrected state totals by December 31, 1990 and corrected local area counts by April 1, 1991, evaluations must take place at the same time as the processing and tabulations.

Integrating evaluations and counting is not without risk. Beginning the evaluations work (e.g., interviewing and matching) too early can compromise its independence, and therefore its value. Any delay or uncertainty in the evaluation and adjustment program becomes a delay and uncertainty in the census itself. In 1980, the Bureau had great difficulty producing preliminary PEP estimates by October 1981. Even so, no acceptable final estimates are in sight.

What will be needed is a strong, early commitment to incorporate adjustment at all stages of census work after the completion of all field activity. It will require a commitment to close out the field work for the counting operation early enough to allow evaluation field work, follow-up matching, tabulation and analysis to be completed in time to be incorporated into the official census counts.

F. **How should the adjusted figures be published and used?**

1. A choice between adjustment and the census counts

   a) Is the adjustment to be the official census count, with no other numbers published?

   b) Is the adjustment a set of unofficial estimates, not to be published (though later figures like intercensal estimates would be), and the basic enumeration is considered the official census?

   c) Is the adjustment the set of official estimates, but census figures are published also?

   d) Is the choice of figures to be published some combination of a)-c) above?

2. Publications of the 1990 census

   a) How would publications (public use files, STF's) be affected by the adjustment? Would their production be delayed?

   b) Would procedures for providing publications (public use files, STF's) need to be changed to allow for the adjustment?

3. If the census is fully adjusted, will researchers and other users have access to the unadjusted data?

G. **What are the other implications of census adjustment?**

1. Current surveys

   a) How would the adjustment of the census affect the use of the census as a sampling frame?

   b) How would census counts be used for weighting of the data, or for ratio or regression estimation?

   c) Will an adjustment cause discontinuities in the time series the Bureau produces (e.g. unemployment)?

2. Other implications not yet evident

3. Preparation for the year 2000 census

   a) Will adjustment affect cooperation in later censuses because both individuals and groups will see the actual count as less important?

   b) What longer term research should be considered?

V. DETAILED DISCUSSION OF PLANS

Decisions

Four critical decisions are needed in the planning for adjustment. In September 1985 the Bureau will announce its decision on the role of adjustment. This decision will not be

whether or not to adjust, but rather to what extent the Bureau will relv on adjustment in the overall planning.

The next important date comes at the beginning of 1987. The main design for 1990 will have to be decided in order to allow it to be incorporated into the 1988 Dress Rehearsal.

In the fall of 1988, the Bureau will make a tentative decision on adjustment. This decision will be discussed before Congress, at stakeholders conferences, and with other appropriate groups. This process will lead to an official announcement in January 1989 on the plans and standards for adjustment. Exactly what these plans and standards are will be the result of the decade's research.

Even this final decision may not be a yes/no decision on whether to adjust. Rather, the decision might specify under exactly what conditions the Bureau would or would not adjust. These conditions would relate to the nature of the 1990 undercount and the coverage measurement process. The decision would specify as precisely as possible the nature of the decision and the data upon which it would be based. For example, the Bureau might announce that if the differential undercount conformed to certain specified conditions, the census would be adjusted. If the results of the enumeration and evaluation were radically different from expected, no adjustment would be made.

This section provides a general framework for resolving the critical issues listed in Section IV. This section proposes the necessary research to help answer the adjustment issues and the general timing for testing major issues in the pretests for the 1990 census.

A. How would adjustment affect critical uses of census data?

1. Research on fund allocation

Fund allocation programs are extraordinarily complicated, and it is difficult to assess the effect of undercount on such programs in general. Rather, one must assess the effects on a program by program basis. The project would:

- devise a list of federal programs which dispense money on the basis of census results
- assess the effects of undercount upon the programs
- assess the effects of adjustment on the programs
- implement explicitly the adjustment for a set of geographical areas, with the goal of measuring the overall effect of undercount adjustment from all programs on these areas.

The project would conclude with a final report summarizing all of the findings.

2. Develop alternative adjustment scenarios

This project will develop a number of explicit adjustment scenarios. The scenarios will go into enough detail to give a user an idea of the implications for their uses, but will not go into as much detail as might be needed by a computer programmer. The final report from this project will describe clearly each of the scenarios. These scenarios will be input to projects A.3, A.4, and A.6 listed in section A. In these latter projects, we will solicit comments on the scenarios from the various users. The present project is also related to project D.3.

3. Research on PL 94-171 requirements

This project will assess the effect of a census adjustment on the PL 94-171 program and to solicit comments from those responsible for implementing the program. The project will:

- develop a list of contacts in each of the states
- obtain a written statement from each of the state contacts commenting on the adjustment scenario developed in project A.2.

This project will conclude with a final report summarizing both the reactions of the various state users to adjustment and the effects of adjustment on the PL 94-171 program.

4. Effects on other census users

This project identifies other key government and nongovernment users of decennial census data. The project will:

- identify other key government and nongovernment users
- contact each of the users
- solicit written comments from the users regarding their reactions to the adjustment scenarios

The project will conclude with a final report summarizing both the effects of adjustment on the various key uses of census data and the reactions of the users to the scenarios.

5. Research on apportionment

The main objective of this project is to explore further the effects on congressional apportionment of census undercount and undercount adjustment. Much work has already been done in this area, including the work reported by Gilford, Causey, and others, and the work reported in the Wolter affidavit. This project will undertake additional work in regards to effects on apportionment. All findings will be summarized in a final written report.

6. Research on adjusting housing data

Most of the Census Bureau's efforts and attention regarding the 1990 undercount have been directed at the problem of population undercount and possible adjustments to make up for the missed population. But coverage errors also affect census housing data, and these errors have received little attention either within the Census Bureau, the court cases or elsewhere. This project will look into the housing issues and the effects of undercount adjustments on housing data. In particular, the project will:

- identify the main uses of housing data, particularly those that involve funds distribution
- assess the impact of undercount upon the uses
- contact users of housing data, e.g., HUD, soliciting comments on the adjustment scenarios developed in project A.2.

This project will conclude with a final report and recommendations.

7. Stakeholders' conference on adjustment
This project will:

- consolidate the results of projects A.1 through A.6 and produce an overall report addressing the question "How would adjustment affect critical uses of census data?"
- distribute the report to a select group of conference participants
- formulate critical questions, issues, and options on the effect of adjustment on the critical uses of census data
- convene a conference wherein the participants discuss the issues and options that have been formulated.

It is hoped that the conference will achieve a synthesis of opinion on the adjustment issue and the effects of undercount and adjustment on the uses of data. A final written report will be issued following the conference. This report will represent the main input from users to the final decision about whether or not to adjust the 1990 census.

B. **What is the legal and policy context for adjustment?**

1. Establish task force to determine legality of adjustment
The Task Force on legal issues has been established to investigate the legal issues of adjustment. The Task Force consists of representatives from Department of Commerce's Office of General Council, Census Bureau, and Department of Justice. Its purpose is to document the legal requirements or constraints of adjustment.

2. Hold senior staff conference on policy position
This conference will decide the Bureau's official position on adjustment. It will decide whether to recommend any changes in law.

3. Review 1980 criteria for adjustment
Implicit in the Bureau's defense in the New York and Detroit cases was a policy as to the statistical standards for adjustment. This project reviews the testimony, affidavits, and documents of the court case and makes explicit the policy standards for adjustment.

4. Review alternative solutions to the problem of differential undercount
This activity assumes that adjustment is a permissible census taking technique. It is then judged against the other available techniques to eliminate differential coverage such as coverage improvement, sampling, resource allocations. All techniques will be judged based upon data quality, cost, and probability of success using the data available. A decision will be made as to their relative role in census planning. This analysis should be conducted again later in the decade when pretest results are available.

5. Develop mechanisms for making decisions about adjustment
This activity establishes and makes explicit the decision-making process itself, inlcuding:

- Who decides
- When decision is to be made
- When decision is announced
- Criteria to be used.

6. Develop alternative standards (yardsticks) to evaluate adjusted vs. unadjusted population figures

This activity develops conceptual standards to measure when the level and distribution of population is improved by any given adjustment. Without knowing the truth, such a standard can never be applied. However, through modeling and simulation, such a standard can help frame and guide the decision.

C. **How can census coverage best be measured?**

1. Review results from 1980 evaluation studies

This project determines the critical weaknesses in the 1980 PEP that must be corrected in future undercount work. A good starting point is the evidence introduced in the New York law suit. Another resource is to use the PEP to prepare and conduct further tabulations and analyses that identify, quantify, and highlight the key weaknesses in the 1980 PEP.

The final product of this project is a written report. This document will provide an important planning tool for future undercount and adjustment research,and a means for refining other projects mentioned in this research plan and for formulating additional projects to meet key weaknesses in undercount programs.

2. Complete research on forward trace study and retrospective match study

In 1981, the Bureau started two research projects designed to develop and test the Reverse Record Check approach for 1990. The CPS-Census Retrospective Match was a match of the 1977 Current Population Survey to the 1980 Census. The Forward Trace Study used year-by-year tracing. It includes the construction of a sample using the 1980 census, the 1980 PEP missed file, and immigrant records.

A final written report will provide a complete description of the Forward Trace Study and the Retrospective Match Study and make recommendations about the use of a reverse record check for 1990.

3. Develop demographic analysis estimates under alternate assumptions

By varying the assumptions of demographic analysis, alternate estimates of the U.S. population will be produced. The alternate estimates will be based on reasonable assumptions that are consistent with the observed data. The result will be a better understanding of the precision and robustness of the demographic analysis estimates.

4. Develop error models for demographic analysis estimates

This activity will provide explicit formulas for computing the demographic analysis estimates. By varying the input values, it should be possible to construct a measure of uncertainty which can be interpreted as a variance or mean-square error. A written report will specify the techniques used, the possible errors introduced by each uncertainty and a tentative measure of error.

This activity serves two important purposes. First, it is necessary for any possible use of the demographic analysis estimates for adjustment. Second, the results will show how much the final estimates can be improved by

improving the input data. For example, only by knowing how much possible error is introduced by the migration data or birth registration correction is it possible to measure the value of a survey to measure migration or birth registration completeness.

5. Design and implement a 1985 PES Pretest

The 1985 pretest will consist of a sample of blocks that will be completely listed and matched to the pretest census. The matching will be done by computer and followup will be limited to the computer nonmatches after some clerical review.

The research for computer matching will need at least 10,000 persons. A sample of 200 census blocks will yield sufficient persons to conduct the computer matching research.

The matching will be a two-way match between the PES and the census. The PES persons who do not match to the census will estimate the census omissions in the census and the census persons who were not listed in the PES will be checked to estimate the erroneous enumerations. The difference will estimate the net undercount.

Another purpose of the research is to develop methods to minimize follow-up. Only limited number of non-matches will be sent to followup in order to clarify problems in the original questionnaire design. Also we will learn about making combined overcount and undercount estimates as we may also be able to obtain good estimates for duplicates and curbstones.

6. 1985 Pretest Hard to Count Study

One weakness of previous census evaluation studies such as the Post Enumeration Program has been the failure to include a sufficient proportion of certain hard-to-count groups in the evaluation sample. This leads to a systematic underestimate of the corrected population and the estimated undercount. Although this bias is certainly present to some extent in all groups, evidence suggests that it is strongest for young adult males, ages 18 to 40.

In the hard-to-count study, a sample of males who are 18 to 40 years of age will be selected from the 1983 Social Security and IRS files. Other possible sources for sampling young males are:

- Unemployment records
- Immigration and Naturalization Service files on recent immigrants
- Comprehensive Employment and Training Act files
- Draft Registration files
- Driver License files

7. Develop error models for matching studies and methods to measure variables

This project will construct an error model for the principal matching studies which includes sampling error, matching error, response error, and model error. It will also develop methods to measure these errors. These methods will be implemented in later pretests.

8. Design an adjustment prototype based on multiplicity (network sampling)
Multiplicity or network sampling offers the possibility of including hard-to-count groups in the evaluation sample. In such samples, parents are asked to provide names and addresses of children and vice versa. A prototype using multiplicity should be designed that would address omissions, erroneous inclusions and the data needs for small area estimation. Multiplicity may or may not be pretested depending upon such factors as cost, feasibility, manpower, etc.

9. Develop and issue public use tape based on 1980 studies
The results from the 1980 PEP will be documented and placed on a standard public use tape for use by outside scholars. The tape will include those variables which can be disclosed and have greatest value.

10. Research methods of combining demographic analysis and case-by-case studies
This project follows C.4 and C.7 which develop error models for each set of estimates. Using both demography and statistics, this activity will produce a report outlining one or more ways to combine the result.

11. Develop alternate estimation models for matching studies
Since 1960 the Bureau has used a model based upon an assumption of conditional independence between the census and the evaluation studies. Other models exist including those suggested by Horvitz, Ericksen and Kadane, and in the literature on wildlife studies. This activity will study and suggest which can be adapted for use in 1990. The results potentially can be studied in a pretest site.

12. Design a pre-enumeration study to be conducted as part of the 1986 Pretests
A survey will be conducted at one 1986 pretest site some six months before census day. A pre-enumeration survey will examine the independence assumption and whether this technique expedites the evaluation process.

13. Design and implement a test of PES and ARM in 1986 Pretests
Both a PES and ARM should be designed and implemented in 1986 pretests.

14. Pilot Study of "Card Approach"
In order to expedite matching, one suggestion is to send cards along with the census questionnaire. This card would contain all the needed information to match the household and would be retained by the household. The post enumeration survey interviewer would list the information from the card. The interviewer for the E-sample would also verify the information on the card. This approach is useful in rural areas that have few city type addresses and/or named roads. A small experiment should be conducted in either 1986 or 1987.

15. Identify best post-stratification variables for dual system estimation
The dual system estimate assumes independence between the census and the PES. Post stratification relaxes the assumption of independence so that independence is assumed only within the post strata. By identifying variables closely correlated with misses in both systems, a better estimate can be formed. Most systems have used age, race, sex and geographic location. This study will re-estimate the 1980 PEP and the pretests to see if other variables work better.

16. Surveys to supply data needs for demographic analysis
Based on the results of the alternate assumptions studies (C.3) and the error models (C.4), it will be possible to identify where additional information is needed to improve the demographic analysis estimates. Such studies could a birth registration completeness test, a study of undocumented immigration (legal and illegal), a study of emigration, or other studies. These studies should be considered for funding to the extent they narrow the range of uncertainty of the demographic estimates or help produce estimates for groups such as Asians or Hispanics.

17. Design and implement 1987 Pretest
Since the results of the 1987 Pretest will not be available until after the 1988 Dress Rehearsal is designed, the test will be limited to important but not central issues.

One goal will be to test sample designs which may facilitate estimation. By deeply stratifying the blocks, a simple synthetic estimate may be possible within strata.

We will use this opportunity to test special procedures for several special population groups such as military, institutional, and colleges. This study would include sampling, interviewing, matching and reconciliation. If possible, this test would include a pilot PES for Puerto Rico.

18. Design and implement 1988 dress rehearsal
Using the information obtained from the 85 and 86 pretests, an adjustment procedure will be designed and implemented in the 1988 dress rehearsal. The adjustment procedure will be the main procedure that is being considered for the 1990 census.

19. Final design for 1990 census evaluation
For 1990 census evaluation, decisions will be reached that will determine which methods will be used and how they will be designed.

20. Design experimental coverage program for testing 2000 census
Each census should be the pretest for the next. The research program will uncover new issues and new methods to be tested in 1990 for possible implementation with the 2000 census.

D. **How can local area estimates of coverage best be made?**

1. Exploratory analysis of 1980 data
In order to better understand the undercount and its effect on local areas, we need to examine the 1980 PEP. These results will assist in identifying predictor variables and their effect when applied to local areas. The analysis at different levels of geography will show potential faults of the estimation methodologies and where improvements can be made and tested. Also, different models may hold for individuals in counted housing units and individuals in missed housing units. This can be tested with the 1980 PEP data.

2. Pretest regression and other methods
By applying methods of local area adjustment to pretests we can learn of

differences from using a different sample design than used in 1980. Also different correlates of undercount may arise from this analysis. Using the pretests and the dress rehearsal we can test the models developed and studied from the 1980 data.

3. Design adjustment prototypes
Three small area adjustment prototypes will be designed and implemented on the 1980 census data. A detailed explanation of the three adjustments will be given as well as the adjusted counts and graphical displays. The adjustment prototypes will be circulated around the Bureau and to the public. Reactions to the adjustment prototypes will be solicited.

4. Decide on 1990 design
In order to implement a 1990 adjustment the basic implementation of local area estimation needs to be established. The basic framework needs to include the explanatory (or carrier) variables, the functional form of adjustment methodology, and tests of validity need to be established to determine whether adjustment is necessary or at what level adjustment can be justified by the data.

5. Develop and test 1990 software
Computer software that will be used to fit models, check their fit, and examine unusual data points needs to be developed for the computer. The deadlines of producing 1990 census numbers means all local area estimates will be calculated in a very short time. Consequently computer software needed for a complete analysis should be developed and tested on the computer before the data becomes available.

6. Propose any changes to 1990 questionnaire
The variables needed to implement local area adjustment may not currently be available on the short form. Because of the delays generally associated with processing the long form questionnaires, it may be necessary to put these questions on the short form questionnaire. Only then will local area adjustment be able to meet the necessary time deadlines.

7. Identify necessary census processing schedule
To insure all necessary work is completed, a schedule should be developed. Then delays in different phases of the census can be translated into its effects on producing local estimates. This will help those involved understand how much time they have to perform the necessary work and whether the work can be done under 1990 restrictions.

8. Develop overall error models
In order to better understand, compare and develop improvements on small area adjustments, explicit error models will be developed. The assumptions needed to implement the model will also be listed. The model should include:

- Sampling error
- Nonresponse error
- Matching error
- Model error
- Regression error.

This work will help guide and decide the best small area procedure to use for census adjustment.

E. **How should adjustment be implemented as part of the census?**

1. Develop household and housing unit models
This activity will analyse the results of the PEP and the early pretests to build a model which categorizes census misses by principal type, persons missed in counted housing units, and persons missed in missed housing units. This model will take into account the causes and correlates of each type of miss in order to be able to adjust the census consistently for persons, households, and housing units.

2. Simulate adjustment on 1980 census tabulations
The 1980 census results for a few local areas will be reprocessed and adjusted. Each set of adjustments will use a single estimate of undercount based on the 1980 PEP. The adjustment will be implemented in three ways. One method will be an imputation scheme using the 1980 PEP misses as the donor pool. Another method will use the Hot Deck procedure using the locally enumerated people. Finally, a weight adjustment will be used. The results will be analysed to assess the characteristics and feasibility of each adjustment scheme. This project assumes a single set of small area estimates has been made and concentrates on the effects upon characteristics.

3. Conduct 1986 pretest on adjustment
This pretest begins the process of putting the pieces together and integrating adjustment into the census process. Since much of the work on measurement and estimation will not be completed, this test will not be on the accuracy of adjustment, but will highlight the requirements for any serious planning for adjustment. A working group is being formed to begin specific design work.

4. Develop and test alternative treatments of measured overcounts
Net overcounts must be treated as part of the adjustment process. Just as activity E.1 and E.2 looked at models to add people, this activity seeks models to delete erroneous enumerations. The causes and correlates of erroneous enumerations will be studied based on the 1980 PEP and the early pretests. Models will be built and tested.

5. Analyse 1986 adjustment tabulations for consistency and reasonableness
The results of the 1986 pretest will be analyzed to determine the consistency and reasonableness of the adjustment.

F. **How should the adjusted figures be published and used?**

1. Prepare alternative publication stubs to reflect adjustment
This activity will construct a range of publication stubs to report the adjustment. These stubs will be circulated to principal users for comment and then incorporated into the 1986 pretest.

2. Analyse the tabulations produced from the 1986 pretest
The results from the pretests will be studied and improved for use in the Dress Rehearsal.

G. **What are other implications of census adjustment?**

1. Solicit input from current surveys and other users
This activity seeks to identify all effects of adjustment. Will adjustment pose new problems for current surveys redesign, population projection, or other program other than those already discussed and handled under the term "error-of-closure". Any special requirements identified will help guide the research on adjustment so as to minimize the effect.

VI. BASIC ASSUMPTIONS AND POTENTIAL OBSTACLES:

A. Basic Assumptions

1. The new matching system is developed and has appropriate error rates and timing and cost characteristics

2. Adequate staff in the various divisions are assigned to carry out the proposed work.

3. Adequate resources are devoted to conduct pretests.

4. The true differential undercount will be at least as large as it was in the 1980 Census.

B. Potential Obstacles

Many of the critical issues above entail determining whether potential obstacles can be overcome. There are other potential obstacles to adjustment that are beyond the scope of this research, such as:

1. A system that allows rapid, accurate, and inexpensive geocoding of addresses is not in place.

2. Address registers and census questionnaires (or their 1990 equivalent) are not maintained in a manner that allows rapid, accurate, and inexpensive access.

3. Census enumeration field work continues throughout the summer, which would delay evaluation interviewing and matching until after the critical dates.

4. Because of lack of public cooperation or because of lack of field control, a high proportion of census enumerations lack name, adequate address, characteristics or are listed in an incorrect area.

5. Public opinion and reaction is such that record linkage is viewed unfavorably and could jeopardize the counting itself.

6. The nature of the 1990 undercount is sufficiently different from that of 1980, so that all plans are inappropriate.

7. The Bureau is unable to recruit and retain an adequate technical staff (mathematical statisticians, programmers, etc.).

8. The problem may be irresolvable, i.e., no amount of testing can guarantee that the estimated results are accurate.

VII. CRITICAL DECISIONS AND REQUIRED ACTION:

Critical Decisions

1. Decision on role of adjustment	9/85
2. Decision on main design for 1990	1/87
3. Tentative decision on adjustment	10/88
4. Plans and standards for adjustment officially announced	1/89

Required Action	Participating Divisions	Dates Start	Dates Complete
A. How would adjustment affect critical uses of census data?			
1. Research on fund allocation	POP, PPDO	--	9/86
2. Develop alternative adjustment scenarios	SRD	4/85	9/86
3. Research on PL 94-171 requirements	DUSD	10/84	6/85
4. Effects on other census users	DUSD	--	6/85
5. Research on apportionment	PPDO, SRD	10/85	6/85
6. Research on adjusting housing data	POP, HOUS	4/85	9/86
7. Stakeholders' conference on adjustment	DUSD		10/87

Required Action	Participating Divisions	Dates Start	Dates Complete
B. What is the legal and policy context for adjustment?			
1. Establish task force to determine legality of adjustment	Commerce, PPDO, Justice		6/84
2. Hold senior staff conference on policy position	SRD		7/85
3. Review 1980 criteria for adjustment	SRD		12/84
4. Review alternative solutions to problem of differential undercount	DIR, DPLD, SRD, SMD		9/85
5. Develop mechanisms for making decisions about adjustment	PPDO		10/87
6. Develop alternative standard (yardsticks) to evaluate adjusted versus unadjusted population figures	SRD, Outside experts	6/84	10/87
C. How can census coverage best be measured?			
1. Review results from 1980 evaluation studies	SRD	--	9/85
2. Complete research on forward trace study and retrospective match study	SRD		9/85
3. Develop demographic analysis estimates under alternate assumption	POP, SRD, Outside Experts		9/85
4. Develop error models for demographic analysis estimates	POP, SRD		9/85
5. Design and implement a 1985 PES pretest	SRD, CMSR, DOD, DPLD, SMD, FLD	5/84	6/86
6. 1985 pretest hard-to-count study	DPLD, SRD, SMD	5/84	6/86

Required Action	Participating Divisions	Dates Start	Dates Complete
7. Develop error models for matching studies and methods to measure variables	SRD, SMD Outside Groups		9/88
8. Design an adjustment prototype based on multiplicity (network sampling)	Outside Groups	11/84	9/85
9. Develop and issue public use tape based on 1980 studies	SMD		9/85
10. Research methods of combining demographic analysis and case-by-case studies	SRD		6/87
11. Develop alternate estimation models for matching studies	SRD Outside Experts		6/87
12. Design a pre-enumeration study to be conducted as part of 1986 pretests	SRD	1/85	4/87
13. Design and implement a test of a PES and ARM in 1986 Pretests	CMSR, DPLD, SMD		
14. Pilot study of "Card Approach"	DPLD, CMSR	1/85	12/86
15. Identify best post-stratification variables for dual system estimation	SRD, SMD	--	1/87
16. Surveys to supply data needs for demographic analysis	POP, DSD	10/85	9/87
17. Design and implement 1987 Pretest	SRD, SMD, DPLD	4/86	4/88
18. Design and implement 1988 dress rehearsal	SRD, SMD, DPLD	1/87	4/89
19. Final design for 1990 Census evaluation	SRD, SMD, DPLD	1/88	1/89
20. Design experimental coverage program for testing 2000 census	SRD	10/88	12/89

Required Action	Participating Divisions	Dates Start	Dates Complete
D. How can local area estimates of coverage best be made?			
1. Exploratory analysis of 1980 data	SRD, Outside Experts		4/85
2. Pretest regression and other methods	SRD		
3. Design adjustment prototypes	SRD, SMD		9/86
4. Decide on 1990 design	SRD, DIR, SMD, POP		1/89
5. Develop and test 1990 Software	SRD, DPLD		1/90
6. Propose any changes to 1990 questionnaire	SRD, DOD		1/87
7. Identify necessary census processing schedule	DPLD, SRD	10/8	4/85
8. Develop overall error models	SRD, Outside Experts	10/84	9/88
E. How should adjustment be implemented as part of the census?			
1. Develop household and housing unit models	HOUS, SRD	10/84	9/86
2. Simulate adjustment on 1980 census tabulations	DPLD, POP, HOUS, SRD	9/84	10/85
3. Conduct 1986 pretest on adjustment	DPLD, SMD, SRD	10/84	4/86
4. Develop and test alternative treatments of measured overcounts	DPLD, POP, SRD	10/86	9/87
5. Analyse 1986 adjustment tabulations for consistency and reasonableness	POP, HOUS	10/86	9/87

Required Action	Participating Divisions	Dates Start	Dates Complete
F. How should the adjusted figures be published and used?			
1. Prepare alternative publication stubs to reflect adjustment	POP	10/85	9/86
2. Analyse the tabulations produced from the 1986 pretest	POP, DUSD	10/86	9/87
G. What are other implications of census adjustment?			
1. Solicit input from current survey and other users	SMD, DUSD, NCHS, AVRHS	10/86	9/87