OVERSIGHT OF THE 2000 CENSUS: ACCURACY AND COVERAGE EVALUATION [ACE]—STILL MORE QUESTIONS THAN ANSWERS

HEARING

BEFORE THE

SUBCOMMITTEE ON THE CENSUS

OF THE

COMMITTEE ON
GOVERNMENT REFORM

HOUSE OF REPRESENTATIVES

ONE HUNDRED SIXTH CONGRESS

SECOND SESSION

MAY 19, 2000

Serial No. 106–207

Printed for the use of the Committee on Government Reform

Available via the World Wide Web: http://www.gpo.gov/congress/house
http://www.house.gov/reform

U.S. GOVERNMENT PRINTING OFFICE

71–389 DTP WASHINGTON : 2001

For sale by the Superintendent of Documents, U.S. Government Printing Office
Internet: bookstore.gpo.gov Phone: (202) 512–1800 Fax: (202) 512–2250
Mail: Stop SSOP, Washington, DC 20402–0001

COMMITTEE ON GOVERNMENT REFORM

DAN BURTON, Indiana, *Chairman*

BENJAMIN A. GILMAN, New York
CONSTANCE A. MORELLA, Maryland
CHRISTOPHER SHAYS, Connecticut
ILEANA ROS-LEHTINEN, Florida
JOHN M. MCHUGH, New York
STEPHEN HORN, California
JOHN L. MICA, Florida
THOMAS M. DAVIS, Virginia
DAVID M. MCINTOSH, Indiana
MARK E. SOUDER, Indiana
JOE SCARBOROUGH, Florida
STEVEN C. LATOURETTE, Ohio
MARSHALL "MARK" SANFORD, South Carolina
BOB BARR, Georgia
DAN MILLER, Florida
ASA HUTCHINSON, Arkansas
LEE TERRY, Nebraska
JUDY BIGGERT, Illinois
GREG WALDEN, Oregon
DOUG OSE, California
PAUL RYAN, Wisconsin
HELEN CHENOWETH-HAGE, Idaho
DAVID VITTER, Louisiana

HENRY A. WAXMAN, California
TOM LANTOS, California
ROBERT E. WISE, JR., West Virginia
MAJOR R. OWENS, New York
EDOLPHUS TOWNS, New York
PAUL E. KANJORSKI, Pennsylvania
PATSY T. MINK, Hawaii
CAROLYN B. MALONEY, New York
ELEANOR HOLMES NORTON, Washington, DC
CHAKA FATTAH, Pennsylvania
ELIJAH E. CUMMINGS, Maryland
DENNIS J. KUCINICH, Ohio
ROD R. BLAGOJEVICH, Illinois
DANNY K. DAVIS, Illinois
JOHN F. TIERNEY, Massachusetts
JIM TURNER, Texas
THOMAS H. ALLEN, Maine
HAROLD E. FORD, JR., Tennessee
JANICE D. SCHAKOWSKY, Illinois

BERNARD SANDERS, Vermont (Independent)

KEVIN BINGER, *Staff Director*
DANIEL R. MOLL, *Deputy Staff Director*
DAVID A. KASS, *Deputy Counsel and Parliamentarian*
LISA SMITH ARAFUNE, *Chief Clerk*
PHIL SCHILIRO, *Minority Staff Director*

SUBCOMMITTEE ON THE CENSUS

DAN MILLER, Florida, *Chairman*

THOMAS M. DAVIS, Virginia
PAUL RYAN, Wisconsin
MARK E. SOUDER, Indiana

CAROLYN B. MALONEY, New York
DANNY K. DAVIS, Illinois
HAROLD E. FORD, JR., Tennessee

EX OFFICIO

DAN BURTON, Indiana

HENRY A. WAXMAN, California

JANE COBB, *Staff Director*
LARA CHAMBERLAIN, *Professional Staff Member*
MICHAEL MIGUEL, *Professional Staff Member*
ANDREW KAVALIUNAS, *Clerk*
MICHELLE ASH, *Minority Counsel*

CONTENTS

OVERSIGHT OF THE 2000 CENSUS: ACCURACY AND COVERAGE EVALUATION [ACE]—STILL MORE QUESTIONS THAN ANSWERS

FRIDAY, MAY 19, 2000

HOUSE OF REPRESENTATIVES,
SUBCOMMITTEE ON THE CENSUS,
COMMITTEE ON GOVERNMENT REFORM,
Washington, DC.

The subcommittee met, pursuant to notice, at 9:30 a.m., in room 2247 Rayburn House Office Building, Hon. Dan Miller (chairman of the subcommittee) presiding.

Present: Representatives Miller, Ryan, Maloney, and Davis of Illinois.

Staff present: Jane Cobb, staff director; Chip Walker, deputy staff director; Lara Chamberlain, Michael Miguel, and Amy Althoff, professional staff members; Andrew Kavaliunas, clerk; Michelle Ash, minority counsel; David McMillen and Mark Stephenson, minority professional staff members; and Earley Green, minority assistant clerk.

Mr. MILLER. Good morning. We will begin with opening statements first, and then we'll hear from Director Prewitt and then we'll proceed.

Director Prewitt, thank you for being with us here today. I'm pleased to hear that the Census Bureau is proceeding on schedule for nonresponse followup. This is the most difficult part of the census with respect to ACE, or the estimation plan.

As we move into the politically charged arena of the Bureau's estimation plan, one of the greatest concerns to Congress and the scientific community is that we will not be provided all of the information and data necessary to evaluate the results in a timely manner.

I want to get the Director's assurances today that these numbers will be fully scrutinized by the Bureau and the scientific community at large prior to their release for public use.

As many of you know, in 1990 there were numerous errors found in the sampling plan known as the PES. After the census in 1991 the Bureau discovered a computer error in the PES that inflated the undercount by 1 million people. Then, during a series of evaluations that took almost 2 years, the Bureau discovered more errors in the PES leading to even more erroneous enumerations.

In 1992, the experts at the Bureau who reviewed the 1990 census estimation plan issued what's known as the CAPE report. These experts reported that, "About 45 percent of the revised estimated

undercount is actually measured by us and not measured undercount." In other words, in 1990 almost half of the statistical adjustment was wrong.

Once States draw their district lines you can't come back a year later and say, "sorry, we made a mistake, we added or subtracted too many people." If it took almost 2 years in 1990 to find the errors, how can you ensure that we don't have the same errors this time in only a few months?

From a practical perspective, there is no guarantee that this plan is even viable. Despite claims to the contrary, the National Academy panel has not given the ACE a full blessing. While certain groups have endorsed statistical estimation as a concept, this is a far cry from an endorsement of the actual plan.

To give you an analogy, we all know that it's possible to build a spaceship to go to the moon or Mars. Yet, as with all very complex scientific tasks—and the estimation plan is immensely complex—your spaceship could blow up on the launching pad or burn up in the atmosphere as has happened twice recently.

What assurances do we have from the Director that their scientific plan won't blow up? Just because something may be theoretically possible doesn't mean it can be done.

Despite claims by the Democrats, Republican opposition to the estimation plan is based on fundamental, unresolved problems. Is the plan constitutional, is it legal, is it in the best interest of our Nation as a whole, and simply, will it work?

In January 1999, the Supreme Court ruled that sampling or estimating portions of the population was illegal. Democrats read the decision to outlaw the issue of sampling for purposes of apportionment only, while Republicans read the decision to prohibit the use of estimation for redistricting as well.

In the wake of that decision, the nonpartisan Congressional Research Service issued an opinion, "A closer examination of other parts of the Court's opinion indicates that it did not interpret those other purposes as necessarily including at least intrastate redistricting."

Unfortunately, this administration was not going to be deterred by even the Supreme Court. In a political move clearly against the best interests of the Nation, the Clinton-Gore administration decided to conduct a two-number census. This unwise decision will clearly throw the States into legal turmoil over the census, the likes of which this Nation has never seen.

And while the Democrats and their so-called experts have claimed that it is perfectly legal to use estimation for purposes of redistricting, I would simply offer a few words of caution. These are the same so-called experts that said estimation could be used for apportionment. Those on the estimation side of this disagreement have yet to win a court case.

The fundamental purpose of the census is to fairly and accurately apportion and distribute political representation. Our political system, for the most part, is the envy of many other nations. One of the foundations of our system is its relative transparency. Our elections are carefully scrutinized and the appeals process clear. If warranted, for example, an election can be challenged, voter registration records can be checked and rechecked, ballots re-

counted. With estimation, it's simply not possible to verify whether or not a person added actually exists or if a person subtracted was done so rightfully.

Additions and subtractions exist only in a virtual world, a world based not in reality, but in the complex mathematical formulas that could be right or wrong and understood only by a few select statisticians and government bureaucrats. Census estimation, no matter who is crunching the numbers, is not a system that lends itself to trust and integrity, two cornerstones of our electoral process.

And while we have spent billions of dollars to motivate people to participate in the census, something that all sides agree is a civic ceremony, what would motivate someone to participate in the census when they can sit back and be estimated? Why fill out your census forms at all if the government will compensate for you anyway? And even worse, how can it be acceptable that someone does their civic obligation, fills out their form on time and sends it in, only to have the government say they count as less than a whole person? Is this not a violation of one man, one vote? Can the Director guarantee that every person who filled out a form and only one form will be counted as one person and not less? The answer, disturbingly, is no.

The fact not widely talked about in the Bureau is, there will be people who do everything right, fill out their census forms, send them in on time and will be counted as less than a whole person.

While I fully support expending the resources to reach the undercounted, I wholeheartedly oppose the concept of counting someone as less than a whole person.

The census has traditionally been constructed of millions and millions of individuals. However, this estimation plan has introduced a new level of demographic grouping that is very dangerous in its assumptions. The assumption that people within racial groups act alike and have the same tendencies is something that this Nation has been trying to overcome for over 100 years, but now the Census Bureau has gone down that exact path.

No longer are we individuals. Now we are White males, 25 to 35, who rent or own. We are Cubans in Miami, Mexicans in Texas, Puerto Ricans in New York, that according to the Census Bureau are all alike and thus are grouped together. We are Asians, including Chinese, Japanese and Koreans from Seattle to Washington, DC, grouped together like so many choices.

The by-products of this estimation plan are not healthy for our Nation. From civic disengagement to simply throwing one man-one vote out the window, we in the long run hurt our Nation.

We must do everything possible to eliminate the undercount. We must also remain faithful to our Constitution, the law and the civic health of our Nation. Clearly this administration is putting politics ahead of sound public policy. Unfortunately, it will take the courts once again to protect the integrity of our census.

[The prepared statement of Hon. Dan Miller follows:]

SUBCOMMITTEE ON THE CENSUS
The Honorable Dan Miller, Chairman

H1-114 O'Neill House Office Building, Washington, D.C. 20515

FOR IMMEDIATE RELEASE
May 19, 2000

Contact: Chip Walker
(202) 226-1973

STATEMENT OF CHAIRMAN DAN MILLER

Oversight of the 2000 Census: A.C.E. – Still More Questions Than Answers

Director Prewitt, thank you for being with us again today. I am pleased to hear that the Census Bureau is proceeding ahead of schedule for non-response follow-up. This is the most difficult part of the Census with the exception of the A-C-E or estimation plan.

As we move into the politically charged arena of the Bureau's estimation plan, of greatest concern to Congress and the greater scientific community, is that we will not be provided all information and data necessary to evaluate the results in a timely manner.

I want to get the Director's assurances, today, that these numbers will be fully scrutinized by the Bureau and the scientific community, at large, prior to their release for public use.

As many of you know, in 1990 there were numerous errors found in the sampling plan known as the PES. After the census, in 1991, the Bureau discovered a computer error in the PES that inflated the undercount by one million people. Then, during a series of evaluations that took almost two years, the Bureau discovered more errors in the PES leading to even more erroneous enumerations.

In the 1992 the experts at the Bureau who reviewed the 1990 Census estimation plan issued what's known as the "CAPE Report." These experts reported that, "about 45% of the revised estimated undercount is actually measured bias and not measured undercount." In other words, in 1990, almost half the statistical adjustment was wrong.

Once states draw their district lines, you can't come back a year later and say, "Sorry. We made a mistake; we added or subtracted too many people."

If it took almost two years in 1990 to find the errors, how can you ensure we don't have the same errors this time in only a few months?

From a practical perspective there is no guarantee that this plan is even viable. Despite claims to the contrary, the National Academy Panel has not given the ACE its full blessing. While certain groups have endorsed statistical estimation as a concept, this is a far cry from an endorsement of the actual plan.

To give you an analogy, we all know that it's possible to build a spaceship to go to the moon or Mars. Yet as with all very complex scientific tasks, and the estimation plan is immensely complex, your spaceship could blow up on the launching pad or burn up in the atmosphere as has happened twice recently.

What assurances do we have from the Director that their scientific plan won't blow-up? Just because something may be theoretically possible, doesn't mean it can be done.

Despite claims by the Democrats, Republican opposition to the estimation plan is based on fundmental, unresolved problems:
Is the plan constitutional? Is it legal? Is it in the best interest of our nation as a whole? And, simply, will it work?

In January of 1999, the Supreme Court ruled that sampling, or estimating portions of the population, was illegal. Democrats read the decision to outlaw the use of sampling for purposes of apportionment only, while Republicans read the decision to prohibit the use of estimation for redistricting as well.

In the wake of that decision the nonpartisan Congressional Research Service issued an opinion saying, Quote " . . . a closer examination of other parts of the Court's opinion indicates that it did not interpret those other purposes as necessarily including, at least, intrastate redistricting."

Unfortunately, this Administration was not going to be deterred by even the Supreme Court. In a political move, clearly against the best interest of the Nation, the Clinton/Gore Administration decided to conduct a two-number census.

This unwise decision will clearly throw the states into legal turmoil over the census, the likes of which this nation has never seen.

And, while the Democrats and their so-called experts have claimed that it is perfectly legal to use estimation for purposes of redistricting, I would simply offer a few words of caution. These are the same so-called experts that said estimation could be used for apportionment. Those on the estimation side of this disagreement have yet to win a court case.

The fundamental purpose of the census is to fairly and accurately apportion and distribute political representation.

Our political system, for the most part, is the envy of many other nations. One of the foundations of our system is its relative transparency. Our elections are carefully scrutinized and the appeals process clear. If warranted, for example, an election can be challenged. Voter registration records can be checked and rechecked, ballots recounted.

With estimation it's simply not possible to verify whether or not a person added actually exists, or if a person subtracted was done so rightfully.

Additions and subtractions exist only in a virtual world. A world based not in reality, but in the complex mathematical formulas that could be right or wrong and understood only by a select few statisticians and government bureaucrats. Census estimation, no matter who is crunching the numbers, is not a system that lends itself to trust and integrity, two cornerstones of our electoral process.

And, while we have spent billions of dollars to motivate people to participate in the census, something that all sides agree is a civic ceremony, what would motivate someone to participate in the census when they can sit back and be estimated?

Why fill out your census forms at all if the government will somehow compensate for you anyway?

And, even worse, how can it be acceptable that someone does realize their civic obligation, does fill out their form on time and send it in, only to have the government say they count as less than a whole person. Is this not a violation of one man, one vote? Can the Director today guarantee that every person who filled out a form and only one form, will be counted as one person and not less? The answer, disturbingly, is no.

The fact not widely talked about by the Bureau is that there will be people who do everything right, fill out their census forms, send them in on time, and will be counted as less than a whole person.

While I fully support expending the resources to reach the undercounted, I wholeheartedly oppose the concept of counting someone as less than a whole person.

The Census has traditionally been constructed of millions and millions of individuals. However, this estimation plan has introduced a new level of demographic grouping that is very dangerous in its assumptions. The assumption that people within racial groups act alike and have the same tendencies is something that this nation has been trying to overcome for over a hundred years. But now the Census Bureau has gone down that exact path.

No longer are we individuals, now we are white males, 25 to 35 who rent or own. We are Cubans in Miami, Mexicans in Texas or Puerto Ricans in NY that according to the Census Bureau are all alike and thus are grouped together. We are Asians, including Chinese, Japanese and Koreans from Seattle to Washington, D.C. grouped together like so many clones.

The byproducts of this estimation plan are not healthy for our nation. From civic disengagement, to simply throwing one man one vote out the window we, in the long run, hurt our nation.

While we must do everything possible to eliminate the undercount, we must also remain faithful to the Constitution, the law and the civic health of our nation. Clearly this Administration is putting politics ahead of sound public policy. Unfortunately it will take the courts, once again, to protect the integrity of our census.

Mrs. Maloney.

Mrs. MALONEY. Thank you, Mr. Chairman. I must say I find the title of today's hearing very curious. It's called Status of the 2000 Census, Accuracy and Coverage Evaluation, Still More Questions Than Answers. Yet it seems to me, Mr. Chairman, that there are almost no unanswered questions, only questions which you don't like the answers to.

Despite the cautious stance taken by the Census Bureau, I believe that the 2000 census may well be the best, fairest and most accurate census ever, a fitting way to start the 21st century. It will be that not just because of the operational successes we have seen to date, but because it incorporates modern scientific methods into its design.

We all know the problems of the 1990 census. It contained millions of errors and was the first to be less accurate than the census before it. The 1990 census had an error rate of over 10 percent, 8.4 million people were missed, 4.4 million people were counted twice and 13 million people were counted in the wrong place. And we know who the people were that were missed. They were children. They were minorities in urban and rural areas.

The Census Bureau, working with the National Academy of Sciences at the direction of Congress in a bipartisan way, has tried to fix these errors, but there are politicians who, for partisan reasons, have tried to make sure that it doesn't happen. They have tried to make sure that minorities in poor, urban and rural areas and children are not undercounted.

The Census Bureau first discovered the problem of the undercount during World War II, 60 years ago, when more young men showed up for the draft than the Census Bureau thought existed. For young Black men, nearly 13 percent more showed up for the draft than they expected, showing that there was a disproportionate undercount, particularly for minorities.

We now know that the people missed in the census are the urban and rural poor and minorities. We also know that the people counted twice are primarily people who are fortunate enough to have two homes. They're suburbanites. Those errors shift economic resources and political representation unfairly. Those who oppose the use of modern scientific methods in the census would ensure that millions of people missed in the census are left out permanently and the millions of people counted twice are forever kept in. That is fundamentally unfair and it is unjust and it must stop, and that is why this census has been called the civil rights issue of the decade—whether we will correct, knowing people that are left out, whether we will do what every scientist says needs to be done to make sure that they are counted and represented.

The closer the Census Bureau has gotten to developing a way to fix these errors, the harder the opponents of a modern census have worked to stop them. In 1987, the professionals at the Census Bureau proposed a 300,000-household survey to measure and correct for the errors of the census. The politicians at the Reagan Commerce Department stopped that planning dead in its tracks. Correcting the 1990 census would have been stopped for good had not the great city of New York—and I am proud to be a Representative from that great city—the city of New York sued.

Finally, in late 1989, the Commerce Department allowed planning for the quality control survey to go forward. However, instead of allowing it to be a 300,000-household survey, the politicians at then-President Bush's Commerce Department cut it in half, and the Secretary reserved the right to block the use of the corrected, modern, scientific numbers. Not surprisingly, he overruled his own Republican-appointed Census Director, Dr. Barbara Bryant, and the professional nonpartisans at the Census Bureau, and he did block their use.

In 1997, the opponents of a fair and accurate census, they held up a disaster relief bill to the Midwest because they attached language to this important bill that would have prohibited the use of modern scientific methods. They believed that the President of the United States, when the country was in disaster, people were suffering, their homes were under water, that he would not have the nerve to veto the disaster relief bill over the census, over an accurate census, yet the President vetoed the disaster relief bill over the census because of their crass attempt to manipulate the numbers in the census; and he received editorial support clear across this country.

And I would like permission, Mr. Chairman, to put all of those editorials in the record of this hearing.

Mr. MILLER. OK.

Mrs. MALONEY. In 1997 and again in 1998, opponents of a fair and accurate census, those who did not want minorities and the poor and children in urban and rural areas counted, tried to use the appropriations process, the budget process, to legislate how census 2000 would be conducted by threatening to hold up two budgets and close down the government. Their attempt to block the use of modern scientific methods failed again.

Principally at the direction of Congress, the National Academy of Sciences has conducted extensive research and review of the planning and implementation of the 2000 census throughout the decade of the 1990's, working with the Census Bureau. Four separate panels of independent experts have consistently supported the use of modern scientific statistical methods, in general. More recently, a fifth panel, the Academy's panel to review the 2000 census has endorsed the Bureau's specific plans for the ACE program in census 2000; and I would like permission to put in the record the five statistical reports and scientific reports that have come out in support of the Bureau's plans.

May they be put in the record?

Mr. MILLER. No objection.

Mrs. MALONEY. The overwhelming majority of the scientific community has concluded that if we are to have a 2000 census that is fair and accounts for all residents of this country, regardless of race or economic status, it must be a census that uses modern methods. And I would now like to put into the record this list of organizations that supports the use of modern scientific statistical methods, and it includes all kinds of associations from across this country.

Mr. MILLER. Without objection.

Mrs. MALONEY. The General Accounting Office, the Commerce Department's Inspector General and George Bush's Census Director, Dr. Barbara Bryant, are all on record in their support.

The Census Bureau has presented its plans for the use of modern methods to the scientific community on a continuing basis since 1996. This subcommittee and the Census Monitoring Board have been kept apprised of those plans since their inception, and the Secretary's 2000 Census Advisory Committee, Race and Ethnic Advisory Committees and Census Advisory Committees of Professional Associations have all been briefed on these plans as well.

Again, Mr. Chairman, there are not any unanswered questions about the ACE program, only answers that the Republican Conference and the RNC doesn't like. I have heard the opponents of modern methods say repeatedly that they are a, "risky experimental plan that is inaccurate, that all we need to do is use old methods and try really hard to just count everyone." Well, the only thing that is risky is not using modern methods.

Over the course of the years, working on this issue, I have repeatedly heard from people, how do you know that you missed people, how do you know that there is an undercount? Well, the Census Bureau is unique among government agencies in that they tell you how well they have done. And the only way we know that the 1990 census was less accurate than the one before it was the 1990 post-enumeration survey, the use of modern statistical methods.

And the only way we will be able to determine the most accurate count for the 2000 census is from the results of the ACE program, the use of modern statistical methods. In the end, it is only through those methods that we will have the most accurate census possible.

I would also like to comment briefly on the Supreme Court case. Very simply, in the Supreme Court case, the Republicans won one and the Democrats won two. The Court held that you could not use modern scientific methods for the apportionment of seats between the States, but if at all feasible, you could use it for the distribution of Federal funds which is tremendously important since the Federal Government distributes over $185 billion a year based on funding formulas that are tied to census numbers. That means that over $3 trillion in the next decade will be distributed on these numbers, and we need to make them accurate; and they also held that it could be used for redistricting within the States, and that is why we have to come forward with two numbers, one for reapportionment between the States and one for redistricting and the distribution of funds.

I'd also like, Mr. Chairman, to briefly comment on the e-mail that you made public last week. Mr. Chairman, I continue to believe that it was inappropriate to make this public without first talking to its author, especially since you used it to imply a vast conspiracy by the census to hide information from Congress. And I remain disturbed by the fact that I had little more than an hour's notice of this e-mail's existence prior to our hearing. And I still am disturbed that GAO, a supposedly nonpartisan independent body, contacted the staff of the majority but did not contact the staff of the Democratic minority.

The e-mail may be poorly worded, but after speaking to its author, I called the gentleman. He is an honest, hardworking American. He is a former Marine. He is working very hard now for his country on this great civic ceremony, the census. I believe him, that

he made an honest mistake that is not in any way evidence of a systematic attempt to deny information to Congress.

Nevertheless, it seems to have raised questions in your mind which should be put to rest. Therefore, I have written to the Comptroller of GAO asking him to investigate this incident as soon as possible and to have the GAO determine if there is an intentional, systematic attempt to hide information from Congress. I'd like to introduce that letter into the record.

[The information referred to follows:]

CAROLYN B. MALONEY
14TH DISTRICT, NEW YORK

2430 RAYBURN BUILDING
WASHINGTON, DC 20515–3214
(202) 225–7944

COMMITTEES:
BANKING AND FINANCIAL SERVICES

GOVERNMENT REFORM

JOINT ECONOMIC COMMITTEE

Congress of the United States
House of Representatives
Washington, DC 20515–3214

DISTRICT OFFICES:
1651 THIRD AVENUE
SUITE 311
NEW YORK, NY 10128
(212) 860–0606

28–11 ASTORIA BOULEVARD
ASTORIA, NY 11102
(718) 932–1804

May 18, 2000

The Honorable David M. Walker
Comptroller General of the United States
U.S. General Accounting Office
441 G Street, NW
Washington, D.C. 20548

Dear General Walker:

At an oversight hearing held by the Subcommittee on the Census on May 11, 2000, Subcommittee Chairman Dan Miller opened the hearing by introducing into the record a copy of an email communication from a Census 2000 area manager in the Los Angeles region to local census office managers. I was surprised by this action, having been informed of the email's existence only an hour before the hearing, even though it was apparently shared with your staff the previous afternoon. Chairman Miller alleged that the email documented "a clear attempt to prevent Congress, through the GAO, from having access to the information." He continued by asking whether "this is the first time, or just the first time the Census Bureau has gotten caught?"

I believe that while this email may have been poorly worded, it provides scant evidence of a systematic attempt by the Census Bureau to deny GAO access to information. As Mr. Mihm testified at the hearing, the GAO already has direct access to all of the information contained in the form in question though the cost and progress data base. It is also important to note that this email was from a mid-level manager with no responsibility for articulating Census Bureau policy, particularly on an issue as sensitive as the access of various oversight bodies to Census 2000 information. Nevertheless, this email has clearly raised questions which it is important to answer.

As the Ranking Member of the Subcommittee on the Census, I request that you determine if there is currently an intentional and systemic attempt by the Census Bureau to prevent Congress, the GAO, or any other oversight body from having access to information.

Please contact Mark Stephenson at 225-5420 at your earliest convenience to discuss this request.

Sincerely,

Carolyn B. Maloney

Carolyn B. Maloney
Ranking Member
Subcommittee on the Census

cc: Rep. Dan Miller, Chairman, Subcommittee on the Census
Dr. Kenneth Prewitt, Director, U.S. Census Bureau

Mrs. MALONEY. I'd like to thank the chairman, and I would also like to put into the record, since at times this issue has been called partisan, and regrettably sometimes it has been partisan in our comments—so, therefore, to bring the debate above a partisan level, I would like to introduce into the record all of the editorials from across this country in support of the Census Bureau's plans, in support of the use of modern scientific methods to correct the undercount, and it comes from the Miami Herald, the Houston Chronicle, newspapers all across this country.

Mr. MILLER. I think we've already accepted that.

Mrs. MALONEY. Thank you.

[The prepared statement of Hon. Carolyn B. Maloney follows:]

14th District • New York

Congresswoman

Carolyn Maloney

Reports

2430 Rayburn Building • Washington, DC 20515 • 202-225-7944
1651 Third Avenue • Suite 311 • New York, NY 10128 • 212-860-0606

Statement of the Rep. Carolyn B. Maloney
Hearing on the Status of the Accuracy and Coverage Evaluation

May 19, 2000

Thank you Mr. Chairman.

I must say I find the title of today's hearing very curious. It's called "Status of the 2000 Census: Accuracy and Coverage Evaluation – Still More Questions than Answers." Yet it seems to me Mr. Chairman that there are almost no unanswered questions, only questions which you don't like the answers to.

Despite the cautious stance taken by the Census Bureau, I believe that the 2000 census may well be the best, fairest and most accurate census ever -- a fitting way to start the 21st century. It will be that not just because of the operational successes we have seen to date, but because it incorporates modern scientific methods into its design.

We all know the problems of the 1990 Census -- it contained millions of errors, and was the first to be less accurate than its predecessor. The 1990 census had an error rate of over 10 percent -- 8.4 million people were missed; 4.4 million people were counted twice; and 13 million people were counted in the wrong place. The Census Bureau, working with the National Academy of Science at the direction of Congress has tried to fix these errors, but there are politicians who for partisan reasons have tried to make sure that didn't happen.

The Census Bureau first discovered the problem of the undercount during World War II -- 60 years age -- when more young men showed up for the draft than the Census Bureau thought existed. For young Black men, nearly 13 percent more.

We now know that the people missed in the census are the urban and rural poor and minorities. We also know that the people counted twice are primarily suburbanites. Those errors shift economic resources and political representation unfairly. Those who oppose the use of modern scientific methods in the census would insure the millions of people missed in the census

are left out permanently, and the millions of people counted twice are forever kept in. That is fundamentally unfair and must stop.

The closer the Census Bureau has gotten to developing a way to fix these errors, the harder the opponents of a modern census have worked to stop them. In 1987, the professionals at the Census Bureau proposed a 300,000 household survey to measure and correct for the errors in the census. The politicians at the Regan Commerce Department stopped that planning dead in its tracks. Correcting the 1990 census would have been stopped for good had not the City of New York sued.

Finally, in late 1989, the Commerce Department allowed planning for the quality control survey to go forward. However, instead of allowing it to be a 300,000 household survey, the politicians at President Bush's Commerce Department cut it in half, and the Secretary reserved the right to block the use of the corrected numbers. Not surprisingly, he overruled his own Census Director, Dr. Barbara Bryant and the professionals at the Census Bureau and did block their use.

In 1997 the opponents of a fair and accurate census held up flood relief to the Midwest as they tried to prohibit the use of statistical methods in the census. The President vetoed their crass attempt to manipulate the census. In 1997 and again in 1998 opponents of a fair and accurate Census tried to use the appropriations process to legislate how Census 2000 would be conducted by threatening to hold up two budgets and close down the government. Their attempt to block the use of statistical methods again failed.

Principally at the direction of Congress, the National Academy of Sciences (NAS) has conducted extensive research and review of the planning and implementation of the 2000 Census throughout the decade of the 1990s working with the Census Bureau Four separate panels of independent experts have consistently supported the use of modern statistical methods in general. More recently, a fifth panel, the Academy's Panel to Review the 2000 Census has endorsed the Bureau's specific plans for the A.C.E. program in Census 2000.

The overwhelming majority of the scientific community has concluded, that if we are to have a 2000 census that is fair, that accounts for all residents of this country, regardless of race or economic status , it must be a census that uses modern methods. The General Accounting Office, the Commerce Department Inspector General, and George Bush's Census Director Dr. Barbara Bryant are all on record that to that effect.

The Census Bureau has presented its plans for the use modern methods to the scientific community on a continuing basis since 1996. This Subcommittee and the Census Monitoring Board have been kept appraised of those plans since their inception. And the Secretary's 2000 Census Advisory Committee, Race and Ethnic Advisory Committees, and Census Advisory Committee of Professional Associations have all been briefed on these plans as well.

Again Mr. Chairman, there aren't any unanswered questions about the ACE program only answers that the Republican conference and the RNC doesn't like. I have heard the opponents of modern methods say repeatedly that they are a "risky, experiential plan that is inaccurate, that all we need to do is use old methods and try really hard to just count everyone." Well the only thing that is risky is not using modern methods.

Over the course of the years working on this issue I have repeatedly heard from people, "How do you know that you missed people, how do yo know that there is an undercount?" Well the Census Bureau is unique among government agencies in that they tell you how well they have done. And the only way we know that the 1990 Census was less accurate than the one before was the 1990 Post-Enumeration Survey, the use of modern statistical methods.

And the only way we will be able to determine the most accurate count for the 2000 Census is from the results of the ACE program, the use of modern statistical methods. In the end it is only through those methods that we will have the most accurate census possible.

I'd like to also comment briefly on the email you made public last week, Mr. Chairman. I continue to believe that it was inappropriate to make this public without first talking to its author, especially since you used it to imply a vast conspiracy by the Census to hide information from Congress. And I remain disturbed by the fact that I had little more than an hour's notice of this email's existence prior to our hearing.

The email may be poorly worded, but after speaking to the author, I believe that to be an honest mistake, and not evidence of a systematic attempt to deny information to Congress. Nevertheless, it seems to have raised questions in your mind, which should be put to rest.

Therefore, I have written to the Comptroller of GAO asking him to investigate this incident as soon as possible, and to have the GAO determine if there is a intentional systemic attempt to hide information from Congress. I'd like to introduce this letter into the record.

Thank you Mr. Chairman.

Mr. MILLER. Mr. Ryan.

Mr. Ryan, before you begin, you mentioned the editorials. There's some very large number of editorials opposed to the concept of statistical adjustment and manipulation, and I ask consent that we enter those in the record, and without objection, those editorials will be included.

[The information referred to follows:]

The Washington Post, May 14, 2000

The Washington Post

May 14, 2000, Sunday, Final Edition
SECTION: OUTLOOK; Pg. B02
LENGTH: 1878 words

We're Overstating The Importance Of the Undercount
By Peter Skerry

The undercount hunt is underway. But as **Census** 2000 enumerators knock on doors and try to coax information out of the 34 percent of the population that did not return the forms mailed out nine weeks ago, it's worth setting the record straight on a couple of issues. First, the benefits for minorities in maximizing their **census** totals have been wildly exaggerated; and second, important causes of the undercount are quite beyond the bureau's power to remedy.

There is no disputing the reality of what the bureau calls "the differential racial undercount." According to its own estimates, the bureau missed 1.6 percent of the total population in 1990. A disproportionate number of those missed were minorities: 2.3 percent of Asian/Pacific Islanders went uncounted, 4.4 percent of blacks, 4.5 percent of American Indians and 5.0 percent of Hispanics, compared with 0.7 percent of non-Hispanic whites. And, although preliminary evidence indicates an unusually good minority response to the 2000 **Census,** over the previous five censuses the differential has been widening.

Understandably, minority advocates seek to increase group numbers in the hope that doing so will benefit their people fiscally and politically. But such hopes are ill-founded.

In terms of the fiscal stakes, it is usually reported that $ 185 billion in federal funds are distributed annually on the basis of formulas that rely on **census** data. True enough. Yet research by the National Academy of Sciences and the General Accounting Office consistently concludes that less than 0.5 percent of those funds are affected by the overall undercount.

This is in part because federal grant formulas are not simply based on population, but on numerous other factors, including income and age. Moreover, in programs designed to help distressed communities, a gain in population may be seen as a sign of health and can actually lead to a reduction in funding. Even when grant increases are pegged to population gains, the critical factor for a given jurisdiction is not merely its absolute population, but its population relative to other jurisdictions--a result that can obviously hurt as well as help minorities. Further, grant programs typically have funding ceilings,

so larger numbers simply mean a fixed pie divided into smaller pieces.

What about the political benefits? How does the minority undercount affect the drawing of congressional districts--one of the primary purposes of the **census?** And would the proposed remedy of statistical sampling and adjustment actually empower minorities, as many believe?

It is widely assumed that the undercount disadvantages Democrats by disadvantaging minorities. But as Tom Hofeller, redistricting director for the Republican National Committee, once said, "The gerrymander overcometh all. What demographics give, legislatures can take away in the dead of the night." The critical factor in each of the 50 states is not simply numbers but which party controls the redistricting process. In 1990 Democrats controlled both legislative chambers and the executive mansion in 19 states, Republicans only three. Today the situation is substantially reversed, with Republicans dominating 14 states, Democrats just 11.

When Republicans are in control, the minority undercount makes scant difference, because they have long since mastered the art of packing minorities into "majority minority" districts. While such districts make it easier for minority Democrats to get elected, they also mean the creation of more homogeneously white districts, which tend to hurt the electoral chances of non-minority Democrats and help those of Republicans. On the other hand, when majority minority districts are not an option, Republicans have been equally adept (within the constraints of the 1965 Voting Rights Act) at diluting minority numbers by dispersing them among several districts. In either case, the critical factor is not marginal differences in minority population totals, but who controls the redistricting process.

Among the states currently under Democratic control, the most important is, of course California, where Republicans will be slow to forget the redistricting battle they lost in 1980. According to conventional wisdom, minority concern over the undercount is shared by California's Democratic leaders. But the interests of minorities have rarely been congruent with those of party leaders whose main concerns, after all, are protecting incumbents and broadening the party's electoral appeal. But even a skillful Democratic gerrymander dedicated to minority empowerment eventually collides against the obdurate fact that undercounted minorities tend not to vote.

What can the bureau do about the minority undercount? Not surprisingly, **census** critics explain the causes of the undercount in terms of what they think the bureau ought to have been doing. Is there a language barrier to filling out **census** forms? Print them in languages other than English. Do illegal immigrants fear being reported to the Immigration and Naturalization Service? Tell la migra to back off during **census** season. Do respondents worry about disclosing information that might jeopardize social welfare benefits? Ensure the confidentiality of **census** data.

These tactics are all good ones--except that they barely scratch the surface. In 1990, the bureau sponsored dozens of experimental "alternative enumerations" by social scientists

who did field research to double-check results in various hard-to-count neighborhoods. The results show the causes of the undercount to be more complex and daunting than many had supposed.

For example, virtually all of these studies indicate that among minorities distrust of government is only one factor in their reluctance to fill in forms. Another powerful factor is distrust of one's neighbors. Koreans have a saying: "Don't let the government know how many sons and how many cows you have." But among Korean immigrants another cause of underreporting the number of people in apartments is their reluctance to reveal overcrowded living conditions. Similarly, some Korean women are reluctant to admit having married outside their ethnic group and therefore neglect to report non-Korean partners to the **census.**

Another reason for the low response to enumerators is that many minority "communities" turn out to be far less cohesive than that term implies. Crime is common enough that residents fear opening their doors to strangers, including **census** enumerators--even, or perhaps especially, if the enumerators share their language and background. Households are frequently complicated, with individuals coming and going constantly. In many immigrant enclaves, single men share cramped quarters with strangers, often sleeping in shifts. And even when men live with their families, the level of trust may not be high. One study of San Francisco's Mission District describes an apartment occupied by three Salvadoran families who, while sharing the kitchen and bath, each then retreated behind the locked door of their own bedrooms where they spent their time and stored their food. In suburban Long Island another group of Salvadorans sharing quarters avoided conversing with one another for fear of reigniting hostilities from their country's civil war.

The high level of distrust in some neighborhoods was driven home to two researchers doing an alternative enumeration in a San Francisco public housing project. Despite the fact that they were extremely well-plugged into the neighborhood, with a team reflecting its diverse cultures and languages, the researchers, who had access to the usually confidential **census** forms, were astounded to find that in some instances residents had been more forthcoming on the questionnaire than with the research team.

Nonetheless, the undercount is an important symbolic issue for many minorities. Many African Americans are especially sensitive about the undercount because they trace it back to the Constitution's original stipulation that each slave be counted as three-fifths of a person. Doubtless this is why many of those concerned about the undercount frame the issue in terms of rights. Sen. Charles Schumer (D-N.Y.), for example, has asserted, "The Constitution, of course, guarantees the right of every person residing in the United States to be counted." Indeed, some of those who are concerned about the undercount equate this "right to be counted" with the right to vote, and argue that not being counted in the **census** is equivalent to being disenfranchised.

Framed thus, the minority undercount presents a major challenge to political elites. To Democrats, it is one malady for which there exists a clear remedy: adjustment by means

of statistical sampling. Given the seemingly intractable nature of many minority and urban problems, and the potential cost of proposed solutions such as affirmative action, increased minimum wages, greater educational expenditures and the like, adjustment seems relatively inexpensive. It is also, like redistricting, a highly technical issue unlikely to arouse much opposition among the general public. But adjustment is a complex, error-prone procedure, the risks of which greatly outweigh its presumed benefits.

Republicans are loath to appear to be against minority enfranchisement, so their arguments against adjustment have focused on statistical and legal technicalities. Their failure to address the question of whether the undercount really hurts minorities has left Republicans in their assigned role as mean-spirited opponents of racial justice.

But symbolism can be taken too far. A moment's reflection makes it clear that there is no such thing as a right to be counted. Quite the opposite: Cooperating with the **census** is a legal obligation. Granted, individuals not counted in the **census** are not included in the one-person-one-vote calculations by which district lines are drawn. But there is nothing to prevent such citizens from voting or from going out and organizing others to vote.

Nor is being counted tantamount to being given the franchise--especially since the **census** enumerates many individuals, such as children and non-citizens, who are not permitted to vote. Simply being counted does not "empower" these people in any meaningful way.

Numbers are important. But more important is what people do with them. Think of the memorable formulation of Rep. Barney Frank (D-Mass.): "How many members of the NRA are there? I don't know. I don't think my colleagues know. What's important politically is not how many there are, but what you do about it. The extent to which you mobilize enormously outweighs the numbers." Population totals simply don't translate automatically, without group organization or effort, into political power. Yet rather than confront the genuine social and political problems of the disadvantaged, we have all but convinced ourselves that tweaking population totals at the margins will result in minority empowerment.

Peter Skerry teaches at Claremont-McKenna College in California and is a senior fellow at the Brookings Institution, which has just published his book "Counting on the **Census?** Race, Group Identity and the Evasion of Politics."

Mr. MILLER. I am glad you talk about getting away from the partisanship, but it's something you make a statement and it really bothers me to set claims, Mrs. Maloney, that we don't want to count people. This Congress has provided every penny and provided all the resources the Bureau has asked to get everybody counted. That is our objective, that is our goal; and we are not trying to not count people, and to say that is just political rhetoric. So I just want to make sure we clear the record.

Let me clear one other record and that is the question about scientific endorsement. The National Academy of Sciences panel to review the 2000 census has not endorsed the ACE. I had a long meeting with Janet Norwood only 2 days ago, and she emphatically stated several times that neither she nor any member of the panel has made any determination as to the quality or outcome of the ACE. She explained what is quite obvious to most people, that you can't evaluate a statistical process if, one, it's not complete and, two, you don't have the data. So please stop misrepresenting the truth here. The ACE is a statistical plan, but at the moment it is mostly just that, a plan.

Mr. Ryan.

Mr. RYAN. Thank you, Mr. Chairman.

Mr. Chairman, let me followup on this. We are going into that touchy part of the census where I think we have done very remarkable accomplishments in the enumeration process, and I'm excited about hearing more about how well the enumeration is working, but now we are getting into that touchy area and now we are hearing the kinds of discussions, the kind of political rhetoric that is unfortunate.

First of all, it's not all Democrats against all Republicans. Democrats in my home State of Wisconsin are against sampling because they know it is not good for the State of Wisconsin. So I object to the characterization that opposition to the ACE plan means opposition to counting people. The same people who oppose the ACE plan, the chairman and myself, are the same people who have provided $7 billion to improve the census, especially in the traditionally undercounted areas and communities in the inner cities, $7 billion provided for advertising, $7 billion provided for partnership, hiring that far exceeded any previous census. So saying that the people who provided these resources don't want the people to be counted is wrong and is actually racially divisive.

Post-census local review, we passed that out of Congress. We can't get it signed into law. Post-census local review, in my opinion, is a very good idea. It simply says in those hard-to-count areas, the inner city of Milwaukee, the inner city of New York, please look at the data, local, elected officials, tell us if we missed anybody. Did we miss a public housing complex? Did we miss a neighborhood that's tough to reach? Did we miss a Latino neighborhood that didn't want to comply with the census? If so, we'll go back and get those people counted. That's a very common-sense idea.

It's a common-sense idea that was supported in 1990 by Democrats and Republicans alike. The mayor of Chicago, mayors all across the country supported a common-sense idea like post-census local review. We passed post-census local review to try and improve the count, to try and make sure that those people who are histori-

cally undercounted get counted, that local officials, mayors, county board members, city council people, pastors in inner-city churches get a chance to look at the data before it's finalized to make sure that their citizens weren't missed. Well, this administration blocked post-census local review. We don't have post-census local review.

LUCA was a good idea. LUCA worked well, but it can be improved upon. I still think we should do post-census local review. So to suggest that those of us who have questions about the statistical adjustment are somehow against getting the most accurate census is a ridiculous claim, No. 1.

No. 2, the scientific community is clearly not unified on this point, so also to suggest that the scientific community is completely behind statistical sampling is not correct. I have a letter here from the Statistics Department of the University of Berkeley in California, Dr. David Freeman and Kenny Walter. I ask unanimous consent that this letter be inserted into the record, Mr. Chairman.

Mr. MILLER. Without objection.

[The information referred to follows:]

UNIVERSITY OF CALIFORNIA, BERKELEY

BERKELEY • DAVIS • IRVINE • LOS ANGELES • RIVERSIDE • SAN DIEGO • SAN FRANCISCO

SANTA BARBARA • SANTA CRUZ

TELEPHONE: (510) 642-2781
CABLE: UNIVCALB
FAX: (510) 642-7892
URL: http://stat-www.berkeley.edu

UNIVERSITY OF CALIFORNIA
DEPARTMENT OF STATISTICS
367 EVANS HALL # 3860
BERKELEY, CALIFORNIA 94720-3860

17 May 2000

Chairman Dan Miller
Subcommittee on the Census
Committee on Government Reform
U.S. House of Representatives
H1-114 O'Neill HOB
Washington, DC 20515

Dear Congressman Miller:

Census 2000 and ACE (a method for adjusting the census using "capture-recapture") are quite similar in design to Census 1990 and the PES, as discussed by Stark (2000). Data from 1990 and previous censuses suggest there will be a net undercount in 2000, differential by gender, age, race, ethnicity, and-- most importantly-- by geography. History also suggests that efforts to adjust the census are unlikely to improve its accuracy.

There are three main problems with capture-recapture.

1) Processing error. On 15 July 1991, Secretary of Commerce Mosbacher turned down a proposal from the Census Bureau to adjust Census 1990 by adding 5.3 million people (net, nationwide). Each and every state would have had its population increased, some more than others. Research by the Bureau and others, summarized in Wachter and Freedman (2000), has shown that 3.0 to 4.2 million out of this 5.3 million figure represents errors not in the census but in the adjustment process. Our best estimate for total processing error is 3.6 million (net, nationwide).

To mention only one issue, movers must be identified and matched to the census at their census-day address. Failure to identify movers' can artificially inflate the adjustment. In 2000, identifying movers will be even more problematic than it was in 1990, because information on movers will largely be gleaned from "proxy" interviews with neighbors, rather than direct interviews with the movers themselves.

2) Correlation bias. Some people are missed both by the census and the adjustment process. To determine the size of this bias, we need an estimate of the population that is independent of both the census and the adjustment. One resource is Demographic

Analysis (DA), keyed to administrative data like birth and death certificates. DA provides national population estimates by gender, age, and race. (In 1990, DA did not provide estimates for state or substate areas, but that may change in 2000.) DA suggested that correlation bias was in the range from 2.4 to 3.6 million, compared to the estimated net undercount of 5.3 million. Bias was especially marked for Black males. This and other evidence suggest that the size of the bias was quite different in different parts of the country.

Thus, the Census Bureau proposed to adjust Census 1990 by adding 5.3 million people. However, this adjustment was itself in need of correction: 3.0 to 4.2 million people needed to be subtracted from the adjustment in various states (point #1), and some similar number of very different people needed to be added back in other states (point #2). For details, see Wachter and Freedman (2000). Since most of the 1990 adjustment was bad data, Secretary Mosbacher was right to reject adjustment.

3) Heterogeneity. ACE estimates undercount rates for broad demographic groups called "post strata." For example, one post stratum consists of non-Hispanic Asians age 0-17 living in rental units anywhere in the U.S. Another post stratum consists of Hispanic females age 18-29 in owner-occupied units in large cities across the country, where the census mail-back rate was high.

Undercount rates for a post stratum are estimated by pooling data from the whole ACE sample. Small areas-- blocks, counties, congressional districts, states-- are adjusted by assuming that undercount rates are constant within post strata across geography. For example, it is assumed that all non-Hispanic Asians age 0-17 living in rental units are equally likely to be undercounted-- from the suburbs of Honolulu to Chinatown, New York. This assumption is plainly false. Freedman and Wachter (1994) used 1990 census data on variables thought to behave like undercount rates, and demonstrated huge variations within post strata across states.

The Bureau has created separate post strata for Puerto Rico, indirectly acknowledging the heterogeneity issue. But why Puerto Rico and not Alaska, Hawaii, or Los Angeles?

This completes our summary of the main problems with adjustment. In 2000, will the census headcount be more accurate, or the adjustment? This is a critical policy question, and there is unlikely to be any definitive answer in the short term. Results for 1990 are the best guide to the future, because data for 1990 were freely available and intensively studied. Even for 1990, however, there is no general agreement: compare Anderson and Fienberg (1999) with Brown et al. (1999).

To address questions of comparative accuracy, the following sorts of data would be needed.

1) Adjustments to the census, as estimated by ACE and by Demographic Analysis, tabulated in a common set of categories (gender crossed with age and race).

2) Raw and adjusted census counts, by area and post stratum.

3) Detailed adjustment data for the various post strata, as in the 1990 Advisory Use File; and for the various sample blocks, as in the file provided to Congressional Monitoring Board.

4) Rematching and reinterview studies like the Evaluation Followup of 1990, as described in the P-studies. The 1990 error analyses (like P-16) had little geographic detail, except via rather questionable imputations; see, for instance, Freedman et al. (1994). In 2000, the Bureau should provide real geographic detail in its error studies.

5) Adjustment factors by post strata, estimated biases in these factors, and the corresponding covariance matrices.

6) To study heterogeneity, proxy variables can be tabulated by post stratum and state. Given the size of ACE, it should also be possible to compare direct and synthetic population estimates, at least for some number of large geographical areas.

Yours sincerely,

David A. Freedman
Professor of Statistics

Kenneth W. Wachter
Professor of Demography
and Statistics

References

Breiman, L. (1994). The 1991 census adjustment: undercount or bad data? Statistical Science 9: 458-537.

Anderson, M. and S.E. Fienberg (1999). Who Counts? The Politics of Census-Taking in Contemporary America. Russell Sage Foundation, New York.

Brown, L.D., M.L. Eaton, D.A. Freedman, S.P. Klein, R.A. Olshen, K.W. Wachter, M.T. Wells, and D. Ylvisaker (1999). Statistical controversies in Census 2000. Jurimetrics 39: 347-75.

Freedman, D.A. and K.W. Wachter (1994). Heterogeneity and census adjustment for the intercensal base. Statistical Science 9: 476-485.

Freedman, D., K. Wachter, R. Cutler, and S. Klein (1994). Adjusting the census of 1990: loss functions. Evaluation Review 18: 243-280.

Stark, P.B. (2000). The 1990 and 2000 Census adjustment plans. Technical report 550, Department of Statistics, U.C. Berkeley.

Wachter, K.W. and D.A. Freedman (2000). The fifth cell. Evaluation Review 24: 191-211.

Mr. RYAN. Also, the National Academy of Sciences has not endorsed the ACE plan. So to suggest that the scientific community believes that this is a unified point, that's just not the case.

Sampling didn't work in 1990. We found that years later we had dire problems. So one thing that I think in today's hearing, hopefully we can get into, is the compressed time line that this plan involves. I am concerned that this rushed time line is going to give us the errors we will discover down the road when it's too late.

So with that, Mr. Chairman, I'd like to yield back the balance of my time. My friend from New York, the national academy of scientists hasn't officially endorsed the ACE plan. The scientific community is divided on this. So I hope that we can move forward on an even keel, on an objective basis; and I hope that we won't get into this heightened political rhetoric where we are impugning the motives of each of the two parties involved.

All of us want an accurate count. All of us want everybody to be counted in the neighborhoods where they live. Democrats and Republicans in Wisconsin want that. All of us want that. So let's keep the discussion at that level if we may.

With that, I yield back the balance of my time.

Mrs. MALONEY. I request permission to respond. I think that I should be able to respond.

Mr. MILLER. Mr. Davis.

Mr. DAVIS OF ILLINOIS. I yield time, Mr. Chairman.

Mrs. MALONEY. Very briefly I feel that we should let the facts speak for themselves, so I would like to put into the record the legislation, bipartisan, that went—that passed calling on the National Academy of Sciences to come forward with a plan to correct the undercount. The plan that they came back with, which was the use of modern scientific methods, I would like to put into the record the language that the Republican majority attached to the disaster relief bill that would have prohibited the use of modern scientific methods. I would like to put into the record that the Republican majority tried to attach to two budget bills, holding up two budgets, that would have limited and prohibited the use of modern scientific methods.

The facts speak for themselves, and I will put that in the record, and it is clear—and it is clear what the intention of such actions would do and how it would affect and continue an undercount. Knowingly, the majority tried to stack the deck so that millions of Americans would be intentionally left out of representation and funding dollars in this country. It is unfair. It is unjust. It is the civil rights issue of this decade.

Mr. RYAN. Will the gentlelady yield for an honest point of clarification—not a tit for tat, just an honest point of clarification?

Mr. DAVIS OF ILLINOIS. Reclaiming my time, I yield.

Mr. RYAN. Thank you, Danny. Appreciate it. That was 2 years ago in the last Congress. That was then; this is now. Let's move forward with not a lot of political rhetoric. Let's move forward and debate this objectively, and let's not impugn the motives of each other. We all want an accurate count. With that, I yield.

Mr. DAVIS OF ILLINOIS. Reclaiming my time, Mr. Chairman, let me thank you for convening this hearing regarding oversight of the 2000 census and the impact of the accuracy and coverage evalua-

tion, that is, the ACE process. As enumerators begin the process of going door-to-door to those households that did not send in their census forms, it is important that we examine the ACE operations. The ACE process was added to the 2000 census to replace the post-enumeration survey of 1990 in an effort to improve the accuracy of the census.

We all know that accuracy must be the goal. We can ill afford to go back to the days of 1990 when too many people lost from an inaccurate census. The constituents of my district, the Seventh Congressional District in Illinois, deserve and need an accurate count of the entire population. They realize that too much is at stake to get a less-than-accurate count.

In 1990 Chicago lost millions of dollars in Federal funds because of a census undercount. According to the Bureau, at least 10 people, including at least 113,831 in the State of Illinois, 81,000 in Cook County and 68,000 in the city of Chicago were not counted in the 1990 census. Many of those missed were obviously children and women who live in minority communities. Because the 1990 census missed counting millions of people in Chicago, every one of our residents were shortchanged on money to repair roads and streets. They were shortchanged on money for mass transit and senior citizen homes. They were shortchanged on money for schools, parks and job training. Perhaps the most egregious shortchange was that of political representation, and in a democracy representation is essential to having a voice in local, State and Federal Government.

I represent many hard to count people. According to the Census Bureau 165,000 of them live at or below the poverty level in my district. I'm pleased that we're holding these hearings in an effort to make certain that the Census Bureau and others are doing everything possible to get an accurate count. Yes, many people are indeed difficult to count. Therefore, we must use every effort to try and make sure that the past evils and transgressions of our Nation do not continue to negatively impact upon the reality of our being, and if there is to be fairness, we must indeed make use of every method available to us.

Just 2 days ago the mayor of my city, Mayor Daley announced a $400,000 radio and television advertising campaign to be funded by the city to encourage people to cooperate and participate in the 2000 census. This advertising campaign coupled with what the Census Bureau has already committed to will go a long way toward a more accurate census. However, I tell you that you cannot undo with radio and television ads what 400 years of slavery, deprivation, discrimination, denial of equal opportunity, lack of opportunity to be educated, to understand, to be a part of the mainstream, you cannot undo that with radio, television and newspaper advertising. You cannot even undo it by sending people door to door, looking for people that you cannot find, people who in many instances are unreachable, untouchable and unfindable, and I don't care how much we put in, unless we make absolutely certain that every available technique, every scientific advancement, every opportunity exists to count every single person, account for every single person in this country, then we once again will come up lacking. Once again, individuals will be left out. Once again, individ-

uals will be shortchanged and once again, this Nation will have shortchanged itself.

So, Mr. Chairman, I am pleased that we're having this hearing and look forward to the information that Dr. Prewitt will share with us. So I thank you and yield back the balance of my time.

Mr. MILLER. Thank you, Mr. Davis. You do outline the real challenges of this massive undertaking that we're in the process of. Dr. Prewitt, if you and Mr. Hogan and Mr. Thompson would stand and raise your right hands, we'll get you sworn in and proceed, in case Mr. Thompson and Mr. Hogan is needed to assist you.

[Witnesses sworn.]

Mr. MILLER. Thank you and please be seated. The record will reflect that Mr. Thompson, Dr. Hogan, Dr. Prewitt answered in the affirmative. Welcome. Thank you. You may proceed with the opening statement.

STATEMENT OF KENNETH PREWITT, DIRECTOR, BUREAU OF THE CENSUS, ACCOMPANIED BY JOHN H. THOMPSON, ASSOCIATE DIRECTOR FOR DECENNIAL CENSUS; AND HOWARD HOGAN, STATISTICIAN, CHIEF, DECENNIAL STATISTICAL STUDIES DIVISION

Mr. PREWITT. Thank you, Mr. Chairman, and I did solicit the chairman's permission this morning to spend just a few minutes returning to the e-mail incident that was addressed last week at the subcommittee hearing with the GAO, and I have also informed the minority and also the leader and also Mr. Davis and Mr. Ryan that I would like to address that quickly.

I'd like to start by saying that it is understandable that in the absence of facts the offending sentence instructing the LCO managers not to share a given report with the GAO could have led to the strong reaction of the chairman, of Congressman Ryan and of Mr. Mihm of the GAO. But the facts do in fact mitigate this reaction and I would like to quickly put them in the record, and here I paraphrase from a subsequent e-mail by Mr. Rodriguez, who was the author of that initial e-mail.

The report in question, he explains is, a regional level report and information from it should be shared only by the regional level. It in turn generates local office reports, and it is this information that can be shared by local managers. As he writes in a subsequent e-mail to us, that, as per our instructions on May 8th at about 3 p.m., the report, the offending report—not the offending report but the one that initiated the incident—was to be shared at the area manager's discretion with their local census office but any one local census office was not to share the production information of a different LCO; that is, each LCO was only to share its own information and not other patterns of information. If anyone outside the Census Bureau wanted to obtain a regional level report, they can get that from the regional census center, and then he goes on to explain why he sent the report he did: "My intentions were to provide the report to my offices as a tool to encourage friendly competition and thereby productivity, nothing more. I regret that my intentions have been misinterpreted and rightly so because of the way my e-mail was worded, and I apologize for any inconvenience I may have caused."

I would like to say, Mr. Chairman, that obviously I join in that apology but nevertheless do want to make certain that we understand that the facts themselves give no warrant for the accusation that the Census Bureau is preventing the GAO from discharging its responsibilities.

Mr. Chairman, we have taken this accusation so seriously that yesterday with Deputy Director William Barron I met with David Walker, the Comptroller General, with Nancy Kingsbury and with Chris Mihm to reiterate the Census Bureau policy in regard to access, and I do believe that we all fully understand that there's nothing in our policies that are designed to prevent any access by the GAO.

Mr. Chairman, more than a week ago at the subcommittee hearing to which I've referred, you said that there were Census Bureau employees, "in very influential positions who are dangerous." This I take as a very serious charge. If substantiated, I would take corrective action. Obviously if I am, myself, the person who is dangerous, then I would expect you to bring that to the attention of the Secretary and he would take corrective action.

In that same hearing, you asked the GAO to investigate the incident that led you to make this extremely serious charge. Yesterday I asked David Walker, Comptroller General of the GAO, if his organization had any evidence that would corroborate your charge that the Census Bureau has people in very influential positions who are dangerous. He replied in the negative, and in this he was seconded by Christopher Mihm, who also was present at the meeting.

Mr. Chairman, it is now more than a week after your charge. I know that you and your staff had conducted your own independent investigation, and I wait for the evidence on which this accusation is based, for I am unable to take corrective action until I know who these people are and what it is that makes them dangerous. So may I respectfully request that you please provide me, as soon as possible, the names of the dangerous people, the nature of the danger they pose and of course the evidence that would substantiate this charge, and I promise to you that I will take corrective action. Thank you, sir.

Now, if I may, I'll turn to my opening comments on the topic of this hearing.

Census 2000 operations continue on track and on budget. I earlier reported that the mail-back response rate at 66 percent was very encouraging, and in my written testimony I indicated that we had completed 39 percent of the nonresponse followup workload. I would like to update that number through yesterday. The workload is now 50 percent complete. Putting these two numbers together, the census enumeration is now approximately 85 percent complete. That is, 85 percent of the housing units have now responded or we have identified them as vacants.

Still, the Census Bureau does not anticipate at this point that census 2000 will have better coverage than the 1990 census because many of the factors that led to the undercount in 1990 are still present in American society, and indeed, as a proportion of the population have grown—more gated communities, more recent immigrants, more linguistically isolated households, more persons liv-

ing in irregular housing and perhaps more anger toward the government.

The Census Bureau has both measured and documented the existence of a substantial undercount since 1940, and this has already been referenced in the opening comments. The Census Bureau has been running harder but believes this will only allow us to stay even. That is, we expect that neither the overall coverage levels nor the differential undercount rates in census 2000 will show improvement over 1990. The Census Bureau strongly hopes to be proven wrong in this assessment, and the ACE will give us the information to determine whether this is so.

The ACE provides a final quality check on how well we have done in the initial census. The alternative is not to do the ACE and never know how we have done below the national level where demographic analysis does provide a benchmark. The ACE also provides the means to generate a more accurate count.

The 1990 version of the ACE, or the accuracy and coverage evaluation, was called the post-enumeration survey. It provided information that was used during the 1990's to improve statistical programs. The population estimates the Bureau of Labor Statistics asked us to incorporate into the current population survey program following the 1990 census were corrected for the undercount identified through the 1990 PES. The Bureau of Labor Statistics also requested adjusted population controls for the consumer expenditure survey. All other major national demographic surveys conducted by the Census Bureau or other agencies in the Federal statistical system also were converted to this adjusted population base. And Katherine Abraham, the current Commissioner, of course testifies that in the absence of this correction their published demographic distribution of unemployment and other measures would have been inaccurate.

We believe we have an obligation to the Bureau of Labor Statistics and the many, many other users of our data products to make our data as accurate as possible. I have said previously that the Census Bureau currently expects that the corrected numbers using the accuracy and coverage evaluation will be the more accurate numbers. If the Census Bureau does not have confidence in the results, we will not use them. The decision whether to release the statistically corrected data should take into consideration operational data to validate the successful conduct of the ACE, whether the ACE measurements of undercount are consistent with historical patterns of undercount and a review of selected measures of quality.

In the fall of this year the Census Bureau will discuss the review process and criteria with the statistical community and other interested parties. We will set forth how we will assess whether our operational functions for the ACE were met. All major operations have been designed and documented and the details have been available for review and comment. Every document requested by the subcommittee has been forwarded. Here, however, Mr. Chairman, is a complete set. It is possible there are documents here that you have not yet requested, but we can provide you the entire complete set of our decision documents that go into the design of the ACE.

Now, let me very quickly try to describe the operations as requested in your invitation letter. Several major operations have now been completed. One is ongoing and others will follow the completion of nonresponse followup. All operations are currently on schedule.

The basic concept behind the ACE is the comparison of the data from two systems, an independent survey and the initial census. Because of its small size relative to the initial census, we believe we can do a better job enumerating people in the housing units in a sample. We can be more selective about the interviewers, train them longer, pay them more and provide more quality assurance.

The first step in the ACE process is to design and select a sample which consists of approximately 314,000 housing units or about one-fourth of 1 percent of the total housing units. The basic units of this sample are what we call block clusters, and there will be about 11,800 block clusters in our sample. The sample was designed and selected to provide sufficient precision to estimate the true population for various groupings of the population that we call post-strata which I will describe below.

The next step in the process is to create an independent listing of housing units. By independent we mean that we do not start with or refer to the master address file from census 2000 but instead have census staff systematically canvass the block clusters to list the addresses. This operation was completed in the fall of 1999, checked and keyed and 100 percent quality controlled in our national processing center.

The Census Bureau then matches this list of housing units to the master address file, first by computer and then clerically if necessary, using this additional information, we continue to improve our address list. The purpose of this housing unit match is to create an accurate linked list of housing units in the block clusters. This work was also completed on schedule.

To provide sufficient data to compare the ACE to the initial census the Census Bureau of course must conduct interviews to collect data from each of the housing units that were independently listed. We initiated the ACE interviewing with a telephone phase using laptop computers in a technique we call computer assisted personal interviewing. This is the laptop computer that we're using that in that procedure, and we would be delighted, of course, to provide a staff briefing of how it is used.

We began in late April telephoning households in the ACE sample at unique addresses for which a census 2000 questionnaire had been mailed back, processed through data capture and for which a telephone number was provided. As of today we have completed over 60,000 interviews by telephone, more than 20 percent of our workload for the ACE. In addition to getting an early start on interviewing, the benefits include providing experience for our supervisors and a final testing of our automated system. The Bureau, of course, has had extensive experience with telephone interviewing. We designed this phase of the ACE based upon our testing of the methodology in the dress rehearsal.

As you know, we do not begin personal visit interviewing until nearly all nonresponse followup work is completed in all the ACE block clusters in an LCO. This is one of the ways we preserve inde-

pendence between the ACE and the census. If nonresponse followup and ACE, field interviews are working the same areas simultaneously, they could affect each other's work, and that's why we wait to complete nonresponse followup before starting the personal visits.

Interviewers, whether on the telephone or personal visit, focus on reconstructing the Census Day household; that is, determining who lived at the address on Census Day at the time of the ACE interview and collecting as much information as possible for those who lived at the address on Census Day but have moved out, so we also have special procedures, of course, for movers.

All of these interviewers will use the CAPI technique. This is a technique that improves the accuracy of the operation because it permits a more structured interview and more probing questions. We have extensive processes for conducting quality assurance to identify data quality or falsification problems, though for data quality purposes we do not widely publicize these processes. Most personal visit interviewing will be conducted in late July or August, but some may begin in mid-to-late June. Personal visit interviews are conducted only with a household member during the first 3 weeks if the case is available. If an interview is not obtained after 3 weeks, interviewers will attempt to interview another knowledgeable person, and during this latter part we use our very best interviewers of course who are trained to convert reluctant respondents.

We then do person matching. This will occur in October and November. Census Bureau staff conduct the various stages of the matching of persons listed in the ACE interviewing to those persons counted in the same block clusters as part of the initial census, and we have designed, of course, these matching processes to minimize errors. Incorrect matching determinations generally result from incomplete, inaccurate or conflicting data or from poor judgment, and so we have several stages at which we conduct this matching process each with its own quality assurance process.

We then turn to dual system estimation. We use data from the ACE and the census to estimate the true population using a statistical technique called dual system estimation. The DSE will be conducted for each of over 400 groupings of people or post-strata. The dual system estimator of true population is then used to calculate a coverage correction factor for each post-stratum, which is the ratio of the DSE to the initial census count. The variables that define the post-strata grouping include race, ethnicity, age, sex, owner and nonowner, return rates, whether in or out of a metropolitan area and, if in, the size of the area, the type of census enumeration method.

These are characteristics that our research indicates are correlated with a likelihood of inclusion in the census. An example of one post-stratum is non-Hispanic Black males age 18 to 29 in nonowner units in mail-out mail-back areas of metropolitan areas with 500,000 or more people in a tract with a low return rate in the census.

Coverage correction factors are then applied the census files. For example, if the coverage correction factor for a non-Hispanic Black male in the specific post-stratum described above is 1.02, this means the Census Bureau measured an undercount of 2 percent for

this post-stratum and for every 100 people counted in the census in these areas two records will be added. This process is sometimes called synthetic estimation. After this, the corrected census file can be used to produce the corrected tabulation for all uses of census data.

Mr. Chairman, I have tried to give a rather simple and quick basic description of the ACE and the documents listed in the appendix, and, of course, this fuller set of documents can be investigated for any more detailed questions that you might have. Thank you sir.

[The prepared statement of Mr. Prewitt follows:]

PREPARED STATEMENT OF

KENNETH PREWITT

DIRECTOR, U.S. BUREAU OF THE CENSUS

Before the Subcommittee on the Census

Committee on Government Reform

U.S. House of Representatives

May 19, 2000

Mr. Chairman, Mrs. Maloney, and Members of the Subcommittee:

It is a pleasure to be here today to report on the Accuracy and Coverage Evaluation (A.C.E.), which will provide a report card on Census 2000 enumeration operations. The Census Bureau takes much professional pride in carefully reporting to the Congress and to the American people how it has done. At previous hearings, we have discussed the many operations that the Census Bureau is undertaking to make Census 2000 as complete and accurate as possible. Census 2000 operations are more robust and innovative than those for any previous census. Operations are going well. I have noted that the mailback response rate to the census--now 66 percent--is encouraging, as is the fact that, to date, we have completed 39 percent of the nonresponse followup workload.

Still, the Census Bureau does not anticipate at this point that Census 2000 will have better coverage than the 1990 census because many of the factors that led to the undercount in 1990 are still present in American society today. The Census Bureau has both measured and documented the existence of a substantial undercount since the 1940 census. This documentation reveals the persistence of a large differential undercount, existing for at least the last 60 years, between the Black and non-Black populations. The 1990 census indicated that the differential undercount is not purely a phenomenon of the Black population. Children, renters, Hispanics, Asians and Pacific Islanders, and American Indians were also found to be disproportionately undercounted.

The Census Bureau has been running harder, but believes this will only allow us to stay even. That is, we expect that neither the overall coverage levels nor the differential undercount rates in Census 2000 will show improvement over 1990. The Census Bureau strongly hopes to be proven wrong in this assessment and the A.C.E. will give us the information to determine this.

The A.C.E. will provide a final quality check on how well we have done in the initial census. The alternative is to not do the A.C.E. and never know how we have done below the national level (where demographic analysis provides a benchmark). The A.C.E. will also provide the means to generate more accurate counts. The 1990 version of the A.C.E., the Post-Enumeration Survey (PES), provided information that was used during the 1990's to improve statistical programs. The population estimates that the Bureau of Labor Statistics (BLS) asked us to incorporate into the Current Population Survey program following the 1990 census were corrected for the "undercount" identified through the PES. The BLS also requested adjusted population controls for the Consumer Expenditure Survey. All other major national demographic surveys conducted by the Census Bureau for other agencies of the federal statistical system also were converted to this adjusted population base.

Katherine Abraham, Commissioner of the Bureau of Labor Statistics, wrote me on March 11, 1999, that if the BLS had used uncorrected numbers, its estimates of the overall level of employment and unemployment would have been too low and the published geographic and demographic distribution of unemployment and other measures would have been inaccurate.

Mr. Chairman, in your letter of invitation you asked that I **be prepared to provide the status and a brief overview of the A.C.E. 2000 methodology and operational time line, and readiness for key activities and dates that lay ahead. Please include, but do not limit your testimony to the following activities: address listing, sample size and selection, person interviewing, person matching, Dual System Estimation, synthetic estimation, release of adjusted population counts and underlying data, and plans to evaluate the adjustment.** I will discuss the last two points first.

The Census Bureau is committed to making its data as accurate as possible for all uses of the data. In accordance with the 1999 Supreme Court ruling, the Census Bureau will not use statistical sampling (the A.C.E.) to produce state population totals used for congressional apportionment, which must be produced by December 31 of this year. The Census Bureau does plan to use statistical sampling techniques to produce the more detailed data required for redistricting and federal program purposes because we believe we can produce more accurate data by incorporating proven sampling methodologies into traditional enumeration procedures. Section 141(c) of Title 13 requires the Census Bureau to report redistricting data directly to the states by April 1, 2001. Public Law 105-119, the Census Bureau's appropriations bill for FY 1998, requires the Census Bureau, upon release of redistricting numbers based on statistical sampling, also to release comparable data produced without the use of sampling, and we plan to do so.

I have said previously that the Census Bureau currently expects that the corrected numbers using A.C.E. will be the more accurate numbers. If the Census Bureau does not have confidence in the A.C.E. results, we will not use it. The decision whether to release the statistically corrected data should take into consideration operational data to validate the successful conduct of the A.C.E.,

whether the A.C.E. measurements of undercount are consistent with historical patterns of undercount, and a review of selected measures of quality. In the fall of this year, the Census Bureau will discuss the review process and criteria with the statistical community and other interested parties.

All major operations of the A.C.E. have been designed and documented and the details have been made available for review and comment. The National Academy of Science's Panel to Review the 2000 Census, chaired by Janet L. Norwood, was convened in the fall of 1998 to review the methods, procedures, and results of Census 2000. Among other things, the Panel is reviewing the statistical methods and operations of the A.C.E. and Dual System Estimation. The Census Bureau prepared several documents for a Panel workshop in February, and these documents, listed in the Appendix, were made available then and by separate letter in early April to you, Chairman Miller, and to Mrs. Maloney, Chairman Rogers, and Mr. Serrano. Subcommittee staff were briefed on these documents by John Thompson and Howard Hogan on April 17. (The Appendix also lists and notes one document that was prepared subsequently.)

STEPS IN THE A.C.E. PROCESS

Now, I will describe the basic A.C.E. operations. Several major A.C.E. operations have been completed, one is ongoing, and others will follow the completion of nonresponse followup. All operations are currently on schedule. The basic concept behind the A.C.E. is the comparison of data from two systems--an independent survey of about 314,000 housing units (including Puerto Rico) and the initial census. Because of its small size relative to the initial census, we believe we can do a better job enumerating people in the housing units in the sample. We can be more selective about the interviewers, train them longer and pay them more, and provide more quality assurance.

Sample Design.--The first step in the A.C.E. process is to design and select the sample, which consists of approximately 314,000 housing units, or about one-fourth of one percent of total housing units. We selected the initial sample in the first half of 1999 and subsampled in early 2000 to provide a more efficient sample design and meet our field requirements. The basic units of the sample are what we call "block clusters"--single census blocks or clusters of contiguous census blocks, about 11,800 in all. The sample was designed and selected to provide sufficient precision to estimate the true population for various groupings of the population that we call "post-strata," which I will describe below.

Independent Listing.--The next step in the process is to create an independent listing of housing units in A.C.E. block clusters. By independent, we mean that we do not start with or refer to the Census 2000 Master Address File, but, instead, have census staff systematically canvass the block clusters to list addresses. This listing was completed in the fall of 1999, and then addresses were checked in, keyed, and 100-percent quality controlled in our National Processing Center.

Housing Unit Matching and Followup.--The Census Bureau then matched the A.C.E. list of housing units to the Census 2000 Master Address File, first attempting to match by computer and then clerically, if necessary. Addresses that required additional information for matching were sent to a field followup operation. Clerical staff reviewed the information from the field followup and attempted to match the addresses. The purpose of the housing unit match is to create an accurate linked list of housing units in the block cluster. During the housing unit match stage, we learned that a few original A.C.E. listers may have erroneously listed the wrong area. In these cases, we conducted an independent relisting operation. The A.C.E. housing unit match and followup operation was completed in mid-April.

Telephone and Personal Interviewing.--To provide sufficient data to compare the A.C.E. to the initial census, the Census Bureau must conduct interviews to collect data from each of the housing units that were independently listed. We initiated the A.C.E. interviewing with a telephone phase using laptop computers, a technique called Computer Assisted Personal Interviewing (CAPI). Census Bureau staff began in late April telephoning households at unique addresses in the A.C.E. block clusters for which a Census 2000 questionnaire has been mailed back, processed through data capture and for which a telephone number was provided. As of May 15, we had completed over 56,000 interviews by telephone, or approximately one-sixth of our total A.C.E. workload. In addition to getting an early start on interviewing, the benefits include providing experience for our supervisors and a final test of our automated systems. The Census Bureau has extensive experience with telephone interviewing. The Census Bureau designed this phase of the A.C.E. interviewing to avoid violations of the independence requirement and tested this methodology in the Dress Rehearsal. We are confident that the gains in quality from using the telephone outweigh any potential loss in independence.

We do not begin personal visit interviewing until nearly all nonresponse followup work is completed in all the A.C.E. block clusters in an LCO. This is one of the ways we preserve independence between the A.C.E. and the census. If nonresponse followup and A.C.E. field interviewers are working in the same area simultaneously, they could affect each other's work. That is why we are waiting to complete nonresponse followup before starting A.C.E. interviews by personal visit.

Interviewers (telephone or personal visit) focus on reconstructing the Census Day household, that is, determining who lived at the address on Census Day and at the time of the A.C.E. interview and collecting as much information as possible for those who lived at the address on Census Day but have moved out by the time of the A.C.E. interview. People who moved into an address after Census Day are also interviewed to provide additional information on movers. Only short form questions are asked together with more probing questions to establish correct residence.

Interviewers (telephone or personal visit) with laptop computers will use the CAPI technique. The questions to be asked are displayed on the computer screen and responses are entered directly into the computer. The Census Bureau believes this technique improves the accuracy of the operation, because it permits a more structured interview and more probing questions. We

have extensive processes for conducting quality assurance to identify data quality or falsification problems. For data quality purposes, we do not widely publicize these processes.

Most personal interviewing will be conducted in July and August, but some may begin in mid- to late June. Personal interviews are conducted only with a household member during the first 3 weeks that the case is available for interviewing. If an interview with a household member is not obtained after 3 weeks, interviewers will attempt to interview another knowledgeable person. During the latter part of the operation, the best interviewers are used to convert as many noninterview cases as possible to completed interviews, either by talking to a household member or another knowledgeable person. This nonresponse conversion has been planned to improve the completeness of data for matching.

Person Matching.--In October and November, Census Bureau staff conduct the various stages of matching persons listed in the block clusters during A.C.E. interviewing to those persons counted in the same block clusters as part of the initial census. The Census Bureau has carefully designed the A.C.E. in order to minimize matching errors. Incorrect matching determinations generally result either from incomplete, inaccurate, or conflicting data or from poor judgment in deciding whether a match occurs. Thus, we have several stages in the process of person matching.

First, we attempt to match people by computer. Those that do not match or are only "potential" matches are reviewed by clerical staff who use an automated computer match and review system to identify additional matches. After the initial stages of computer and clerical matching, selected cases--including some nonmatches and potential matches--are sent to field followup to obtain additional information that would facilitate matching. After followup, the Census Bureau conducts a final clerical matching operation.

At the conclusion of this intensive matching and followup effort, some information will still be missing. This may involve characteristics for individuals or cases where it could not be determined whether a person matched or was correctly enumerated in the initial census. The next step in the process is to statistically impute missing characteristics for individuals and for cases where it could not be determined if there was a correct enumeration.

Dual System Estimation.--Using data from the A.C.E. and the census, the Census Bureau will estimate the true population using a statistical technique called Dual System Estimation (DSE). The DSE will be conducted for each of over 400 groupings of people or post-strata. The Dual System Estimator of true population is then used to calculate a "coverage correction factor" for each post-stratum, which is the ratio of the DSE to the initial census count.

The variables that define the post-strata groupings include race, ethnicity, age, sex, owner/nonowner, return rates, whether in or out of a metropolitan area and, if in, the size of the area, and type of census enumeration method--characteristics that our research indicates are correlated with a likelihood of inclusion in the census. An example of one post-stratum is

non-Hispanic Black males, age 18-29, in nonowner units, in mailout/mailback areas of metropolitan areas with 500,000 or more people in a tract with a low return rate in the census.

Synthetic Estimation.--Coverage correction factors are then applied to the census files. For example, if the coverage correction factor for non-Hispanic Black males in the specific post-stratum described above is 1.02, that means the Census Bureau measured an undercount of 2 % for this post-stratum and for every 100 such people counted in the census in those areas two records will be added. This process is sometimes called "synthetic estimation." After this, the corrected census files can then be used to produce corrected tabulations for all uses of census data.

Mr. Chairman, I have tried to give a simple, basic description of the A.C.E. and to address each of the issues raised in your letter of invitation. The documents listed in the Appendix, all of which have been publicly available for some time, go into much more detail about a wide range of A.C.E.-related issues that we have examined. I will now answer any questions you may have.

APPENDIX

Summary of Documents for National Academy of Science
Dual System Estimation Workshop
February 2 and 3, 2000
(With additional document noted.)

	Title	Author	Content
1	A.C.E.: Theory and Application	Howard Hogan	Gives a detailed discussion of the model for dual system estimation for census applications. Provides the theoretical foundation for the operational aspects as related to A.C.E.
2	A.C.E.: Overview of the Design (S-DT-2)	Danny Childers, Debbie Fenstermaker	A very brief summary of the basic A.C.E. operations. Shows how the various stages of sampling, matching and field work interrelate.
3	A.C.E.: The Design Document (S-DT-1)	Danny Childers	A very detailed description of the A.C.E. design. Gives a thorough accounting of procedures and rules for nearly all facets of A.C.E. Can be used as a reference manual.
4	A.C.E.: Targeted Extended Search Plans (# Q-18)	Alfredo Navarro	Provides a description of plans and conditions for extending a search for person matching outside the sample block cluster. Plans call for a certainty and sample based components. Estimator spelled out.
5	A.C.E.: Missing Data Procedures (# Q-19)	Patrick Cantwell	Describes plans for dealing with household noninterviews, missing data items including unresolved residency, match and enumeration status.
6	A.C.E.: Dual System Estimation (# Q-20)	Richard Griffin	Includes all the details for calculating DSEs. Shows how the mover Procedure-C, missing data and targeted search are reflected in the DSE.
7	A.C.E.: Post-Stratification for Dual System Estimation (# Q-21)	Richard Griffin Dawn Haines	Describes the post-stratification model for 2000 . Shows and defines post-stratification variables as well as how multiple race responses are handled.
8	A.C.E.: Synthetic Estimation (# Q-22)	Dawn Haines	Describes the application of synthetic estimation to census data. Includes how overcounts and undercounts will be handled and also addresses controlled rounding.
9	A.C.E.: Post-Stratification Preliminary Research Results (# Q-23)	Eric Schindler	Describes a number of post-stratification models under consideration. Summarizes variance, MSE, and bias.
10	The following document was produced on April 19, 2000: A.C.E.: Final Post-Stratification Plan for Dual System Estimation (#Q-24)	Richard Griffin Dawn Haines	Presents the final post-stratification plan for the Accuracy and Coverage Evaluation Survey.

Mr. MILLER. Thank you, Director Prewitt. I'm not going to enter all that in the record. It will make it too long and lengthy. So I will appreciate having the access to that. Let me briefly make a comment about the issue of transparency and openness, and I hope like you that we get beyond this very quickly.

However, let me just explain the foundation for my concerns. Issues of transparency and access go right to the heart of one of the reasons of why I called this hearing today. The census is like a business placing an order. Last year the Congress placed a $7 billion order with the Census Bureau. This Congress had done everything it can to make sure we had the money to pay. Now the inventory is coming in, and I am equally responsible to those Senators and Representatives and the people they represent to make sure we get what we paid for. It's my job to check the inventory.

It's also the job of the GAO, the monitoring board and the Inspector General and the National Academy of Sciences to review the 2000 census. Unfortunately, when I tried to check the inventory, the Bureau tells me I can't open certain boxes or I have got to wait 3 weeks while they check with headquarters or there are some boxes that are off limits. Let me give you a few examples.

Last year the Bureau refused to provide this subcommittee with data from the 1990 census. They claimed it was protected by title 13 and it wasn't. And we are all sworn anyway, but the result was to delay the request for months.

The Bureau also delayed providing information requested for the dress rehearsal for an entire year, effectively preventing analysis. Two months ago, I entered into the record a list of information requested by the monitoring board that were delayed by more than 60 days or refused. The Bureau also produced a set of guidelines that limited access to local offices by the GAO, the board and the subcommittee. And just last week we received a copy of an e-mail that gave me, my staff and representatives of the GAO reason to believe a Bureau employee was instructed to withhold information and was instructing subordinates to do the same.

This week I received a letter from the Director requesting that I not call any Bureau employees without a Democratic staffer or a member of the Bureau present. Is this what is meant by transparent census free from political manipulation? If this is the routine during the relatively simple census, delayed information, limited access and obstructed investigations, how can we have confidence in the extremely complex statistical adjustment?

How can we honestly say this process is free from political manipulation if we are not allowed to review the process, or if we are only allowed to look at certain parts of it under certain conditions with proper supervision of Democrats and Bureau employees?

These developments are increasingly troubling and do not add to the credibility of this or future decennial censuses.

My comments certainly do not reflect on you, Director Prewitt, and I think the people behind you. The nature of the concern is there is a contempt for Congress and the responsibility that we have, as the elected officials have for overseeing, not only want the $7 billion of the census, but the critical role the census is for our entire electoral process.

It is rare that we have a copy of an e-mail like that and it is legitimate for us to be looking at that.

Let me make another comment about the use of the PES and the BLS adjusted numbers. My understanding is that the BLS accepted adjusted numbers for very large areas, national populations, and for large States. But you are proposing releasing adjusted numbers for every State, county, city, and block in the country. That is a completely different use of the numbers than what BLS is using. At small levels of geography, adjusted numbers are not reliable, and in fact the Census Bureau doesn't use the adjusted numbers. Which raises a question of I see in your statement instead of using "adjusted" numbers you started using "corrected" numbers, which implies that you have already decided that the adjustments are the correct ones, which to me almost politicizes the use of that word. So I am concerned that you start saying, well, this is the corrected one. You have obviously made a decision—not obviously but apparently instead of calling it adjusted numbers you are using corrected numbers and that is a political way to refer to those numbers.

Dr. Bryant, the Director of the Census Bureau in 1990, originally supported adjusting the national numbers. But she decided not to adjust the intercensual estimates after extensive evaluation by the CAPE Committee showed the 1990 adjustment was 45 percent error. Let me read from her decision, this is from the CAPE report: "Work suggests, the CAPE committee's work suggests that no survey, either the high quality, well-controlled and interviewed 1990 PES of 170,000 households or a larger one can be used to make a post census fine-tuning of an average undercount as small as 1.6 percent in all types of places, counties, and States. Given that, from little or no evidence that adjustment would improve the quality of sub-state estimates other than for a limited number of large places, the decision is not to adjust."

This is from the December 1992 Federal Register, which I would like to enter into the record, and without objection it will be included in the record.

[The information referred to follows:]

1ST DOCUMENT of Level 1 printed in FULL format.

FEDERAL REGISTER
VOL. 58, No. 1

Notices

DEPARTMENT OF COMMERCE (DOC)
Economics and Statistics Administration
Bureau of the Census

[Docket No. 920895-2347]

Decision of the Director of the Bureau of the Census on Whether To Use Information From the 1990 Post-Enumeration Survey (PES) To Adjust the Base
for the Intercensal Population Estimates Produced by the Bureau of the Census
ACTION: Notice of final decision.

58 FR 69

DATE: Monday, January 4, 1993

[*69]

SUMMARY: This is a notice of the final decision of the Director of the Census Bureau on the issue of whether to use information from the Post-Enumeration Survey (PES) to adjust the base for Intercensal Population Estimates produced by the Bureau of the Census.

DATES: The decision is effective on December 30, 1992.

FOR FURTHER INFORMATION CONTACT:

Peter Bounpane, Assistant Director, Decennial Census, Bureau of the Census, Telephone (301) 763-5613.

December 29, 1992.
SUPPLEMENTARY INFORMATION: Title 13, *U.S. Code, section 181,* states that: During the intervals between each census of population required under section 141 of this title, the Secretary, to the extent feasible, shall annually produce and publish for each State, county, and local unit of general purpose government which has a population of 50,000 or more, current data on total population and population characteristics and, to the extent feasible, shall biennially produce and publish for other local units of general purpose government current data on total population. Such data shall be produced and published for each State, country, and other local unit of general purpose government for which data are compiled in the most recent census of population taken under section 141 of this

title. Such data may be produced by means of sampling or other methods, which the Secretary determines will produce current, comprehensive, and reliable data.

This authority is delegated by the Secretary of Commerce to the Director of the Bureau of the **Census.**

On August 10, 1992 and September 17, 1992, notices were published in the [*070] Federal Register (57 FR, Nos. 154 and 181, pp. 35562-35564 and pp. 42939-42940), informing the public about alternatives available to the Director of the **Census** Bureau for potential improvement in the intercensal estimates of population and to seek comments on the alternative options on this issue. In addition, a public hearing was held on August 31, 1992 at the Bureau of the **Census** to provide the public the opportunity to present views on this matter and to give the Bureau of the **Census** the opportunity to hear the comments of interested parties.

The decision process was divided into several distinct phases:

Research Phase

A program of research was undertaken following the July 15, 1991, decision of former Secretary of Commerce Robert A. Mosbacher not to adjust the 1990 **census,** and continued through November 1992. Following his decision, Secretary Mosbacher urged the Bureau of the **Census** to continue the research that was started in 1990 to determine whether the **census** should be adjusted using data from a post-enumeration survey (PES) and a method of population estimation called the dual system estimate (DSE) which used both **census** and PES results. Although former Secretary Mosbacher determined the research results were not usable for adjustment of the **census,** he felt that -- with continued work -- results might be used to adjust the base of intercensal population estimates to improve those estimates. Intercensal estimates are not prepared for **census** tracts and blocks, or used for redistricting, as are **census** data. The additional research following former Secretary Mosbacker's decision was done under the direction of a senior level group of Bureau of the **Census** statisticians and demographers who made up the Committee on Adjustment of Postcensal Estimates (CAPE). The Director of the Bureau of the **Census** was present at virtually all meetings of the CAPE and examined carefully all research reports of CAPE during the July 1991-November 1992 period of its work.

Public Commentary Phase

The Federal Register notices of August 10, 1992, and September 17, 1992, invited comment from the public on the issue of whether to use data from the PES to adjust the base for intercensal population estimates, and on public preferences among five options available to the Director. 1,118 individuals and organizations responded by the end of a three month period of commentary which ended November 13, 1992.

Public Hearing

On August 31, 1992, 17 individuals, some representing organizations and coalitions, testified at a public hearing on the issue.

Evaluation

In making my decision, I relied on the results of research produced under the direction of the Committee on Adjustment of Postcensal Estimates (CAPE), advice from members of CAPE and senior Bureau of the Census officials, and from a panel of outside experts who studied the matter under the aegis of the CAPE and spent a day at the Bureau of the Census. In addition, I considered all public comments received by mail and at the public hearing. These comments, as well as the appendices, CAPE Report and Addendum, referred to in the following explanation of the decision, are available for public inspection in the U.S. Department of Commerce, Central Reference and Records Inspection Facility, room 6020 Herbert C. Hoover Building, 14th and Constitution Avenue, NW., Washington, DC 20230.

Following is a detailed discussion of the decision and the basis for the decision. The discussion is in six sections:

I. Summary Statement of the Decision

II. Uses of Intercensal Population Estimates

III. Research Input into the Decision

IV. Analysis of Options Available to the Director of the Bureau of the Census

V. Adjusted Population Estimates for Survey Controls but Official Unadjusted Estimates for Use in Administering Any Law in Which Population, or Population Characteristics, Are Used to Determine Benefits

VI. Summary of the Public Commentary and Public Hearing

Date: December 29, 1992.

Barbara Everitt Bryant,

Director, Bureau of the Census.

I. Summary Statement of the Decision

Intercensal population estimates using the 1990 census as their base will not be adjusted to correct for an undercount, currently estimated as 1.6 percent nationally, in the 1990 census.

The small overall undercount makes adjustment impossible to do accurately at all levels of places, counties and states at which intercensal estimates are produced. Intercensal population estimates are made by the Bureau of the Census for the United States, the 50 states and the District of Columbia, and for 44,055 substate areas.

This decision was difficult to make because it is the unanimous opinion of senior statisticians and demographers at the Bureau of the Census comprising the Committee on Adjustment of Postcensal Estimates (CAPE) that adjustment would improve the accuracy of the 1990 census base at the national level. There is substantial consensus, but not unanimity of opinion, among CAPE members that adjustment would improve the distribution of population shares among the states,

but not necessarily the share of every single state. Below the large area level, the research does not show that adjustment improves the accuracy of population estimates.

Having made the decision not to adjust the official intercensal population estimates, I am making the decision that sponsors of Federal sample surveys conducted by the Bureau of the **Census** will be offered the option of having their surveys calibrated to adjusted population estimates. This is described in Section V. The national surveys conducted by the Bureau of the **Census** are calibrated at large, aggregate levels where **Census** Bureau research shows adjusted estimates are, on average, more accurate.

Although the 1990 undercount is estimated as only 1.6 percent for the United States, it is higher among certain demographic groups. At the national level, at which all CADE members have confidence that the adjusted population estimate is an improvement on the **census** count, there are significant differences in the undercount between residents of owner-occupied housing and residents of rental housing. Undercount rates for Blacks, American Indians, and Hispanics, but not for Asians and Pacific Islanders, are statistically significantly different from those for non-Hispanic Whites.

Undercount Rates

Subgroup	Undercount rate (357 post-strata PES) %	Standard error (%)
Total Population	1.58	0.20
Residents of:		
Owner occupied housing	0.07	0.21
Rental housing	4.32	0.39
Persons who are:		
Non-Hispanic Whites	1.18	0.20
Males	1.52	0.23
Females	0.85	0.21
Blacks	4.43	0.51
Males	4.90	0.53
Females	4.01	0.56
Asian or Pacific Islanders	2.33	1.35
Males	3.44	1.59
Females	1.25	1.50
American Indians	4.52	1.22
Males	5.18	1.23
Females	3.86	1.24
Hispanics	4.96	0.73
Males	5.51	0.90
Females	4.39	0.71

*Not statistically significantly different from non-Hispanic Whites (.10 level).

The method of adjustment would correct for the differential undercount measured for these groups, and thus improve the measures of their distribution

throughout the population. Thus, the decision not to adjust is made particularly difficult. Unfortunately, accurate correction of the many substate areas for which the Bureau of the **Census** produces population estimates is not achievable with the present adjustment modeling methodology.

In making the decision not to adjust, I want to acknowledge that the research on undercount and adjustment methodologies conducted at the Bureau of the **Census** over the past two and one-half years has advanced knowledge substantially on both the potentiality and the difficulty of using statistical techniques to improve the accuracy of enumeration. Results of the 1990 **census** show that the population of the United States is very diverse in demographic composition and living arrangements, and becoming increasingly so. It is unlikely that direct enumeration can ever approach 100 percent coverage of the population -- achieving net coverage of over 98 percent in 1990 was an amazing feat. Looking toward the 2000 **census,** the Bureau of the **Census** will need to work closely with the Congress in deliberating over policy alternatives for developing a design for **census** taking which incorporates into the final count some type of statistical estimation of those who are missed, or choose to be missed, by all types of direct enumeration.

At large aggregate levels, one has more confidence in the use of a large scale survey to adjust the **census** than at the smaller aggregate levels. CAPE research shows that the national population estimate produced by the post-enumeration survey and dual system estimation of 252,712,821 is closer to the true population of the United States on April 1, 1990 than the **census** enumerated count of 248,709,873, although neither of these represents precisely true population count. At the national level, the population estimate can be tested by a third, independent method: Demographic Analysis. Demographic Analysis, using vital statistics and administrative records (births, deaths, immigration, Medicare and other current and historical administrative data), provides an estimate of 253,393,786 -- closer to the PES figure than to the **census** count. Demographic Analysis provides a higher figure than either of the other two methods because it is less subject to the bias of missed persons, which affects both the **census** and the post-enumeration survey.

Below the national level, the statistical test used by the Bureau of the **Census** to measure whether the **census** or the post-enumeration survey (PES)/dual system estimate produces the more accurate population shares is called loss function analysis. Loss function analysis depends upon first, building an estimate of the true population then comparing the **census** and the PES/dual system estimate to this estimate. Although the test indicates whether the PES or the **census** produces the more accurate distribution of population between states, it does not show which state estimates are improved or whether any are made less accurate.

As a check on the loss function analysis result, **Census** Bureau demographers and a demographer expert from outside the **Census** Bureau reviewed each state's estimated undercount to see if it made demographic sense, given what they know about the demographic composition of each state. For 44 states and the District of Columbia, the PES/dual system estimates of undercount appear logical. That is, given the proportions and concentrations of different demographic groups in each state, the mix of rental and owner-occupied housing, and measured undercount patterns for these, the undercount in relation to other states was what demographers might expect. For three states, however, the measured undercount was somewhat more than they would anticipate; for three it was

somewhat less than they would anticipate.

From analysis of the extensive research CAPE produced, I conclude, with CAPE, that (a) adjustment would improve the accuracy of the national level count of the total population and of demographic groups by sex, race, Hispanic origin and tenure (residents of owner-occupied or rental housing) and that (b) adjustment would improve the overall distribution of population shares among the states.

At substate levels, however, CAPE was unable to conclude with reasonable certainty whether adjustment would improve or do harm to the estimates for counties and places. CAPE made no recommendation of whether or not to adjust for substate levels. While in some states loss function analysis provides statistical evidence that the population share of large places with high concentrations of those in undercounted demographic groups is improved versus the population share of the balance of their states, this is not the case within all states with such large places. Below the large area level, the research does not show that adjustment improves the accuracy of population estimates.

None of the options available to me -- not adjusting, partial adjustment or full adjustment -- is ideal. There is no perfect answer to, "What was the population of the United States on Census Day, April 1, 1990?" Any statistical and demographic research can only answer that question with some degree of uncertainty. After more than two years of Bureau of the Census research evaluating the 1990 census, the post-enumeration survey which followed it, and statistical models for making an adjustment, I must make a decision now. CAPE's recent work has solved some of the problems and criticisms which caused former Secretary Robert A. Mosbacher to decide not to use the PES/dual system estimate for adjustment of the 1990 census. Unfortunately many issues remain. Some sources of bias cannot be removed. Some of the concerns with the PES/dual system estimate expressed by CAPE and in the public commentary are not resolved and probably could not be with more time and research. While I strongly support continued research on these issues for the year 2000 census, I also feel strongly that the debate over adjustment of the 1990 census and related intercensal estimates should not continue further into the decade.

The CAPE work with the 1990 PES has been intensive for over two years. It has occupied the time and talent of the most senior statisticians and demographers within the Bureau of the Census. Their work suggests that no survey -- either the high quality, well controlled and interviewed 1990 PES of 170,000 households or a larger one -- can be used to make a post-census fine tuning of an average undercount as small as 1.6 percent in all types of places, counties and states at a level of accuracy beyond that by which surveys are usually judged. The preponderance of evidence suggests that adjustment would be a tradeoff of small errors in some states, counties and places for large overall improvements at the state and national level -- but there would be errors. However, as the public commentary gives evidence, adjustment would not be acceptable to those localities harmed by such a decision. The stakes for communities and states [*072] in population figures are high, as the uses of intercensal estimates in the next section demonstrate.

Statisticians among those who made public comments, and members of CAPE express concerns about aspects of the PES/dual system estimate, about the total error model and loss function analysis used to evaluate it and about possible

errors in the PES as a measurement tool compared to the level of undercounts and overcounts it seeks to measure. Given that there is little or no evidence adjustment would improve the quality of substate estimates, other than for a limited number of large places, the decision is not to adjust the base for intercensal population estimates.

II. Uses of Intercensal Population Estimates

The Bureau of the Census produces intercensal population estimates for the United States, metropolitan areas, states, counties, county subdivisions, places (cities, towns, townships, etc.), and consolidated cities. They are used by Federal, state, and local governments, and the public and private sectors in a variety of ways. The Bureau of the Census investigated Federal formula program uses and other major Federal uses n1 and identified five major Federal uses:

n 1 Sources: General Accounting Office, FEDERAL FORMULA PROGRAMS Outdated Population Data Used to Allocate Most Funds. (September 1990).

General Accounting Office, FEDERAL FORMULA PROGRAMS Adjusted Census Data Would Redistribute Small Percentage of Funds to States (November 1991).

General Accounting Office, Catalog of Federal Aid to States and Localities (1987).

Superintendent of Documents, CFDA (1989, 1990, 1991, 1992).

U.S. House of Representatives, United States Code (1988).

Conversations with grant administrators.

1. As the basis for Federal formula program funds allocations. Funding formulas are written by the Congress, not by the Bureau of the Census. Existing formulas were written over many years and incorporate population estimates as all, or part, of the variables used in these formulas in different ways. Some formula program funds allocations use population estimates while others use the decennial census throughout the decade. Title *13, U.S. Code, section 183*(a) requires use of the current population estimate, which according to this decision will be the unadjusted estimate, for programs which provide Federal benefits, unless in section 183(b) the formula has been specifically tied to the decennial census.

Some formulas use population shares for all persons in states, or for particular age groups, or for those with particular characteristics as some portion of the formula; other formulas use per capita figures, such as per capita income, for which population estimates are the denominator, or a ratio of state to U.S. per capita income with state and national population estimates as the denominators. Some of these programs use the population estimates in several ways in their formulas. The largest of these funding programs is Medicaid, title XIX, which uses 50 percent of the ratio of state to U.S. per capita income as one part of its formula.

2. As the denominator for per capita, and incidence rates, such as per capita income; particular health conditions per 1,000 persons; number of crimes per 1,000 persons, etc.

58 FR 69, *072

3. As the determinant of the volume cap for tax-exempt private activity bonds issued within a state. Section 146 of the Internal Revenue Code mandates use of the most recent Bureau of the **Census** state population estimate issued before the beginning of the calendar year in which the bonds are issued for setting the cap for tax-exempt private activity bonds. Small states are guaranteed a minimum funding level, but states with over three (3) million population receive an additional $50 in bonding authority for each person in the population estimate. Unlike uses of population estimates in many Federal funding formulas, bonding authority for each state is based on the absolute size of the state's population, not on its proportional share.

4. As the basis for calibrating Federal government and other sample surveys. By calibrating sample survey estimates to independent estimates of the population, accuracy of survey estimates is improved. For Federal government surveys conducted by the Bureau of the **Census**, the current population estimate is used as the control total for the total population and population groups by sex, age, race, Hispanic origin. They are calibrated geographically by the population estimates for the 50 states, District of Columbia, New York City and Los Angeles. Because Federal surveys conducted by the Bureau of the **Census** are calibrated only to these large aggregated areas, at which CAPE considers adjusted population estimates an improvement over the unadjusted estimates, the Bureau of the **Census** will offer survey sponsors the option of having their surveys calibrated to adjusted population estimates beginning in 1993.

5. As descriptive statistics to provide the most current profile of the population of the Nation, state and substate areas, and as the numbers for determining population gains and losses compared to prior years.

III. Research Input Into the Decision

When former Secretary of Commerce Mosbacher made the decision not to adjust the 1990 **census**, he expressed several major concerns with the PES and dual system estimation model that would have been used for adjustment purposes. However, he did recognize that both the PES and Demographic Analysis showed an undercount and a continuation of the historical differential undercount of certain groups. He directed that research continue at the Bureau of the **Census** to determine the possibility of incorporating adjustment into the intercensal estimates, which are made with less geographic detail than the **census**.

The Bureau of the **Census** established the Committee on Adjustment of Postcensal Estimates (CAPE), comprising 13 statisticians and demographers, plus the Director and Deputy Director of the Bureau of the **Census** as ex-officio members. Eight of the 13 members had also served on the Undercount Steering Committee which directed the research prior to July 1991.

CAPE deliberations were not constrained by pre-specification of procedures and of the dual system estimating model. Such prespecifications were required pursuant to court order in considering whether to adjust the 1990 decennial **census**. CAPE also had the advantage that many of its members had worked extensively with PES data and gained insight on possible ways to improve the dual system estimate model.

CAPE determined that the necessary research could not be completed in time for potential incorporation into 1991 intercensal estimates. Instead, they

58 FR 69, *072

scheduled this work in time to provide it for the Director's decision on whether or not to incorporate adjustment into the estimates to be released in late December 1992 for July 1, 1992.

CAPE research focused on five key areas of technical concern with the population estimate model as of July 1991. This research has been able to answer all of these concerns to some degree, but not all completely. n2

n 2 This research is described in more detail in "Report of the Committee on Postcensal Estimates," August 7, 1992 and Addendum, November 1992 (Bureau of the Census, Washington, DC 20233)

Concern 1

Could the problems in the smoothing model, including lack of robustness, be resolved? [*073]

Background

The pre-July 1991 modeling of the PES was done by dividing survey respondents into 1,392 post-strata, or groupings, based on census division, (geographic), type of place of residence, race, Hispanic origin, sex, age, and some post-strata by whether residents owned or rented their housing units. The rationale for each post-stratum is the assumption that persons within it have similar probability of having been counted in the census or the PES (although not necessarily the same in each). This is called the homogeneity, or synthetic, assumption. An undercount rate was computed for each post-stratum by matching names and characteristics of people enumerated in the PES to the census to determine who had been counted and who not, then producing an adjustment factor for each post-stratum to make the PES dual system estimate of the population. One problem in dividing the sample respondents into 1,392 post-strata was that sample sizes for some strata were very small. A statistical process known as "smoothing" had to be done to reduce variability from sampling error. Results varied by how some of the outliner post-strata (those with unusually high sampling error) were treated. Because of these variations, statisticians felt the model was not "robust" to changes in model specification and procedure.

Results

The PES and dual system estimate have been redesigned to used only 357 post-strata, thereby gaining enough sample size in each stratum to make smoothing unnecessary. Analysis of PES data showed some characteristics more important than others in identifying whether persons had been counted in the census. Besides race and Hispanic origin characteristics, living in owner occupied or rented housing proved to have an effect, as did living in an urbanized areas of 250,000 or more versus living in other urban or non-urban areas. A more effective division made it possible to reduce the number of post-strata. CAPE now considers the 357 post-strata model robust. Sampling error has been reduced as a source of error, with some loss of homogeneity. Both CAPE and I have used some caution in interpreting results of loss function analyses since these are highly dependent upon the homogeneity assumption (See discussion under Concern 4).

Concern 2

58 FR 69, *073

Could the estimated nonsampling error biases in the PES estimate of undercount be removed?

Background

In the research prior to July 1991, Bureau of the **Census** statisticians identified biases (nonsampling errors) that amounted to 0.7 percent of the then reported 2.1 percent undercount.

Results

Matching experts matched PES respondents to the **census** in the 104 blocks that had the most effect on undercount. This removed some bias and reduced the national estimate of undercount by 0.1 percentage points. During this analysis, the Bureau of the **Census** also found and corrected a computer error in the matching that had led to an overstated estimate of undercount. This correction reduced the undercount estimate an additional 0.4 percentage points. Matchers also rechecked matching in blocks with a preponderance of Hispanic surnames. These matches proved to have been done with the same apparent accuracy of the rest of the PES matching. Beyond these improvements, CAPE could find no reliable or expedient method to remove the balance of nonsampling error bias from the PES estimates. A significant amount of bias remains. The research estimates that, at the national level, removing all biases from the PES estimates would lower the estimated undercount from 1.6 to 1.3 percent. When the effect of correlation bias is not taken into account, (CAPE believes correlation bias should be considered) the estimated undercount would fall to 0.9 percent. Correlation bias is discussed under Concern 3.

Concern 3

Were all components of bias adequately reflected in the total error model, and was total error being accurately handled in loss function analysis?

Background

The total error model is used for building an estimate of the "true" population without bias, since the exact truth cannot be known. Loss function analysis is a statistical technique used to compare the **census** count and the PES/dual system estimate to this "true" target population to see which is closest to it (that is, has less "loss"). Loss functions can be run for various geographic levels.

Results

Two new components of error were added to those used in modeling prior to July 1991. With these additions, CAPE felt satisfied that all components of error were represented except for homogeneity bias. (The bias of all persons in a post-stratum not having the same probability of having been counted in the **census** or PES. This is discussed under Concern 4.) CAPE was concerned about the accuracy and variance of the estimates of bias and, therefore, used caution in evaluating the results of loss function analyses since the "true" target numbers are themselves dependent on levels of estimated bias.

The majority of CAPE members felt that correlation bias should be a component

58 FR 69, *073

of total error. Correlation bias has two parts: (1) The dual system estimate assumes a person's participation in the PES is not affected by his or her participation in the census. Based on the evidence available to it, CAPE judged this was not a problem. The PES was conducted 4 months after the census and there were other controls. (2) The dual system estimate assumes people in each post-stratum have the same probability of being included in either the census or the survey. Some people may be virtually impossible to count, leading to a component of correlation bias. A national estimate was made of these people by comparing the PES estimate to the Demographic Analysis estimate, which is less subject to correlation bias. The difference was not added to the PES estimate, but included in the total error model and used to determine the target "true" target population to which the census and the PES estimate were compared. As mentioned, most CAPE members felt correlation bias should be a component of total error. However, concerns continue about measuring it and allocating it to the target "true" numbers. Loss functions, therefore, have been analyzed with and without correlation [*074] bias.

For use in loss function analysis, the "true" target population estimate was made by modifying the PES estimate by removing error in the PES identified in the total error model. A modeling system was used to allocate the direct estimates of bias from 10 evaluation post-strata (aggregated post-strata) to smaller levels of geography. Such modeling and error measurement obviously have errors themselves. CAPE decided to run loss functions in a variety of ways to judge results. CAPE ran loss functions using different ways of allocating bias to the target "true" numbers; with and without correlation bias; using absolute and squared error, as well as variations of these to take account of variations in state (or other areas of interest) size. They looked at aggregate loss (difference from truth), then ran statistical hypothesis tests and reported the significance levels. These were run to look at whether the census or the PES estimate showed the greater or lesser difference from the "true" target population shares for a variety of levels of geography for places and counties by size and for states.

Using hypothesis tests with 10 percent significance, loss function analysis excluding correlation bias does not support adjustment. Including correlation bias, all but one of the loss function analyses favors adjustment at the state level when examining aggregate loss. These results show that the state level analysis is sensitive to assumptions about correlation bias. Below the state level, CAPE was generally unable to conclude that adjustment was better.

Concern 4

Could more be learned about whether or not the homogeneity assumption holds sufficiently to support adjustment?

Background

The homogeneity (or synthetic) assumption is that every person in a post-stratum has the same probability of having been counted in the census or PES (though not necessarily the same probability in both). There were 1,392 post-strata in the earlier model. There are 357 in the current model produced by CAPE.

Results

The PES data show clearly that the variables used for dividing the country into poststrata -- particularly race, Hispanic origin, tenure (owner/renter), and type of place of residence -- account for considerable variation in census coverage. The homogeneity assumption assumes, however, that there is no further variation in coverage within each poststratum separately, that is, that everyone within a poststratum has the same probability of being counted in the census. This assumption is central to the way that PES estimates have been produced for states, counties, towns, and all other units important in postcensal estimation. The PES data by themselves have not offered enough evidence to evaluate the homogeneity assumption adequately, especially to answer the most important question for me -- does the assumption hold sufficiently so that our PES estimates based on it represent an actual improvement to the census?

Consequently, I have relied upon the available results on the homogeneity assumption from the analysis of the artificial populations, discussed extensively in the CAPE report and the Addendum to that report. n3 Because the artificial populations are based on known quantities (such as percent mail response, percent of multi-unit housing, percent in poverty, etc.), we can put the adjustment procedure based on 357 post-strata to a test. For every one of the eight surrogate variables that we examined, our results indicate that we can make adjustments that, on average, are closer to the truth than not adjusting. To represent the adjustment to the artificial population, we used the rates for the surrogate variable in each of the 357 post-strata and the homogeneity assumption to construct adjustments in the same way as the PES does for estimating undercount. I feel very confident about this finding for the artificial populations because we do know the actual values of the artificial population surrogate variables.

n 3 Ibid.

Any problem cased by departures from the homogeneity assumption, which we have called heterogeneity bias, would be less troublesome if we had been able to represent its effects in the loss function analysis. In fact, the loss function analysis for the PES data could not include heterogeneity bias because the PES data did not furnish adequate evidence on this question. Consequently, I have again relied on evidence from additional research on the artificial populations. I have considered the recent results discussed in the CAPE Addendum that compare 1) the loss function analysis as it is seen in the PES context, i.e., as the comparison of the census and PES values to targets based on the homogeneity assumption; to 2) the actual losses comparing adjusted population estimates and the census to the true values. I have been encouraged that, for 7 of the 8 surrogate variables available for study, the loss function analysis is robust. In other words, in all but one case, the loss function analysis understates or approximately correctly measures the level of improvement from adjustment, although the loss function also understates the true errors from non-adjustment and from adjustment.

The August CAPE report stated that research showed that the artificial population analysis supported the homogeneity assumption but, once 25 percent bias was introduced, the support for the assumption broke down. Furthermore, these results were then compared to estimated levels of bias in the PES, which were near the 25 percent level. Because of the clarification in the CAPE Addendum of the interpretation of these earlier findings for the artificial

populations, one should not place so much weight on the results for 25 percent bias, since the comparisons were based on numeric accuracy whereas our primary focus has been on the accuracy of population shares. In August, with the information available to us, the results for 25 percent basis, reported in the manner that they were, represented a strong case against adjustment, but I now believe that the CAPE Addendum should alleviate many of the concerns about this issue.

Concern 5

Could CAPE resolve the inconsistencies between the PES and Demographic Analysis (DA) estimates of undercount?

Background

At the time of the July 1991 decision, there were differences in PES/DSE and Demographic Analysis estimates of population for some major subgroups. Also, at that time, the PES/DSE showed a higher undercount rate than Demographic Analysis, which was unexpected because Demographic Analysis accounts for people missed in both the census and the PES.

Results

The corrections to the PES and subsequent reduction of the PES estimate of undercount to 1.6 percent cleared up the major inconsistencies between the PES and DA estimates. Demographic Analysis which, as pointed out earlier, is less subject to correlation bias, should measure a somewhat higher undercount than the PES since it includes those who are missed or impossible to count. It now does, 1.8 percent undercount for Demographic Analysis compared to 1.6 percent for the PES/DSE. Comparisons of PES and DA for states and large areas (DA cannot go down to small areas) meet face validity expectations for most states. By "face validity" I mean that the patterns of high and low undercounts are logical demographically.

In summary, CAPE developed a new estimation model based on the modeling of the PES to 357 post-strata and a more thorough identification of biases resulting from sampling and nonsampling error.

The Bureau of the Census conducted a number of studies to evaluate and measure errors in the PES. A total error model was then used to combine the [*075] results from these studies to produce measures of net error on the estimated undercount rates. The estimated level of bias in the PES that was estimated from the total error model was -0.7%, without correlation bias, or -0.4% when the effects of correlation bias were taken into account. These estimates reflected the measures of bias in national estimates of undercount. Thus the national estimated undercount rate of 1.6% is, without bias, between 1.2% and 0.9%.

Both the unadjusted and adjusted census results are imperfect. The unadjusted census is subject to undercount, and the adjusted census is subject to the error in the PES -- sampling error and bias. A key question to answer is which of these two imperfect systems will provide a more accurate base for use in the intercensal estimates program. This question is particularly relevant for population shares and was addressed through loss function analysis. n4 The loss function analysis accounts for both the sampling error and bias in the PES. The

CAPE examined the results of the loss function analysis, and concluded that an adjustment would make the distribution of population shares for states more accurate. However, they could not reach consensus on substate areas.

n 4 Loss function analysis is described in the August 7, 1992 CAPE report.

An important issue associated with loss function analysis is the methodology associated with the creation of the target population (the estimates of the true unknown population). The measures of bias in the PES were used to "correct" the PES/dual system estimate to produce the target population. Unfortunately, the PES bias was not directly measured for each state, county or city. Models were used to allocate the directly measured bias at 10 sub-national levels to each state, county, and place to produce the target population. Questions have been raised regarding whether these models reflect an accurate distribution of the bias.

I am also concerned about this issue, as was the CAPE. The CAPE recognized that there were potential deficiencies in the modeling methodology. The CAPE therefore examined a number of different models for creating target populations. They concluded in August that the bias modeling methodology was sufficiently robust to conclude that at the state level, an adjustment would improve the state base for intercensal estimates. However, they could make no conclusion about the effects of adjustment at substate levels other than documenting in the November addendum to their August report that an adjustment would also improve the distribution of population shares for large places (100,000+) in aggregate compared to state balances in aggregate.

I have considered this modeling of bias issue at great length. There are certainly problems with the modeling methodology. In most statistical applications, we never know the true situation, and no model is perfect. Despite the extensive research, too many concerns remain about the level of bias; the estimate of the true population used as the target in loss function analyses; allocation of correlation bias; and whether the homogeneity assumption holds, given the levels of bias, to make a decision to adjust intercensal population estimates defensible across the board to all 44,055 substate areas.

IV. Analysis of Options Available to the Director, Bureau of the Census

A notice that the Bureau of the Census placed in the Federal Register August 10, 1992, and repeated on September 17, 1992, described five decision options available to the Director, Bureau of the Census.

Option One: Incorporate the results of the PES into the base for intercensal estimates at all levels of geography.

CAPE research, which provided much of the input to my decision, did not support adjustment below the state level. Although taking this option would have made some improvements at the state and national level, taking this option for all levels of geography is not defensible.

Option Two: Incorporate the PES results in the intercensal base at the national and state levels. At the sub-state level, use a simple synthetic estimate based on the percentage of state-level estimated undercount.

58 FR 69, *075

I rejected this option because the simple synthetic estimate would distort within state counts artificially. For example, a small city or county that had been well counted would have its population inflated because of the undercount in a large city in the same state. Conversely, the city with a large undercount would have its population understated because of low undercounts in other parts of the state.

Option Three: Incorporate the results of the PES into the intercensal base for national and state level estimates, but not for substate levels (counties, cities, etc.).

Although this is the level of adjustment which CAPE research very clearly supports, I rejected this option because it would not be additive at state levels; for example, the sum of the population of all counties within a state would not sum up to the state total. Public commentary from data users, user groups, Federal statistical agencies, and analysts within the Bureau of the Census said nonadditive population counts would be very difficult for them to use.

Option Four: The base for intercensal estimates for all levels of geography would be a simple average of the 1990 census count and an estimate incorporating the results of the PES.

I gave serious consideration to this option. CAPE conducted additional research on this option and a modification of it after the initial CAPE report of August 7, 1992. There is precedent in the statistical community for averaging two estimates. This would have had the effect of reducing the effect of biases in the PES. However, neither I as Director nor CAPE were comfortable with a "partial" correction at the national level. This option would have had the effect of raising the population by approximately two million when the undercount is a measured four million. CAPE considered the modification of averaging the census and PES, then ratio adjusting to the PES/dual system estimate total proportionally by eight racial, Hispanic, owner/renter subgroups (what CAPE refers to as a raked composite estimate in the November Addendum to the CAPE report). I decided against this modification because, according to CAPE, the "raking" to national totals introduced more dependence on the homogeneity assumption, possibly offsetting gains from reducing the effects of bias. This treatment of a census and an evaluation survey (the raked composite) is a treatment I hope the Bureau of the Census will continue to study in designing ways to make the 2000 census more accurate.

Option Five: Do not incorporate the PES results into the intercensal estimates for any jurisdiction.

This is the option I have chosen. While adjustment would make improvements at the national and state level, it is not defensible for all levels of geography. Concerns remain about the level of bias in the PES/dual system estimate and in whether the homogeneity assumption holds, i.e., whether all persons in a post-strata in the PES have, in fact, the same probability of having been counted in either the PES or the census. These concerns do not have answers nor are further research avenues available with which to answer them in the near future. This option is not ideal, as noted previously, because it foregoes the [*076] opportunity to improve estimates at national and state levels.

V. Adjusted Population Estimates for Survey Controls But Official Unadjusted Estimates for Use in Administering Any Law in Which Population, or Population Characteristics, Are Used To Determine Benefits

Having made the decision not to adjust intercensal population estimates, I am making the decision that sponsors of Federal sample surveys conducted by the Bureau of the Census will be offered the option of having their surveys calibrated to adjusted population estimates. Federal statistical surveys are calibrated at the level of national sex, age, race, Hispanic groups, and population totals for the 50 states, District of Columbia, New York City and Los Angeles. These are the levels at which CAPE judges the adjusted estimates are more accurate than the unadjusted estimates. Reaching the demographic groups undercounted in the census is a greater problem in taking surveys than in taking the census.

In the public commentary, Senator Herb Kohl of Wisconsin and Senator Arlen Specter of Pennsylvania suggested decoupling the official population estimate from the survey controls. This is what this decision does. Unadjusted population estimates are the official estimates, required to be reported by the Secretary to the President under title *13, U.S. Code section 183*(a) for use in administering any law of the United States in which population or other population characteristics are used to determine the amount of benefit received by any State, county, or local units of general purpose government. Sponsors of surveys conducted by the Bureau of the Census will have the option of calibration to the adjusted estimates.

VI. Summary of the Public Commentary

As Director, Bureau of the Census, I received letters from 1,118 individuals or groups in response to our Federal Register notices of August 10 and September 17, 1992. I have personally read each of these.

Table 1 shows the distribution of the 1,118 responses -- 58 from Senators, 176 from Members of the U.S. House of Representatives, 37 from Governors, 349 from State Representatives, state and local government officials and agencies, 498 from all other respondents, including organizations, private citizens, and a small number from Federal agencies. Some individuals wrote a letter as well as co-signing a group letter, or wrote a second letter responding to the second Federal Register notice. Such persons are counted only once in Table 1.

Most responses stated a position for or against adjustment rather than responding to the five options of the Bureau of the Census' Federal Register notices. Preferences for the options, as well as positions for or against adjustment are summarized in Table 2. Overall, 995 favored adjustment; 123 opposed it.

In making my final decision, I considered this overwhelming preference for adjustment. However, I felt I should not use a "popular vote" as a determinant of a decision made on the basis of statistical and demographic criteria. With the exception of a few statisticians and a very few others, virtually everyone took the position which maximized the population estimate for his or her state or city rather than considering the merits of the issue. This is a problem which will persist in trying to make any adjustment after a census.

Fifteen to 20 letters had substantial statistical commentary. Some brought

out issues that CAPE was already researching. Others presented new ideas that led to some additional CAPE research. I asked CAPE statisticians to analyze these letters for me, in addition to my reading each of these letters at least twice. A summary of the comments with statistical content is available as Appendix A: Discussion of Technical Issues Raised by Outside Comment. What follows is a more abbreviated summary.

The majority, but not all, of the technical comments came from those who oppose adjustment.

Homogeneity Assumption

Several reviewers were concerned about the critical nature of the dependence upon the homogeneity assumption, questioning whether persons living in post-strata which fell across several states had the same probability of being counted. CAPE agrees that this is one of the more vulnerable aspects of the PES/dual system estimate design and recognizes that each post-strata cannot be precisely homogeneous. The key issue is whether the assumption represents an adequate approximation to the distribution of undercount. While some of the concerns expressed with homogeneity were general comments, one commenter created a U.S. map showing the high degree of association between the adjustment at the state level and the groupings of states into the four Census regions. The reviewer showed maps of other characteristics, such as poverty rate, which do not exhibit as marked a regional character as the undercount rates. Census Bureau researchers subsequently reexamined the series of characteristics employed in defining the 8 surrogate artificial populations. In varying degrees, the Census Bureau's investigations confirm that the adjustment methodology tends to emphasize regional aspects of the characteristic being estimated while missing or understating other aspects of state-to-state variation. However, among the variables defining post-strata, region captures some of the geographic variation. Others -- race, Hispanic origin, tensure type of place of residence -- are highly correlated with undercount.

Another reviewer provided calculations showing that it was possible that departures from the homogeneity assumption, that is heterogeneity, might account for more error in the PES adjustments of states than all the components of error estimated and included in the Bureau of the Census' total error model. Research on artificial populations continued after this commentary was received. These added investigations by the Bureau of the Census showed that the error due to heterogeneity tended to be large for the artificial populations. Applying these results to the PES itself, it is possible that errors due to heterogeneity in fact could be larger than all other sources of error in the adjustment. However, the investigations supported the premise that the PES adjustment could still, on average, make improvements to the overall state population shares. When heterogeneity bias is present, the results for artificial populations showed that the loss function [*077] tended to understate the errors of both the adjustment and non-adjustment. If it understated these by equal amounts, the net difference would be correct. In fact, the analyses showed that, for a majority of the 8 artificial populations (surrogate variables), the loss function would approximately correctly indicate or understate the net advantage from adjustment.

Sample Size

Some writers argued that the PES sample size was insufficient to permit an adjustment. These arguments were not reinforced by explicit calculations showing that the sample sizes were too small. The issue of sample size is linked directly to the level of sampling variance. The use of the 357-post-strata design reduced the effect of sampling variability considerably. Sampling error was a component of the total error model and was included in the loss function analysis. The loss function analysis indicates that at the state level the adjustment results in an improvement.

Post-Stratification

Several comments were received applauding the revised post-stratification (357 post-strata compared to 1,392 in the original design). However, some of these reviewers claimed that the post-stratification was data-driven. The end result of this was that the estimates of sampling error would be too low, therefore, causing the loss function analysis to unduly favor the adjustment.

One reviewer found the new post-stratification acceptable but would have preferred the Bureau of the Census to continue to use "smoothing." Alternatively he proposed to control sampling variability by collapsing the original 1,392 post-strata to a smaller number to retain greater homogeneity within post-strata. However, this reviewer felt that the revised 357 post-stratification scheme for adjusting was superior to not adjusting.

These post-stratification ideas were discussed at various times by CAPE. The clear consensus of the CAPE was not to use "smoothing," and CAPE was pleased with the new post-stratification scheme. Members recognized the danger of post-stratification after data had been examined. This had some bearing on their general concerns regarding the loss function analysis.

Correlation Bias

Correlation bias was widely discussed both internally and externally. Some reviewers noted that an adjustment based on the Bureau of the Census estimate of correlation bias would be conservative; it would not go far enough in correcting the undercount.

Other reviewers noted that at the national level there was clear evidence of correlation bias. However, they claimed that problems resulted because there were no direct measures of correlation bias sub-nationally. It was not clear to these reviewers that the method of modeling correlation bias to produce sub-national estimates was appropriate. These reviewers were not convinced that the adjustment would improve the distribution of population shares subnationally. CAPE had previously expressed many of these same concerns. The general conclusion was that correlation bias should be a component of total error. However, there were concerns about the methods of estimating and allocating it. CAPE requested that loss function analysis be done with and without correlation bias.

Total Error Model

Some reviewers viewed the total error model as being complete, and when combined with the loss function analysis supportive of an adjustment. One reviewer felt that the total error measure of correlation bias was understated and a more accurate measurement would favor adjustment more than the current

estimates.

There were others who did not believe that the total error model covered all sources of error. They cited sources of error they felt were omitted. They also felt many of the sources of error included in the total error model were not measured accurately, and cited specific ones. CAPE discussed the total error model at great length. Members felt that all components of error were included except for bias due to failure of the homogeneity assumption. However, CAPE could come to no agreement about the adequacy of the levels of error measured. CAPE concluded to use caution in evaluating the results of the loss function analysis since the target numbers used as truth were so dependent upon the levels of estimated bias. The uncertainty in the levels of estimated bias thus affects the measures of accuracy based on these target numbers.

Loss Function Analysis

Some reviewers viewed the loss function analysis as being very supportive of adjustment, and that the improvement indicated by the loss function analysis was an understatement because correlation bias was underestimated in the total error model.

Other reviewers generally had two major sources of concern regarding the loss function analysis: (1) There are components of error in the adjusted estimates that are not included in the loss function analysis, such as uncertainties from failure of the homogeneity assumption and from the choice of post-stratification. (2) There are concerns with the methods used to model the total error estimates of bias to create the target "truth" populations. One reviewer expressed concerns regarding the levels of significance for the loss function test.

CAPE discussed the loss function analysis in great detail. In particular, the comments regarding uncertainty due to failure of the homogeneity assumption led to some late research. In general, CAPE accepted the loss function results keeping in mind the caveats reflected in this report.

Other comments, which did not address the technical merits of this decision are summarized and available as Appendix B.

Public Hearing

Testimony at the public hearing August 31, 1992, came from individuals and groups, mostly groups, representing both those in favor of adjustment and those against it. All presented reasoned arguments that I listened to at the hearing and have read and weighed since. A summary of this testimony is available as Appendix C: Public Hearing, August 31, 1992, Summary of the Proceedings.

Table 1. -- Summary of Mail Response

Type of respondent	Pro	Con	Total
U.S. Senate -- Total	31	27	58
U.S. House -- Total	127	49	176
Governors -- Total	20	17	37
State/Local Government -- Total	332	17	349
Florida	205		205

Texas	55		55
California	37		37
Other States	35	17	52
Others * -- Total	485	13	498
Florida	462		462
Other States	23	13	36
Total	995	123	1118

* Organizations, Individuals, and Federal Agencies.

Table 2. -- Summary by Option

	Congress	State & local	Federal agencies	Organizations	Individuals	Total
PRO:						
Letters:						
Option 1		43	1	5	2	51
Option 2					1	1
Option 3						
Option 4		1				1
General	158	308	4	33	439	942
Total pro	158	352	5	38	442	995
CON:						
Letters:						
Option 5	76	34		1	12	123
Total Opinions	234	386	5	39	454	1118
General/Other	52	21	2	3	7	85

[FR Doc. 92-31890 Filed 12-29-92; 11:37 am]

BILLING CODE 3510-07-M

Mr. MILLER. Director Prewitt, if you plan to release adjusted numbers at block level, please be prepared to defend the accuracy at that level. Don't tell me that because some agencies have elected to use adjusted numbers at the national level we should adjust the population for all 6 million blocks across the country.

A part of your statistical design hinges on using race, age and other characteristics to characterize people on what you call post-strata. For example, one category would be Asian women living in small metropolitan areas, age 30 to 49, renting their living space. And while the exact number has changed several times, I believe the current design is made up of 448 such categories; is that correct?

Mr. PREWITT. Yes, sir.

Mr. MILLER. In these post-strata or profiles, when you say Asian, do you make distinctions between Japanese, Chinese, Laotians, Koreans or other Asian cultures?

Mr. PREWITT. No, sir.

Mr. MILLER. Do you make a distinction between Japanese, Chinese, Laotian, Koreans or others, or is it all just Asian?

Mr. PREWITT. It's all Asian.

Mr. MILLER. On Hispanics, do you lump confidential Cuban Americans, Puerto Rican Americans and Mexican-Americans?

Mr. PREWITT. Yes, sir.

Mr. MILLER. And they are all Hispanics?

Mr. PREWITT. They're all Hispanic.

Mr. MILLER. And so the assumption is they all respond in a similar manner?

Mr. PREWITT. Correct. That is the assumption, yes, they have similar capture probabilities, is how we would put it.

Mr. MILLER. Well—so, the Cubans in Miami, the Puerto Ricans in New York, the Hispanics in Houston or Los Angeles all have the same characteristic response rates? Guatemalans, Hondurans in Miami, they all have the same as the Cubans? I find that hard to understand and grasp.

I think if you talked to the Cubans in Miami or the Mexicans in Los Angeles, they may not totally agree with that, that they are all homogeneous as you assume.

Mr. PREWITT. No, I said they all have the same capture probability, which is to say the chances of enumerating them in the basic census are roughly similar. Now, let us make certain that we understand that we're not just talking about Cubans or Hispanics, we are also talking about 17 or 16 other sets of characteristics. Do they own? Do they rent? What is their age? What is their gender? So it is not simply the ethnic or racial characteristics. It is a cluster of characteristics that create a post-stratum.

Mr. MILLER. As I understand, the post-strata design adjustments for various categories will be applied the same way across the entire Nation. For example, Hispanic men age 18 to 29 in large cities in rental housing in areas with high mail return rates. As I read that, it sounds like as long as they fit that description, Cubans living in Tampa will get the same adjustment as Mexicans living in Houston and Puerto Ricans living in New York and so on across the country. That is correct?

Mr. PREWITT. Yeah, it turns out they have very similar capture probabilities. That is why it is correct.

Mr. MILLER. Let me ask you about Puerto Rico. You have 84 post-strata classifications for Puerto Rico; is that right? For Puerto Rico, you have 84?

Mr. PREWITT. Checking.

Mr. MILLER. OK. OK. My understanding is there was 84. Does Puerto Rico have the separate strata but Texas, Chicago, California and New York are treated separately? How is that? Why don't we have separate classifications of post-strata for Texas, California, Chicago or New York City, but we have all of these strata classifications for Puerto Rico? Why is Puerto Rico singled out? I mean, they are all Hispanic or most all Hispanic. But then——

Mr. PREWITT. Before we turn to Puerto Rico, let me address the first part of your question. Can I just go back a bit? You have asked a large number of questions to get to this very particular one. And I don't want to readdress all the access questions, but I do think that nearly every one of those things that you mentioned in your response on the access question have been answered before and continue to be answered. We talked about a terrabyte of information, 52 million yellow pages worth of information. We have met repeatedly with the GAO and the monitoring board. There are no access questions, there are no transparency questions that I know to be on the table right now, sir.

And you chide us about having guidelines. On the other hand, you were the one who asked me in a hearing to please create some guidelines so that we could all sort of understand and move forward in this census without disrupting it. So I find it a little odd that now we try to have some guidelines and we are chided for that. So it is kind of either way.

Mr. MILLER. It is interesting that you have given me guidelines on my behavior. I mean, my letter that you received the other day is that you are telling me who I can call and who can be present in a phone call. I'm the elected representative of the people.

Mr. PREWITT. Of course, you can do whatever you want and will do whatever you want. We then get pressure from the minority wanting to make exactly the same calls. I don't know, I don't know whether you want to call 100 people and put them on tape, or 10 people, or 200 people.

And we're just trying to sort of manage the process right now. We're trying to finish the census. It seems kind of reasonable to us to say, if possible, let's try to coordinate these phone calls. But no, you can call anyone you want to in the country. We will give you the phone numbers of 500,000 employees.

Mr. MILLER. By the way, it was very easy to locate the gentleman in question and he was a very pleasant call.

Mr. PREWITT. I am also very pleased that you made reference to the fact this is a rare e-mail. Of course it is rare. That is why you had to pay such attention to it because there is not a pattern of such e-mails.

But back to the ACE. You said in your opening comments that the Bureau of Labor Statistics used the data at the national level and State level and substate levels, but primarily you are absolutely correct. I would just like to make certain that you under-

stand that the Census Bureau believes that local area data are unstable irrespective of whether they are corrected or not.

We think that the apportionment number, which is based upon the basic enumeration, is unstable at the local level. We don't have a whole lot of confidence in block level data, period, however they are collected, because it is the nature of very small area data that that is where errors can get magnified.

So we are just as worried about block level data preadjusted or precorrected as we are post adjusted and post corrected. It is just a fact of the nature of statistical operations.

So when you say that we have to be absolutely correct at the local level, we would—if that were the standard, we would not be able to give this country the redistricting data based upon the initial census, because we couldn't stand behind that data at every local level. We just couldn't do it. So it is not an issue of whether it is adjusted or unadjusted——

Mr. MILLER. Which is more correct—which is more accurate at the block level, the actual count or the adjusted, or you want to call it the corrected number at the block level? Which is more accurate?

Mr. PREWITT. Oh, undoubtedly the adjusted number is more accurate across——

Mr. MILLER. At the block level?

Mr. PREWITT. Oh, yes, absolutely. At the block level we are missing—we know in certain blocks in Chicago, heavily comprised of African-Americans who rent their housing who are young males, we're missing a large number of them.

Mr. MILLER. You have already decided then that the adjustment is going forward and that is going to be more accurate, the adjusted numbers at the block level. Why did the Census Bureau not use the adjusted numbers for the intercensual estimates? Why did the BLS use it only for the national numbers or very large State numbers and not for—and we have all used the unadjusted numbers? Now you are saying it is more accurate or corrected, the political term you are using.

Mr. PREWITT. I am sorry that you are concerned about that. We use the word "adjusted" and "corrected" interchangeably.

Mr. MILLER. "Corrected" is a new use of the word, I think. You have been using "adjusted," and I think that is more appropriate.

Mr. PREWITT. Well, I have been using the word "corrected" since I got here. I am happy to use the set of terms you want. They are interchangeable as far as we are concerned.

The decision has not been made, sir. The decision that has been made is to proceed with the ACE procedure. And indeed, as I said, we are now over 20 percent finished with the collection of the data.

We have theoretical reasons for believing that this will produce more accurate numbers. We also have to test the operations. Your concern about whether this is a spaceship to Mars that may blow up is an understandable concern. That is also true of the census. It is also true of the enumeration. Any one of our operations could have blown up. We are very pleased that we are now about 85 percent finished with the census and it hasn't blown up. And we've had many, many hearings about why that is so. We are really very pleased with the operational robustness of this census.

But it is not in the nature of one operation versus another operation that it can turn into difficulties. Any operation will run into difficulties, including the new operation—not a new, but an operation that we have not talked about in this subcommittee called coverage improvement followup, which is going to have 7½ million households in it. We have not done that one yet. It may not work well. If it doesn't work well, that will have an impact upon the quality of the apportionment numbers.

So it is not something special with the ACE which makes it vulnerable. Any big complex field operation is vulnerable. The good news is we're 85 percent finished with the census without having had one. We still have 15 percent to go and it is a very, very hard 15 percent because we are now down to the difficult cases.

So we have made the decision that based upon statistical theory, capture-recapture technologies are able to improve a basic count. That's the decision that we have made and we have designed an operation to do that.

Mr. MILLER. Well, I think we are all pleased that the full enumeration is proceeding as it is. I think from the mail response rates to the nonresponse followup is proceeding apparently ahead of schedule. And that is the positive thing.

But ACE, there is legitimate differences within the statistical community, as you are well aware of, and to say that this is already going to be the corrected number, and the other one is an incorrect number by inference—when you say one is correct that means the other one is incorrect—is politicizing the process, and that is unfortunate.

Mrs. Maloney.

Mrs. MALONEY. Thank you very much. I feel that the chairman often will invoke a person's name who is not here and use it as evidence and they're not here to speak for themselves. He did it last week with Mr. Rodriguez, a Marine, the civil servant, the young man who was working in the census office. He wasn't here to speak for himself. Now we have his written letters. We know what he said. But earlier he mentioned the names of Dr. Barbara Bryant, the Census Director under former President Bush, and let's have her come here and speak for herself on how she feels about modern scientific methods.

And he mentioned the name of Dr. Janet Norwood and his conversation with Dr. Janet Norwood from the National Academy of Sciences, implying that she did not support the plan of the agency. Well, may I suggest, respectfully, that we invite Dr. Norwood to come and speak for herself. I have a May 3rd, 1999, letter from Dr. Norwood that I would like to put in the record that appears to indicate that she supports the plan of the agency, "In general, the panel concludes that the ACE design work to date is well-considered. It represents good, current practice in both sample design and post design as well as the interrelationship of the two." And I'd like to put that in the record.

So very respectfully, I suggest that we have these people come and enter their own testimony as opposed to an interpretation by the chairman.

And I would like to ask you, Dr. Prewitt, have you had any conversations with Dr. Norwood? And what is your interpretation to date on her support of the agency's plan?

Mr. PREWITT. Well, Dr. Norwood is of course the chairman of the standing committee now of the National Academy of Sciences that is looking over all of our plans for the accuracy and coverage evaluation as well as the basic census, so obviously I go to all of their meetings that are publicly open. I am frequently asked to testify or to present materials before that committee, so yes, I have attended every one of the meetings of this committee and have had conversations with Ms. Norwood in that context.

I believe that to try to get the facts exactly on the table, the letter to which you refer, which is of course over a year old, written May 3, 1999, was based upon the degree of work that had been done to that date, and there has been a lot more work done on the ACE design since that date. And that's what is represented by this stack. That is—the size of the stack about a year ago would have been, you know, a quarter to a fifth of this size because we hadn't done a lot of the technical work then. So based upon the technical work that had been done, which was the early sampling design this is the judgment that the committee wrote in that letter of May 3rd.

I think it is correct to say that the National Academy committee has not, "signed off on" the full design because they haven't met since the full design has been completed. There are now 106 major decisions that have to be made with respect to the ACE design; 104 of those have been made. The two which have not been made have to do with weight trimming methodology, and varience estimation, the specific criteria for weight trimming methodology. And under variance estimation we haven't fully finished talking through technically the specific criteria for incorporating controlled rounding into generalized variance estimation. Those are the only 2 out of 106 major decisions that have not yet been completed.

You know, we're talking about those all the time right now. In the next couple of weeks we will have resolved those. We will put them in a piece of paper and they will join this stack and they will be sent to this subcommittee if they want them and also to the National Academy.

So I think it is fair to say that they haven't, "signed off on" the final design because no one could have. The final design did not exist on May 3rd nor indeed, as of yesterday whenever the chairman talked to Ms. Norwood it did not exist. So I think the National Academy committee has been extremely useful to us, looking at our work, judging it, passing back suggestions and recommendations to us, holding public events. They had a major public event on the design earlier this year and another one is scheduled for September.

So we don't expect them to have signed off on a design that has not yet been completed. I think the chairman is quite correct. They need to see it and they will see it as soon as it is completely finished, and we are very close to having it completely finished.

Mrs. MALONEY. Dr. Prewitt, a number of people have suggested that the use of statistical methods to correct errors in a census opens the process to political manipulation. Would you please explain to us whether or not you believe that to be true, and what can be done to assure the public that no such manipulation occurs?

Mr. PREWITT. Well, needless to say, there are few charges that bother the Census Bureau as much as that one does. To my knowledge, the decision memo by Secretary Mosbacher in 1991 was the first time a senior official of the U.S. Government ever put on record the possibility that the Census Bureau could design a procedure in anticipation of it having a given partisan outcome. And what he said is that the political outcome of a choice, that is of a statistical procedure, can be known in advance. He says: "I'm confident that political considerations played no role in the Census Bureau's choice of an adjustment model for the 1990 census. I am deeply concerned, however, that adjustment could open the door to political tampering with the census in the future."

This put on record the idea that the Census Bureau could design something having a known partisan outcome. Let me just say that this strikes me as ludicrous on the face of it. The Census Bureau does not have the competence to predetermine partisan outcomes. It has no statistical expertise in reapportionment or redistricting, no expertise on trends in voting behavior. To predetermine partisan outcomes the Census Bureau would need to bring to bear such expertise when it selected data collection methodologies several years in advance of when the census counts are actually to be used for reapportionment or redistricting.

It is simply way beyond the Census Bureau's competence or capacity. Even if the Census Bureau intended to do it, it would not know how to do it. It doesn't have the competence, it does not have the interest, and it certainly does not have the professional position that that is what this job is.

I would like to point out that there are a large number of oversight agencies—this subcommittee, the Congressional Monitoring Board, the Inspector General—and there are some several dozen reporters who follow the census very closely. There are public watchdog organizations. There are National Academy committees. There are a large number of people who, if they could find partisan manipulation of this census, would be the first to report it. And I can only say we are now getting toward the end of the census and no such incidence has ever been revealed. Where is the evidence that the Census Bureau is designing things to have a partisan outcome? What kind of capacity would we have to have? We don't have it. We wouldn't know how to do it. We don't care about it. We don't pay any attention to redistricting data or voter trend data or what Governor controls what State. We don't pay any attention to that stuff, because we're actually trying to do a census and that is the kind of capacity we have.

So I just think the charge that the Census Bureau itself has a partisan agenda should be dismissed. I invite the Congressional Monitoring Board to try to determine this and find it and reveal it. I begged the cochairmen to have a hearing just on this issue, and they have not yet done it.

So I just find that concerning this charge that was put in the books 10 years ago, almost 10 years ago, there has never been any evidence put on the table. I just wish someone would put the evidence on the table so we could answer it.

Mrs. MALONEY. I'd like to ask an operational question on a non-ACE topic. You stated today that the Bureau had completed 50

percent of the nonresponse followup, and that sounds remarkably good at this point. Are you ahead of schedule, behind schedule? Could you comment on this number and exactly where are you?

Mr. PREWITT. Well, we are always cautious, of course. Choosing my words carefully, we are not displeased to be at 50 percent. On the other hand, the hard cases are yet before us. We are now running into gated communities, a much higher density of gated communities, where we are having a very difficult time getting past the doorman. And yet we know we have a low response rate from those areas. We are certainly now running into the difficult cases in the inner city and the African-American population that Congressman Davis just referred to. We are in the difficult cases in the immigrant populations.

So to say that we are 50 percent where we need to be is, indeed, good news. And as I say, when you put that together with the mailout, we are about 85 percent finished with the census. But we are now down to the hard cases and as I have said in my written testimony and I have said many, many times, at the end of the day, we will not get everyone. We would love to be proven wrong, but we have had too many people already say I will not answer this or I don't care what you say, you can put me in jail—I have given you some of that evidence before—so we will not be able to get everyone in the census.

Mr. MILLER. Thank you. Mr. Ryan, if you will let me make one comment first, and that is the question on the political manipulation of the census. I would like to insert in the record the Supreme Court decision concurring that Justice Scalia said, and he used the phrase that an estimation was more likely to be politically manipulated than the full count.

The Supreme Court is even saying that there is a difference of potential political manipulation. And I would also like to include in the record the CRS report on adjustment, because the CRS said, as I said in my statement, that the Supreme Court ruling precluded the use of sampled data for redistricting. Use of money and other things is a different issue. But for purposes of redistricting, it did not rule that it was illegal.

So with that, and with the inclusion of those, without objection I am including both of those statements, Mr. Ryan.

[The information referred to follows:]

Order Code RL2

CRS Report for Congress

Sampling for Census 2000: *Department of Commerce v. United States House of Representatives* and its Ramifications

February 3, 1999

Margaret Mikyung Lee
Legislative Attorney
American Law Division

CRS

Congressional Research Service • The Library of Congress

ABSTRACT

On January 25, 1999, the United States Supreme Court held that the Census Act prohibits sampling in the census for the apportionment of the House of Representatives, but declined to decide whether sampling would also be a violation of the census clause of the U.S. Constitution. This decision was the culmination of two lawsuits which had been brought to challenge the plans of the Census Bureau to use sampling in the 2000 Census. Opponents of sampling claimed victory and promised to focus on improving the traditional headcount through methods such as expanded outreach to undercounted groups and the use of administrative records. But proponents of sampling, including the Administration, noted that the decision did not determine the constitutionality of sampling and did not hold that sampling was prohibited for purposes other than apportionment of the House of Representatives, and indicated that they intended to seek the use of sampling techniques in population counts used for intrastate redistricting and funding allocation formulas. This report will summarize the Court's opinion, its ramifications, and reactions in the wake of the decision. For background of these law suits, see CRS Report 87-871.

Sampling for Census 2000: *Department of Commerce v. United States House of Representatives* and its Ramifications

Summary

On January 25, 1999, the United States Supreme Court held that the Census Act prohibits sampling in the census for the apportionment of the House of Representatives, but declined to decide whether sampling would also be a violation of the census clause of the U.S. Constitution. This decision was the culmination of two lawsuits which had been brought to challenge the plans of the Census Bureau to use sampling in the 2000 Census. Opponents of sampling claimed victory and promised to focus on improving the traditional headcount through methods such as expanded outreach to undercounted groups and the use of administrative records. But proponents of sampling, including the Administration, noted that the decision did not determine the constitutionality of sampling and did not hold that sampling was prohibited for purposes other than apportionment of the House of Representatives, and indicated that they intended to seek the use of sampling techniques in population counts used for intrastate redistricting and funding allocation formulas. This report will summarize the Court's opinion, its ramifications, and reactions in the wake of the decision.

Contents

Sampling for Census 2000: *Department of Commerce v. United States House of Representatives* and its Ramifications

Introduction

On January 25, 1999, the United States Supreme Court held that the Census Act prohibits sampling in the census for the apportionment of the House of Representatives, but declined to decide whether sampling would also be a violation of the census clause of the U.S. Constitution. This decision was the culmination of two lawsuits which had been brought to challenge the plans of the Census Bureau to use sampling in the 2000 census. Opponents of sampling claimed victory and promised to focus on improving the traditional headcount through methods such as expanded outreach to undercounted groups and the use of administrative records. But proponents of sampling, including the Administration, noted that the decision did not determine the constitutionality of sampling and did not hold that sampling was prohibited for purposes other than apportionment of the House of Representatives, and indicated that they intended to seek the use of sampling techniques in population counts used for intrastate redistricting and funding allocation formulas. This report will summarize the Court's opinion, its ramifications, and reactions in the wake of the decision.

The Supreme Court Decision

On January 25, 1999, the United States Supreme Court handed down its decision in *Department of Commerce v. U.S. House of Representatives*, ___ U.S. ___, No. 98-404, slip opinion (U.S. Jan. 25, 1999), together with *Clinton v. Glavin*, No. 98-564. The former lawsuit had been brought by the House of Representatives as an institution, while the latter had been brought by a group of private plaintiffs, both challenging the plans of the Census Bureau to use sampling in the 2000 census as a violation of section 195 of the Census Act[1] and as a violation of the census clause of the Constitution.[2] Both of the three-judge district court panels in the two cases had

---

[1] Codified at 13 U.S.C. § 195, this section provides that "[e]xcept for the determination of population for purposes of apportionment of Representatives in Congress among the several States, the Secretary shall if he considers it feasible, authorize the statistical method known as 'sampling' in carrying out the provisions of this title."

[2] Article I, § 2, cl. 3, as amended by the Fourteenth Amendment, provides that the "Representatives shall be apportioned among the several States according to their respective numbers, counting the whole number of persons in each state excluding Indians not taxed . . .

(continued...)

granted summary judgment to the plaintiffs and held that the plaintiffs had standing to bring the suits and that the use of sampling in the census for the apportionment of the House of Representatives violated section 195 of the Census Act, and each declined to decide the constitutional issue since the statutory interpretation was dispositive of the sampling issue.[3] In a 5-4 decision, the Supreme Court held that the private plaintiffs in *Clinton v. Glavin* had standing to bring the suit and that section 195 of the Census Act prohibited the use of sampling techniques in the census for the apportionment of the House of Representatives among the States, thus affirming the decision of the lower courts. The Court then dismissed the *House of Representatives* case, since it presented the same issues as those in the *Glavin* case and was controlled by the decision in that case. Although the Supreme Court also declined to decide the constitutional issue, Justice Scalia discussed the constitutional issue in dicta in his concurrence, joined by Chief Justice Rehnquist and Justices Kennedy and Thomas, and Justice Stevens also discussed the constitutional issue in his dissent, joined by Justices Souter, Ginsburg, and Breyer.

Justice O'Connor's Opinion for the Court

In her opinion for the Court, Justice O'Connor found standing for the *Glavin* plaintiffs. She noted that to establish standing, "'[a] plaintiff must allege personal injury fairly traceable to the defendant's allegedly unlawful conduct and likely to be redressed by the requested relief [cites omitted].'" Slip opinion at 10. Also, "a plaintiff must establish that there exists no genuine issue of material fact as to justiciability or the merits" in order to prevail on a motion for summary judgment. Slip opinion at 11. Although the lower court did not consider whether there was a genuine issue of material fact, Justice O'Connor affirmed the finding of standing because of ample evidence in the record supporting the plaintiffs' position. She found standing on two grounds. First, the Indiana plaintiff had shown an injury, vote dilution which would result from the virtually certain loss by Indiana of a congressional seat if sampling were used in the census for apportionment, and this injury was concrete and imminent, not conjectural or hypothetical. This injury was traceable to the use of sampling, and the requested relief, the injunction against the use of sampling, would redress the alleged injury. The second ground for standing was the "expected effects of the use of sampling in the 2000 census on intrastate redistricting." Slip opinion at 15. Some of the private plaintiffs would suffer vote dilution in state and local elections from the use of sampling. Many state laws required the use of federal decennial census data for state legislative redistricting. States also use this data for federal congressional redistricting "'because the census count represents the 'best population data available,' . . . [and] it is the only basis for good-faith attempts to achieve population equality.'" Slip opinion at 16, citing to *Karcher v. Daggett*, 462 U.S. 725, 738 (1983), citing in turn to *Kirkpatrick v. Preisler*, 394 U.S. 526, 528 (1969).

---

[2](...continued)
. . The actual Enumeration shall be made within three Years after the first Meeting of the Congress of the United States, and within every subsequent Term of ten Years, in such Manner as they shall by Law direct."

[3] The background of these lawsuits is covered more fully in CRS Report 97-871, *Sampling for Census 2000: A Legal Overview*, by Margaret Mikyung Lee.

On the merits, Justice O'Connor, writing for the Court, found that section 195 of the Census Act prohibited the use of sampling in the census for apportionment. She examined the historical context, noting that until relatively recently, the census statutes had always required personal visits to every household counted. This was changed in 1964 to enable the mail-out census methodology used for the first time in the 1970 census. Justice O'Connor viewed section 195 as a restriction on the general authority to use sampling found in section 141(a) of the Census Act. She observed that until the Administration decided in 1994 to use sampling methodology in the 2000 census, it had previously always taken the position that sampling was statutorily and constitutionally prohibited. Finally, Justice O'Connor found that although the 1976 amendments to section 195 did not alter its originally enacted prohibition on the use of sampling, those amendments served to change the authority to use sampling in all other aspects of the census from permissive to mandatory. She also rejected Justice Breyer's interpretation of section 195, which would have permitted the use of sampling as a supplement to, but not a substitute for, the traditional enumeration methods. The Court declined to reach the constitutional issues and dismissed the *House of Representatives* case. Justice Scalia dissented from part III-B of Justice O'Connor's opinion, which referred to the silence in the legislative history for the 1976 amendments to section 195 as a basis for the Court's finding that the prohibition on sampling remained intact, the opinion reasoning that such a radical departure from traditional enumeration methods would not have been accompanied by silence.

Concurring and Dissenting Opinions

Justice Scalia began his opinion with his dissent from the Court's reliance on the silence in the legislative history of section 195 with regard to any authorized use of sampling for the apportionment census. After painstakingly refuting certain points of Justice Stevens' interpretation of section 195 as permitting sampling in the apportionment census, Justice Scalia stated that "I think it must be acknowledged that the statutory intent to permit use of sampling for apportionment purposes is ***at least*** not clear. In these circumstances, it is our practice to construe the text in such fashion as to avoid serious constitutional doubt [cites omitted]. It is in my view unquestionably doubtful whether the constitutional requirement of an 'actual Enumeration,' Art. I, § 2, cl. 3, is satisfied by statistical sampling." Concurring opinion of Scalia, J., at 3. Justice Scalia continued by discussing the definitions of "enumerate" and "enumeration" that were roughly contemporaneous with the drafting of the Constitution. He noted that the early Congresses required personal visits in the early Census Acts, although estimation techniques were not unknown at that time, indicating that the Framers of the Constitution understood "enumeration" to exclude estimation techniques. Justice Scalia agreed with Justice Stevens' observation that the constitutional goal of equal representation is best served by a census taken in the most complete and accurate manner, but then added that this observation is conditional upon every estimation methodology being more accurate than the headcount and upon Congress' being reliable in permitting only more accurate methods. Since Justice Scalia believes the choice of estimation or sampling techniques is subject to political manipulation, he believes that the headcount, although not necessarily the most accurate census method available, is the most accurate method with minimal possibility of partisan manipulation. In light of the strong argument that can be made that "an apportionment census conducted with the use of 'sampling techniques' is not the 'actual Enumeration' that the Constitution

requires," Justice Scalia would not find that the Census Act permits sampling for apportionment of the House of Representatives.

The dissenters from the Court's opinion expressed themselves in three opinions authored respectively by Justices Breyer, Stevens, and Ginsburg. As discussed above, Justice Breyer would have found that section 195 of the Census Act did not bar the use of sampling in the apportionment census because the reasonable interpretation of section 195, in light of the historical context, is that sampling to supplement the apportionment census is permissible, and only the substitution of a sample for the headcount is prohibited. To Justice Breyer, the relevant legislative history showed that section 195 was meant to focus on and enable the use of sampling as a replacement for enumeration with regard to many types of demographic information collected by the census. Therefore, the exception of the apportionment census was also focusing on prohibiting substitution, not supplementation. Justice Breyer also noted the use of imputation techniques and of corrections in past censuses to refute the argument that only enumeration is permitted. Although apparently finding Integrated Coverage Measurement (ICM) of the undercount more easily justified under his interpretation, Justice Breyer also would have upheld, as a permissible supplement, the Census Bureau's use of sampling for non-response follow-up, which would have been used to determine the final 10 percent of the population.

Justice Stevens would have found that section 195 of the Census Act does not prohibit sampling, but for reasons different from those of Justice Breyer. He would have found that section 141(a) is a general authority to use sampling and that the 1976 amendments had the effect of transforming section 195 into a limited mandate to use sampling where feasible for non-apportionment aspects of the census. Therefore, sampling is neither required not prohibited for the apportionment census; it is left to the discretion of the Secretary of Commerce. Justice Stevens was joined by Justices Souter and Ginsburg in this view. He would have found that sampling is constitutional, because the "actual Enumeration" language did not limit the authority of Congress to direct the manner in which the census should be taken, but simply required an actual population count rather than projections or speculation or bare estimates. He agreed with the view that the Constitutional Convention contained no substantive discussion of census methodology. Finally, the standard for census methodology is whether the constitutional goal of equal representation is served and Justice Stevens believed that the goal would be served best by a manner which is most likely to be complete and accurate, noting that past innovations in the census have improved accuracy. Justice Stevens, joined only by Justice Breyer, would have also upheld standing for the House of Representatives as an institution in the dismissed case, and would have reversed the lower court in that case on the merits.

Justice Ginsburg, joined by Justice Souter, agreed with the Court that the Indiana plaintiff in *Glavin* had standing on the grounds that Indiana would lose a Representative in Congress under the proposed sampling plan, and also agreed with the dismissal of the *House of Representatives* case. However, she would not have decided whether the other plaintiffs in the *Glavin* case had standing on the basis of the expected effects of the sampling plan on intrastate redistricting.

Ramifications and Reactions

Sampling in Intrastate Redistricting

Almost immediately after the Supreme Court issued its decision, the opponents of sampling were claiming victory, but at the same time, the supporters of sampling were downplaying the impact of the decision, by emphasizing the narrowness of the holding. The Court held that the census statute prohibited the use of sampling for the apportionment of the House of Representatives, but declined to reach the constitutional question. The Court had even stated that section 195 *required* the use of sampling for purposes other than apportionment. Slip opinion at 23. The proponents of sampling viewed this as supporting the position that sampling techniques were not only permissible, but were required, in the taking of the census for the purposes of intrastate redistricting and federal funding allocations.[4] However, a closer examination of other parts of the Court's opinion indicates that it did not interpret those other purposes as necessarily including, at least, intrastate redistricting. It refers to these other purposes, noting that the census serves as the "linchpin of the federal statistical system by collecting data on the characteristics of individuals, households, and housing units throughout the country [cites omitted]." Slip opinion at 24.

As discussed above, Justice O'Connor based her standing analysis, at least in part, on the "expected effects of the use of sampling in the 2000 census on intrastate redistricting." Slip opinion at 14. Her discussion of these expected effects appears to indicate that the Court assumed that the federal decennial census figures for apportionment would be the figures used by the States for congressional redistricting and, in many cases, for state legislative redistricting. The Court seems to think that the references to the federal decennial census data in state legislative redistricting statutes and state constitutional provisions are references to the data for apportionment of the House of Representatives. Otherwise, the threatened injury to the plaintiffs would not be redressed by the Court's decision. Certainly, the position of sampling proponents, if officially adopted and carried out, would mean that the threatened injury to voters in state and local elections had not been eliminated by the Court's decision. The issue of redressability and the possibility of a two-number census was raised during oral argument.[5] However, the analysis in this part of the Court's decision deals with standing and not with the merits, therefore, technically, the position of sampling proponents, that sampling in intrastate redistricting is required, is not inconsistent with the Court's holdings on the merits, but is arguably inconsistent with the apparent assumptions and larger scheme underlying the holdings. Another indication of the implications of the Court's standing analysis is the fact that

---

[4] Since the required taking of a traditional headcount for apportionment of the House of Representatives would make the non-response follow-up sampling moot, presumably any contemplated sampling for intrastate redistricting and funding allocation data would be similar in concept to the ICM for the undercount or the Post Enumeration Survey conducted after the 1990 Census.

[5] Oral Argument Transcript, found at 1998 WL 827383 on Westlaw (oral argument of Michael A. Carvin on behalf of the appellees in No. 98-564).

Justice Ginsburg would have declined to base standing in *Glavin* on the expected effect of the sampling plan on intrastate redistricting.

Other case law arguably supports the use of figures other than the official data for apportionment. The Federal Constitution does not require the use of federal decennial census data for intrastate redistricting and federal courts have held that States are not required to use census data, adjusted or unadjusted, for redistricting. In the 1969 Supreme Court decision in *Kirkpatrick v. Preisler*[6] involving Missouri's congressional redistricting plan, the Supreme Court, while invalidating the plan, nevertheless indicated that the use of projected population figures was not *per se* unconstitutional and that States may properly consider such statistical data *if* such data would have a high degree of accuracy (however, the Court also stated that the federal decennial census data were the best data available).[7] In *Senate of the State of California v. Mosbacher*,[8] in which the state senate was suing for the release of adjusted data after the Bureau decided not to adjust the official 1990 census data, the court, citing *Tucker*, noted that if a State knows that census data is underrepresentative of the population, it can and should utilize non-census data, in addition to the official count, for redistricting, but the court also held that the Secretary of Commerce had no affirmative duty to assist the State by providing adjusted census data.[9] Although Congress has not explicitly required States to use federal decennial census data in congressional redistricting, it could arguably do so under the same constitutional powers which give Congress the authority to establish other redistricting guidelines if it chooses, Art. I, § 2, cl. 1, which provides that the Members of the House of Representatives shall be chosen by the People and Art. I, § 4, cl. 1, giving Congress the authority to determine the times, places and manner of holding elections for Members of Congress.

Since, under the Federal Constitution, the States arguably can and should use data other than the official apportionment census data in their own redistricting process if they know the other data to be the best available data, one must look at each State's laws to determine whether the States themselves require the use of official federal decennial census data in the redistricting processes. Although most States prescribe a redistricting procedure by statute for state legislative redistricting, many do not have a statutory procedure for congressional redistricting. The state legislatures in such States simply conduct the congressional redistricting as they decide on an *ad hoc* basis after a federal decennial census. This means that often in such States there is no explicit statutory requirement to use official federal decennial census data for congressional redistricting, although there may be such an explicit requirement for state legislative redistricting. To the extent that a State's own laws

---

[6] 394 U.S. 526 (1969).

[7] *See also Dixon v. Hassler*, 412 F.Supp. 1036, 1040-41 (W.D. Tenn 1976), *aff'd sub nom. Republican Party of Shelby County v. Dixon*, 429 U.S. 934 (1976); *Exon v. Tiemann*, 279 F. Supp. 601, 608 (D. Neb. 1967).

[8] 968 F.2d 974 (9th Cir. 1992).

[9] 968 F.2d at 979 (but Judge Pregerson, dissenting, argued that by refusing to disclose the adjusted data, the Secretary may have impermissibly interfered with the state senate's duty to redistrict under the Federal Constitution and the Voting Rights Act).

do not explicitly require the use of official federal decennial census data for intrastate redistricting, the State is free to use any other data. Even if a State's laws require the use of official federal decennial census data, it is unclear what a reference to official federal decennial census data would mean, if the Federal Government released two official sets of data. This issue was also considered during the oral arguments in the census sampling cases.[10] If the Secretary of Commerce transmits an official, second, adjusted data set, that data arguably could still be considered official federal decennial census data, even if it is not the data used for apportionment of the House of Representatives. Again, one should note that the Court's holding on standing for the *Glavin* plaintiffs indicates that a majority of the Court considers the references to official federal decennial census data to be a reference to the apportionment data.

In any case, statements from the Administration in the wake of the decision focused on the narrow, statutory nature of the ruling and possibility of using sampling in the census for other purposes including intrastate redistricting and funding allocations,[11] leading to reports that there would be a two-number census.[12] Far from settling the census sampling issue, the Supreme Court's decision appears to have merely triggered a new round of maneuvering, with the battles looming once again during the budget and appropriations process,[13] and also with new battles emerging in the state governments over intrastate redistricting.[14]

The Administration's Budget Proposal

Because there was not enough time to revise the proposed budget for the 2000 decennial census operations, the Administration's budget proposal for fiscal year 2000, issued February 1, 1999, included a request based on the use of sampling methodology.[15] This request was for $2.79 billion, not including the $4 million requested for the activities of the Census Monitoring Board. The Census Bureau is working on a revised plan which takes into account the requirements of the Supreme

---

[10] Oral Argument Transcript, found at 1998 WL 827383 on Westlaw (oral argument of Michael A. Carvin on behalf of the appellees in No. 98-564).

[11] *National Journal's Congress Daily* for January 25, 1999, at http://www.cloakroom.com/pubs/congressdaily/dj990125.htm; Press Statement by Commerce Secretary William M. Daley on the Supreme Court's Decision on the Census, for January 25, 1999, at http://www.doc.gov/Opa/Speeches/census125.htm; Press Briefing by Joe Lockhart on January 25, 1999, at http:/library.whitehouse.gov/ThisWeek.cgi?/type=b&date=2&briefing=5.

[12] Barbara Vobejda, *Census Plans To Release Two Sets of Numbers*, Washington Post, January 27, 1999, at A2.

[13] Press Briefing by Joe Lockhart on January 25, 1999, at http:/library.whitehouse.gov/ThisWeek.cgi?/type=b&date=2&briefing=5.

[14] James Dao, *Census Ruling Reignites a Partisan Battle*, N.Y. Times, January 27, 1999, at A17, col. 1; James Dao, *Split Decision Sets Stage for State and Local Battles*, N.Y. Times, January 26, 1999, at A20, col. 2; B. Drummond Ayres, Jr. *Small Strides for Democrats Could Be Big After Census*, N.Y. Times, November 5, 1998.

[15] Office of Management and Budget, *Budget of the United States Government for Fiscal Year 2000* , Appendix, 198-199 (February 1, 1999).

Court's decision and which will certainly propose a much higher budget, given the extensive personal outreach and contact effort which will be necessary for traditional enumeration methods.[16]

Legislative Activity

There will likely be a great deal of activity with regard to the release of funding for the 2000 decennial census programs for the second half of fiscal year 1999 and for fiscal year 2000, since the Administration seems to be seriously considering the release of two sets of numbers or some type of determination or correction of the undercount, while congressional opponents of sampling intend to expend more funds to improve a headcount rather than funding sampling programs. One possible scenario is that the opponents of sampling could agree to funding a sample adequate to conduct a post-enumeration quality check, since such a check has been conducted for at least the past several censuses, but inadequate to adjust the figures for each State. That is, the minimum sample size required for an effective evaluation of the national totals might be funded, but the larger sample required to evaluate the accuracy of the count on a State-by-State basis might not be funded. In a more extreme scenario in which the expenditure of funds on sampling programs would be prohibited, it is conceivable that States or municipalities could provide funds toward a recount or adjustment, since the Census Act authorizes the Census Bureau to conduct surveys for state and local governments which pay for them.

Some proposals for improving the traditional enumeration in a cost-effective manner include the use of administrative records and Americorps workers in conducting the census. The latter is part of the America Counts Today initiative proposed by Congressman Dan Miller, the Chairman of the House Census Subcommittee to improve the accuracy and coverage of the traditional enumeration methods by (1) creating community awareness; (2) increasing the involvement of community leaders; (3) reinforcing community-based enumeration; and (4) strengthening the Census Bureau's commitment to community-based enumeration.[17]

There have been some proposals to amend section 195 of the Census Act, but these would likely be controversial and timely passage would be doubtful. The only legislation introduced in the 106th Congress with regard to the census sampling issue as of the date of this report is S. 166, introduced on January 19, 1999, by Senator Moynihan. This bill would require the Secretary of Commerce to determine any surpluses or shortfalls in certain grant amounts made available to the States by reason of an undercount in the most recent decennial census conducted by the Census Bureau.

---

[16] *Remarks by Deputy Secretary of Commerce Robert Mallett, Budget Press Conference*, February 1, 1999, at http://204.193.243.2/public.nsf/docs/fy2000-unveiling-speech.

[17] *Statement of Chairman Dan Miller of the Subcommittee on the Census at the United States Conference of Mayors*, January 27, 1999, at http://www.house.gov/danmiller/census/jan27b_99.html.

Mr. RYAN. Dr. Prewitt, let me start with a couple of questions and then I would like to—I guess we are getting down to this issue, the politicizing this thing. Have you made the final decision to adjust the census numbers according to the result of the ACE? You have made that final decision; correct?

Mr. PREWITT. No, sir, we will not make that final decision until February, March 2001.

Mr. RYAN. In your written testimony, you indicate that you would not release the adjusted numbers if they did not meet the Bureau's standards of accuracy, and you said a review of the ACE, quote, should take into consideration a review of selected measures of quality.

Specifically what are our measures of quality and when were they established?

Mr. PREWITT. Yes, sir, that is indeed the topic of our next meeting with the National Academy of Sciences. We're working these through right now. We are presenting these publically well before we make the decision in September. Can I give you one example?

Mr. RYAN. Sure.

Mr. PREWITT. Let us say we get the results back from the accuracy and coverage evaluation and we have a higher than expected undercount in inner cities of the White population which owns its own home. That is something we don't expect. Well, it may well be since 1990 to 2000 there has been a lot of gentrification and a lot of this gentrification is now in gated communities and these people are not returning their questionnaires and we get an unexpected undercount in a population group where we did not expect it.

When we get that pattern, what we will do is say can we explain it? If it is a pattern we can't explain, it will make us nervous and we will have to figure it out. If we can explain it because there are now more gated communities in inner cities that happen to be owned and inhabited by Whites who normally give us answers to this, we will say now we have an explanation for something that otherwise looks to be anomalous.

So that is what I would mean by looking at our own results before we make the final decision.

Mr. RYAN. OK. And at your request, the National Academy of Sciences has a special panel that has been convened, headed by Janet Norwood, which we have been discussing today, to review this. How important do you believe is the task of this panel? And do you know what their time line is? What is going to be the time line of the panel for evaluating the statistical methods?

Mr. PREWITT. Well, we think this panel is very important, and our work with the National Academy of Sciences has been very important over the entire decade. However, the decision itself about what the Census Bureau is obligated to do to fulfill its constitutional and other statutory obligations is clearly a decision of the Census Bureau and not an independent committee of the National Academy.

Mr. RYAN. Let me ask you this: Will you wait for the evaluation of the panel before you release your adjusted numbers?

Mr. PREWITT. Well, of course not. We have to release our adjusted numbers according to our statutory deadlines.

Mr. RYAN. Let's get to the political part of this.

Mr. PREWITT. I don't know when they are going to do their evaluation. They are independent of us.

Mr. RYAN. Let me get to the political part of this, and I understand your comments where you say this is ridiculous that the census could be politicized. Well, I don't see you as a political person. I don't see those who work with you at the Census Bureau as political people. I see you as doing a job and you have done a good job of enumeration and I would like to give you credit for that. You are working at a statistical adjustment. You are doing what you have been trained to do.

But your boss is the President of the United States. Your boss below him is the Secretary of Commerce. Very political people, the head of another political party. So you can understand why you would see these kinds of allegations. I don't think people are saying Ken Prewitt is a politician who is seeking political ends with the statistical adjustment. But you can see it is very rational to take a look at the situation and who you work for and then make those conclusions.

The concern that I think many people have is the compressed timetable. In 1991, the Bureau discovered a computer error in the PES system that threw the undercount off by a million people. Then during a series of evaluations that took about 2 years the Bureau discovered more errors in the system that added millions of people erroneously. Now so far in the census research has shown that the Bureau has had two computer errors. One printed 120 million wrong addresses, the other failed to print millions of surnames. These things happen. But given the 1990 experience, and given that small computer errors produce millions of problems very easily in the adjustment, there is cause for concern.

So can you understand that people in Congress and in the scientific community are alarmed at the prospect of making adjusted numbers official after less than 4 months of evaluation? That is the cause for concern. And the other question I have in that is are you trying to have the official numbers done by January 20th? Is that a deadline that you are trying to shoot for?

Mr. PREWITT. You mean the redistricting numbers? The apportionment numbers?

Mr. RYAN. Yes.

Mr. PREWITT. Absolutely not, sir. There is no way we could——

Mr. RYAN. The official adjusted numbers, not just the redistricting numbers. And—well, I will let you answer.

Mr. PREWITT. I think I know where you are going. There is of course the apportionment number, which is December 31, and that will be finished on schedule. There is then the redistricting number, which is April 1st. The current plan for redistricting numbers is that they will be adjusted numbers or corrected numbers. Under no circumstances would our schedule allow us to produce that data tape prior to January 20th.

Mr. RYAN. I understand. I understand that the official adjustment leads to the redistricting numbers. But don't you think it is a reasonable concern that given the problems that can occur with an adjustment, that a 4-month timetable is relatively rapid?

Mr. PREWITT. Oh, yes, sir. In fact, I would say, Congressman Ryan, that trying to get the basic census done in 9 months puts

a lot of pressure on us. And, indeed, a coding error can occur in the enumeration process as well as in the correction process. It can just occur, and the ones that you have cited occurred in the basic census. And indeed it is quite possible that we will find out 2 years from now that we made some error. We don't expect to find that in the enumeration, but if so, we would have already reapportioned and we would have to say "too bad," we made an error and there it is.

It is not something unique to the ACE process, it is something that is a characteristic of the entire process.

We obviously learned a lot based on 1990 and 1980, where we did these exercises, and we have put in place—and this is what this documentation is all about—we have put in place with respect to our software development work enormous layers of redundancy. We are double-coding every piece of software in the ACE estimation process. And then we have compared the results of two completely separate writings of the software code. And we have built in quality assurance processes.

So we know it is a tight time schedule, but so is the census a tight time schedule. Everything in this process is a tight time schedule. We are pleased that the errors that have been discovered so far did not have operational consequences and they were, out of 2,500 different pieces of software, one or two.

Mr. RYAN. The last thing I would want to do is force you to do sloppy work by making you compress into an artificially chosen timetable. Let me go back to the fact that the task of the National Academy of Sciences, at your request, is to evaluate the quality and accuracy of the ACE. Why will you not wait for their review of your data before releasing your official adjusted numbers?

Mr. PREWITT. Because we have a statutory deadline that says we must release numbers by April 1st. The National Academy of Science will take a couple of years. We always ask the National Academy of Science to evaluate our work.

Mr. RYAN. What good is their analysis if you are not going to wait for it?

Mr. PREWITT. Just the way the 1990 analysis helped us plan 2000, their 2000 analysis will help us for 2010. That is just the nature of the system. We cannot delegate—we cannot delegate, as the Census Bureau, the decision about what numbers to give this country to an independent agency.

Mr. RYAN. I'm not saying you're delegating the decision to an independent agency as to what numbers you give. But if you're asking the scientific community to review your data, to review the accuracy of your data before making them official, you ought to wait for them to review your data before making them official. That is where I think the point can be adequately made under reasonable terms that there could be a politicization of this process. That is the concern. If you are not going to wait for the scientific community to look at the data, to look at the accuracy, to make sure things were done correctly, and rush to get these data—these adjusted numbers out there in an official capacity, then why bother? Those questions I think are very serious questions.

One more question and I see you are going to answer it. If you were to release the numbers early how much kind of warning would Congress have when you release the adjusted numbers?

Mr. PREWITT. If we release them early? You mean prior to April 1st?

Mr. RYAN. Yes.

Mr. PREWITT. We historically have released adjusted numbers on a flow basis. That is as soon as we finish a State we release it. And we have certain States that have faster deadlines than other States with respect to redistricting. And that is what we have informed the States, we will get to them as soon as possible. We do not expect to have any State completed before early March. But I can imagine that we will have States out as early as March 5th, some States out March 11th. We are going to be driving toward an April 1st deadline.

I'm not sure what you mean by informing Congress of this. We normally simply release redistricting data tapes on a flow basis starting as soon as we can. I am happy to tell the Congress when we have that schedule. There is nothing secret about that schedule.

Mr. RYAN. That would be appreciated.

Mr. MILLER. Mr. Davis.

Mr. DAVIS OF ILLINOIS. Thank you very much, Mr. Chairman. Mr. Chairman, let me yield for a moment to the ranking member.

Mrs. MALONEY. I thank Mr. Davis for his yielding and for his outstanding leadership on this issue.

I just wanted to respond to the series of questions that my dear friend and colleague, Congressman Ryan, was putting forth and ask you, Dr. Prewitt—Director Prewitt, can we expect from you the same independence and independent judgment and action that we saw in Dr. Barbara Bryant when she opposed former President Bush and Secretary Mosbacher and came out for modern scientific methods because she believed in them? We have a long history of independence in the Census Department and in Census Directors in speaking out for what they think is right for an accurate count for America. Can we expect the same type of independent action on your part?

Mr. PREWITT. Congresswoman Maloney, if the Census Bureau looks at the adjusted data, the corrected data in February-March and if I am then the Director and we decide that these data have some serious flaw in them, we will simply not release them. And irrespective of what the President of the United States wants, whoever that may be at that time, irrespective of what the Secretary of Commerce wants—now if they make us do it, as Mr. Mosbacher overruled Dr. Bryant, I don't know what we would do. But certainly the Census Bureau would not wish to release any data product in which it did not have confidence.

I might say, continuing on this line if I could for a moment, that I believe that in 1980 that the decision about whether to adjust or not was left to Vince Barbara, then the Director of the Census Bureau. I think that is the proper level for this decision. And in 1990, the decision was not left to the professionals at the Census Bureau; instead, it was made at the Secretary of Commerce. I would strongly urge, strongly urge that the decision in 2000 be made by the level of the Census Bureau, regardless of who may be the Secretary

of Commerce at that time. But I believe that this is not a decision that should be made at the level of the Commerce Secretary, but should be made at the level of the Census Bureau itself and its Director.

Indeed we have in place a standing committee that meets every 2 weeks that goes through all of this technical stuff, and it is designed to follow the ACE process very closely, both in terms of its statistical theory, in terms of its operations, and then make a recommendation to the Director as to whether to use it or not.

Just if I could continue for a second, Congressman Ryan, I did not fully—I do understand some of the concerns. I'm not trying to dismiss the concerns. I'm only trying to say that there is no evidence for those concerns. And even if a member of the Supreme Court says that it could happen doesn't mean it could happen. I don't know technically. If you think about it, you are sitting there trying to generate these data and you are now saying in what State is there a redistricting battle in which there is a Governor of this party and a legislature of this party and what are the processes and what would we have to do to get the data—I mean, if you actually think about it for a moment practically, how in the world would we do it?

Are we sitting there sort of looking at voter turnout in different States? Are we sitting there looking at the balance of power between the legislative and executive branches in different States? Do we even understand how different States do redistricting? If you actually look at it practically, it is inconceivable that the Census Bureau in that environment of trying to produce good data is now going to take on this extra task of finding out what are the likely political implications in a given State.

It's just not in the cards, and I don't see how people can think it is in the cards. I don't care if they are on the Supreme Court, sir. I don't know what evidence he has to make that accusation. I simply don't know what the basis of that accusation is.

Mr. DAVIS OF ILLINOIS. Reclaiming my time, you know as I was listening to my friend and colleague from Wisconsin, I was saying to myself, as he described the hierarchy relationship of the executive branch that there is no way that he would think that people with the name Daley and Clinton would be seeing this in a political way.

Mr. RYAN. Never, ever.

Mr. DAVIS OF ILLINOIS. They wouldn't by no stretch of the imagination.

Dr. Prewitt, let me try and make sure that I understand some of the technical language. It is my understanding that capture probability does not necessarily mean that everybody in a category are the same, but there are enough similarities that in terms of the probability of them being counted or enumerated becomes essentially the same.

Mr. PREWITT. That is exactly correct, sir, yes, sir.

Mr. DAVIS OF ILLINOIS. So they don't have to have all of the same characteristics but there are enough factors——

Mr. PREWITT. They are certainly not clones of each other as was suggested. We have done a lot of research for 40 or 50 years on this, and what we do say is what are the probabilities that we will

include in the census people with this set of characteristics. That's all it says. It doesn't say they are alike in all other ways; just how similar are they with respect to the probability of catching them in the census.

Mr. DAVIS OF ILLINOIS. In statistical language, is there a difference between correctedness and accurateness?

Mr. PREWITT. Yes, sir. Accuracy really has to do with the truth. And all statistical operations are estimations of the truth. That is true of the basic census. There is a true number of people who lived in the United States on April 1st. Our census is an estimation of that. We use the ACE to get closer to estimating that truth.

"Accuracy" would be if we actually found and counted every one of them. We will never be able to do that for you. We believe we will get you closer to the truth by using this process, or we wouldn't be doing it. Why else would we do it? We have lots of things to do. We only do statistical procedures because we believe they get us closer to the truth. So accuracy has to do with how close to the truth can we get.

Mr. DAVIS OF ILLINOIS. And so the closest that you could possibly get would be through the use of corrected data? Is that accurate?

Mr. PREWITT. We believe so.

Mr. DAVIS OF ILLINOIS. And so it becomes almost—I mean, we are trying to get as close as we can——

Mr. PREWITT. Yes, sir.

Mr. DAVIS OF ILLINOIS [continuing]. To making sure that every person in the country is, indeed, accounted for.

Mr. PREWITT. Yes, sir.

Mr. DAVIS OF ILLINOIS. And so without using the corrected data, we would obviously then just say to ourselves that we are going to leave those individuals out.

Mr. PREWITT. Mr. Davis, I could answer that as follows: let us say that in 2000 we were not doing an accuracy and coverage evaluation. We were simply doing the basic count and then stopping and then we came up with a number, 275,311,000 or whatever. It would be my judgment that a more accurate number to give to the country would be that number plus 1.6 percent. Which is to say, I would still rather use the 1990 estimate of the undercount even for the 2000 data if we were not doing an accuracy and coverage evaluation. I would be convinced that that number that we counted plus 1.6 percent would be a more accurate number than simply stopping with the basic count. We will do better because the accuracy and coverage evaluation that we have in place for 2000 is a much better tool to use than one from 1990, but even one from 1990 would give us a more accurate count, one that was closer to the truth than simply stopping with the basic enumeration.

Mr. DAVIS OF ILLINOIS. I know that we have talked a great deal about enumeration and there is a cutoff period. There is a time when we expect to have this done. Should we continue to experience difficulty in some areas, will that cutoff date be adhered to or is there any way to continue up to a point of satisfaction?

Mr. PREWITT. We will certainly continue, Mr. Davis. We expect across 520 offices to have completed most of our work in most of them by our cutoff date, which is July 7th. But that is not a cutoff date; that is a date in our master activity schedule. But certainly,

as was true in 1990 and all censuses, there are always some local offices where we have not fully exhausted all of our procedures and we will continue in those areas until we have exhausted all of our procedures, until we cannot think that going back yet again is likely to give a response at that household.

Mr. DAVIS OF ILLINOIS. And so one can expect that every effort or maximum effort will be made to make sure that we even reach those individuals that we are having difficulty with.

Mr. PREWITT. Yes, sir, but at a certain point, we know that we are simply wasting taxpayer dollars. And so at a certain point we are better off—I mean, how many times do you want to go back and knock on a door where nobody ever answers and the person who answers says I don't care what you say to me, I'm not going to give you that information. We could send that person back 4 times, 6 times, 27 times——

Mr. DAVIS OF ILLINOIS. In some instances it would remind me, if you just keep doing it of, you know, a young woman met a soldier and wanted to get married and she said: Soldier, soldier, would you marry me with your fife and drum, and he said no, pretty miss, I can't marry you, I don't have any shoes. So she ran and got him some shoes. Came back, same thing, would you marry me with your fife and drum? No, pretty miss, I can't marry you, I don't have a tuxedo to put on. So she ran and got the tuxedo and came back. And said, soldier, soldier, will you marry me with your fife and drum? Finally, he says no, pretty miss, I can't marry you because I've got a pretty little wife at home.

And so it seems to me that you are saying at some point, people are going to say: Get away from my door, just don't come back——

Mr. PREWITT. Yes, sir.

Mr. DAVIS OF ILLINOIS [continuing]. Anymore. I mean, those individuals who are inclined to do so.

Mr. PREWITT. Yes, sir.

Mr. DAVIS OF ILLINOIS. Let me ask you, of course there has been a lot of conversation about my city, the city of the big shoulders, the city of Chicago, in terms of difficulty that we are having. Could you elaborate on what's going on there and what we are doing?

Mr. PREWITT. Yes, sir, and Congressman Ryan will also be interested because it is not just the city of Chicago but the Chicago region. I do want to say that the Chicago region as a region is actually in fairly good shape. It is not our strongest region but it is certainly not our weakest region. Indeed, when you take into account both the mail-back response rate and the completion of the nonrespondent followup workload, the Chicago region is roughly in the middle right now and since the whole scale is high right now that means we are in good shape. Even in the very worst region we're actually in good shape.

Now with respect to the city itself, I believe there are now four local census offices where we believe that we have had to improve the strength of our local management and we have done so. In some instances we have actually changed the local manager. In another instance we brought in additional management help. Sometimes what happens, Congressman Davis, is that there is more work going on than the system records because stuff just stacks up and somebody doesn't have to process that stuff every 6 hours to

get it and so forth and so on. We are finding that out. That may not be an explanation but that may be part of the explanation in Chicago.

I certainly think that we are running in Chicago into deep resistance to cooperating with the census. And that is actually happening at both ends of the economic scale. We are running into very difficult times in the near north in gated communities. These are people who are very busy. They are, you know, worried about their stock market returns and so forth. They did not send the return in and now we're having a hard time getting past the doormen who guard these buildings and it is extremely difficult.

On the other hand, what we do is we do special things. We go to the building manager. If that doesn't work we go to the owner of the building. If not, we sometimes go to somebody influential in the city and try to get them to make that call. And at the other end of the economic scale, as you well know, the poverty people that you mentioned in your own district, those are very resistant people. They are disconnected from the society. They are indifferent to their obligations. They do not feel that the U.S. Government or the local government cares about them and why in the world should they cooperate with this?

That's why we do an accuracy and coverage evaluation. We are doing one quarter of 1 percent of the households. When you're doing nearly one quarter of 1 percent of the gated communities with your very best people, you have a higher probability of getting in than when you are trying to do the entire universe of gated communities. It is the same thing with the young African-American male in the Robert Taylor home. It is very difficult to get them all.

On the other hand, when you're doing one quarter of 1 percent of them with your very best enumerators the probabilities have just gone way up that you will get them. And that is what we do in order to calculate the undercount.

Mr. DAVIS OF ILLINOIS. After this is all over, does the Bureau have sociologists and researchers and people who will try and study the situation and make some determinations relative to this deep resistance that you spoke of?

Mr. PREWITT. Yes, sir. Yes, we do. We did—after 1990 we had anthropologists and sociologists trying to help us understand these population groups.

Mr. DAVIS OF ILLINOIS. Finally, I think it would certainly be good, and I understand that you are trying to make a trip out to Chicago to give whatever additional assurances to the elected officials and the citizens there that every effort is, in fact, being made to overcome the deep resistance that we might be experiencing, and I certainly look forward to that happening.

Mr. PREWITT. Thank you, sir.

Mr. DAVIS OF ILLINOIS. Thank you, Mr. Chairman. I yield back.

Mr. MILLER. Mr. Director, let me followup one more point on what I was asking earlier about the broad classifications we are using. And I would like to enter into the record a letter I received from Dr. Friedman, who is head of the Statistics Department at University of California at Berkley. He said, "It is assumed that all Non-Hispanic Asians age 0 to 17 living in rental units are equally

likely to be undercounted from the suburbs of Honolulu to Chinatown in New York. This assumption is plainly false."

They have done studies to show that there are huge variations within post-strata across States and so there is a real concern about that. You are well aware of that concern.

Let me now switch to the issue again of transparency with respect to the ACE. You indicated your intention to make the census fully transparent and free from charges of political manipulation. Will you commit to releasing the E sample and the P sample files from the ACE for analysis by the academic and scientific community as soon as they are available to the Bureau? They are not confidential files and for 1990 they were not made available until 1998.

Mr. PREWITT. Yes, I don't know the 1990 to 1998 process, but certainly they will be made available, yes, sir.

Mr. MILLER. There is a real concern about the decision process of which is going to be the more accurate set of data. And if the National Academy of Sciences is not going to be able to make a decision prior to March 2001, who is going to make that decision?

Now, my understanding in 1991, when this decision was made, that there was a panel within the Bureau of experts that was basically equally divided to help, we may need more clarification, but there was some panel of experts within the Bureau. But you are not going to rely on the National Academy of Sciences because they are going to take too long, I gather.

Mr. PREWITT. Yes, sir, we could not wait. The National Academy of Sciences will do an evaluation—there is already a 2010——

Mr. MILLER. So how is the decision going to be made? Is it going to be made—of course you won't be there, unless whoever is appointed President. We know there is going to be a new President. But Mr. Thompson, Mr. Hogan will certainly still be there. What experts? Are they strictly Bureau employees or who is going to come up with the recommendations? I think there was some outside people making recommendations, acknowledged experts.

Mr. PREWITT. What will occur—let's just talk about 2000. What will occur in 2000 and is occurring in 2000, we do have an executive committee that follows the ACE process. As I say it meets every 2 or 3 weeks, many members of which are here behind me. I think there are maybe 9—no, it is larger than that, maybe 13 members of that who represent all Census Bureau employees who are math statisticians, demographers, field operations experts and so forth. And they look at every one of these processes, every one of these processes, and make a judgment and deliberation about what will make the most successful census.

They will continue to meet right through the entire process. The way they are designed, that committee is chaired by John Thompson and it is advisory to the Director. It will make a recommendation to the Director, is the process.

Mr. MILLER. One of the concerns I've had going back a couple of years or so is you can have a bias within a committee. If I select a committee or Mrs. Maloney selects a committee, if we have sole responsibility for selecting it, it will be a bias by who we select. I mean you ask Mr. Davis and Mrs. Maloney, they are going to have one set of opinions. Mr. Ryan and I will have another. My under-

standing was more of a nonpartisan—if you select all people that are already biased in favor of adjustment, you are going to have that conclusion. And I'm sure Dr. Hogan, who is a respected statistician is bias to some extent. Because he has had his heart and soul in this for a decade, he's been working on this program.

But Mr. Friedman—Dr. Friedman at the University of California—Berkeley, who is not going to have any input in it is a respected statistician too. So, I mean are the only people that are going to provide input just going to be people who are "yes" people?

Mr. PREWITT. No, sir, I don't know what you mean by "yes" people.

Mr. MILLER. I don't consider Dr. Hogan a yes person.

Mr. PREWITT. You wouldn't if you met with him. These are professionals.

Mr. MILLER. I want to make sure there is a diversion of opinions in the decision process.

Mr. PREWITT. If you sat and listened to some of the arguments that go into this you would appreciate there is a divergence of opinion. And it certainly includes people who in 1990 thought we should not have adjusted who are employees and very senior, important employees at the Census Bureau. This is not a committee that was sort of put together that way. It is a committee all of whom have defined positions. These are the senior positions, and so they are there by virtue of the position they hold, not the kind of assumption they have. We did not test anybody's viewpoint.

Mr. MILLER. There are no outsiders participating in this?

Mr. PREWITT. No.

Mr. MILLER. But there was in 1990 is my understanding.

Mr. PREWITT. No, that was a different process. I can describe that. That is a different process.

Mr. MILLER. I would be interested to have an explanation of how the process or the decision will be made. In your written testimony you say you will use something called dual system estimation to estimate the degree to which each of the 448 categories of post-strata is overcounted or undercounted. Then you would assign a certain weight to that category which would tell you how many people to add or subtract from that particular segment of the population; is that correct?

Mr. PREWITT. Is that correct? I don't think that is exactly my wording. I don't think I talked about—I would have to look it up, but I think it talks about statistical records, not people.

Mr. MILLER. Let me proceed. If it is possible to have strata with adjustment factors of more than one, you have an adjustment factor of—it may be 1.1, 1.2, it is also possible to have post-strata with adjustment factors of less than 1. That is people fitting a certain description could be multiplied by a factor of 0.8 or 0.9.

Mr. PREWITT. Correct.

Mr. MILLER. Correct? OK. Let's take an example, let's say you are talking about the following: A non-Hispanic White woman age 30 to 49 living in the suburbs who own their own homes in the Midwest. Let's say the Bureau estimates a 5 percent overcount of these women. The Bureau would give this group an adjustment factor of 0.95. So if the unadjusted census counted 100 people—100

women in a block, the adjusted number would show only 95 in that block.

The truth is if the actual census counts 100 people in a block but in the ACE all of these people fall in a category that the Bureau estimates were overcounted, the adjusted population of the block will be less than 100. So we are in effect deleting people from the census.

Mr. PREWITT. No, we're not.

Mr. MILLER. I mean, if we have 100 people and their adjustment factor is 0.95, we are only going to have 95 people counted. No? I know we're not going to destroy forms. We're not talking about the forms being there, but the fact if——

Mr. PREWITT. It is very important to make the American public understand that 72 years from now, when you and I go together to the National Archives, everyone who submitted a form will find their record there. And there will be no form there from anyone who did not submit a form. That is, the actual census file itself will include everyone who cooperated in this census.

Now, we are now talking about a statistical record. And that is a different process. So it is not anything about people being subtracted or virtual people or anything else. We're talking about a statistical process. The answer, sir, is yes. Where we have evidence that a certain population group was double-counted, to leave records for those people in the statistical record, means that we have now inflated some number. We are now giving to the country something which we know to be incorrect, and we don't think we should do that.

Mr. MILLER. If two people completed the form, one in Florida and one in New York and it is the same person, we don't want that. I understand that. But the problem is my understanding is that if you have 100 people living in an apartment high-rise or something. If that is a statistical classification that is considered overcounted—you have 100 people you count, we have 100 forms that are returned. All right? And you have 100 people listed by name. But then because that fit in a classification that is considered overcounted, you are going to subtract people from that so the actual count instead of 100 would be 98, or whatever adjustment number; right?

Mr. PREWITT. Otherwise we would be giving the country incorrect data.

Mr. MILLER. Then you are deleting people from the census.

Mr. PREWITT. No, we were not deleting people.

Mr. MILLER. Wait a minute. We're keeping the forms. I understand the forms are going to be there. I don't know if they are going to be physically kept, that is a different issue. People are going to get counted less than a 1.0. You are counted as a 0.98, a 0.95. You are going to have fewer people. If you have 100 people that fill out that form, you have 100 names in that area and it comes out with a 0.98 adjustment factor because of the statistical analysis, then you are only going to have the number that is going to show up on the adjusted, or you like to call it corrected, the adjusted number is 98 people. Two fewer people.

Mr. PREWITT. Happy to call it adjusted. Again, if I could just take a moment, the people that you are describing, that is the category

of persons that you are describing, we have independent evidence that those kinds of persons were double-counted at the rate of .02 percent, to use your example. And, therefore, to leave statistical records of that category at the level which you are recommending that we do means that basically we're deliberately leaving in the census counts people who have been double-counted, because they counted their college student and we found their college student at the dormitory. That happens.

And what we know from 1990 is there were as many as 4 million cases. So, yes, we have a statistical procedure that for the purposes of giving this country accurate data for reapportionment, for redistricting, for Federal funding, we have a process that does not give the country incorrect data when we know it is incorrect. Nothing more complicated than that.

Mr. MILLER. If you have 100 names in this area, in this block, and your statistical analyses says that is an overcounted population, so even though you have 100 names of 100 separate individuals, you are going to statistically remove two, three, four, whatever the number of people of that overage is. And this is one of the problems about all of these post-strata. It is like the issue Dr. Friedman talks about, your claim that you are getting these numbers from Asians in Hawaii and Asians in Chinatown are the same. They have the same response rate. Some studies show that they don't behave the same. And, you know, I guess you have got proof that shows that the Cubans in Miami respond at the same rate of response as the Mexicans in Los Angeles or in El Paso or somewhere.

I mean you are saying they are exactly the same behavior. Based on that, you can delete people or add people, which is hard to say that—I have not been to Los Angeles——

Mr. PREWITT. That is your characterization, not ours, sir.

Mr. MILLER. But aren't you using—well, you have already said you're using—all Hispanics were one classification. Whether you are a Hispanic in El Paso or Houston or New York City or Chicago or Wisconsin, you get counted the same and you get adjusted the same if you are Hispanic.

Mr. PREWITT. I didn't quite say that. I said——

Mr. MILLER. Well, but aren't all Hispanic one category, period?

Mr. PREWITT. No.

Mr. MILLER. No?

Mr. PREWITT. All Hispanics who also are in census tracts with low response rates who also rent their houses, who also are between the ages of 18 and 29, who also are women, who also are unrelated to anyone else in that household. All of those people who have that set of characteristics constitute a universe of those people.

And then we take a sample of those persons and, on the basis of that sample, estimate for that universe of people who have all of those characteristics—not just Hispanic, but all of those characteristics—what are the probabilities that they were caught in the census. And that is the process that we used. It is not "all Hispanics" because we are—it's like saying all renters or all people between ages of 18 and 29 or all anything else.

Mr. MILLER. But the post-strata for Hispanics is all Hispanics, whether again they are in El Paso or Chicago or Miami, they are all the same. Whether it is Guatemalan, Honduran——

Mr. PREWITT. I just have to say this again. They are not all the same.

Mr. MILLER. Statistically, you're putting them in one classification.

Let me ask another question. When you take someone—subtract someone from the record, you subtract them, but it's not because of a duplicate. It's just that some statistical model says subtract one person. When you have 100 people on a block and the statistical model says subtract somebody, it's not because you have a duplicate. It's because you have 100 specific names there, but it's because of the statistical models, not because of a duplicate.

Mr. PREWITT. That's a separate process. Subtracting duplicates is a separate process.

Mr. MILLER. Go ahead and finish what you were going to say.

Mr. PREWITT. We don't treat all Hispanics like all other Hispanics. We treat Hispanics who rent, who are of a certain age, who are of a certain gender, of a certain relationship to the household. That is the post-stratum. Not all Hispanics. They all have to live in a metropolitan area. So it's simply incorrect to say that all Hispanics belong to the same post-stratum.

Mr. MILLER. But the Hispanics that meet those classifications can be living in Los Angeles, El Paso, Houston, Miami, or New York or Chicago as long as they meet those general classifications. Then they are all adjusted. Asians in Honolulu are being pooled with the ones in New York, and you're saying they respond the same.

Let me go on to Mrs. Maloney.

Mrs. MALONEY. Dr. Prewitt, would you please answer Chairman Miller's line of questioning without interruption? I would like to give you an opportunity to explain the process without interruption.

Mr. PREWITT. I think the particular process we're talking about is the structuring of the post-stratum which, as he said, there are 448. These constitute identifications of population groups, and one of the identifying characteristics is their ethnicity or their race. It's only one of their identifying characteristics. Another is whether it's a metropolitan area or not, the size of the metropolitan area. Another is, as I say, age, renter status and so forth and so on. That constitutes a post-stratum, and it is our judgment that everyone who inhabits that post-stratum has a more similar probability of being captured in the census than someone in a different post-stratum.

Everyone in the country is put into post-stratum. You are. You're put in as a White female between 18 and 29——

Mrs. MALONEY. Why is everyone laughing?

Mr. PREWITT [continuing]. Etc.

And we have—based on our experience, we have an assessment of the probability of having caught you in the census, and that's true for all of these groups. Nothing more complicated than that. That's why they're put together.

We don't yet know until we actually conduct the census how many of them we actually did catch in the census, but we think they constitute a reasonable, plausible, universe of people who have roughly similar probabilities of being captured.

I have not had the chance to read Mr. Friedman's and Mr. Walker's letter. If they are saying that all Asians from every place are put in one post-stratum, they are misreading our post-stratum design. They are very sophisticated statisticians, and I doubt they are misreading it. I doubt that, the way you have characterized their letters, the way they have written it. But I haven't read it, but, my guess, they understand our post-strata structure, and it's not putting all Asians in one post-stratum. It's not.

Mr. MILLER. It's all Asians that meet the large metropolitan areas and age brackets and such, too. But it's correct that a Japanese American in Honolulu that meets that, you know, other demographic characteristics and a China person, a person from China from New York who meets that classification, a large metropolitan area, age brackets——

Mr. PREWITT. Mr. Miller, it's just as—there are lots of ways to rent a home. You can rent a condo, you can rent a co-op, you can rent a mobile home. You can rent different kinds of homes, so all renters are also put into a post-stratum because that's one of our stratification variables. It's not just that all renters constitute a post-stratum. It's that all renters that also have these other characteristics create one.

So there's nothing magic about this process to say there are a lot of different ways in which people rent, but nevertheless we have decided that renters on balance behave differently from owners, and we have a lot of evidence to that effect.

Mrs. MALONEY. Dr. Prewitt, approximately how many post-strata of the 448 include the Hispanic characteristics, approximately?

Mr. PREWITT. Fifty-six.

Mrs. MALONEY. I think this whole issue of the undercount and the deep resistance that my colleague Danny Davis illustrated with the poem—the time that it was clarified to me in the most stark way were the statements of a Republican-appointed member of the Supreme Court, Justice Stevens, when he asked the question of a Republican lawyer—and this was before—the case that we've referred to before the Supreme Court. And he asked her, how would you count a home, an address, where six people lived, yet every time you went and knocked on that door, whether it was in the morning or at night or whatever time, no one answered the door? And she said, zero; we would count it zero. Then he asked, what if you knew and all the neighbors told you that six people lived there? She said, we would count it zero.

Then Justice Breyer asked, what if the lights go on, off and on, every night and you see the lights going off and on every night and you know people live in that home? How would you count that home? And she said, zero.

And that really clarified in the starkest and really simplest of terms why we need to adjust for the undercount when we know that people live there, when we know that people are there. We are being dishonest and unfair and unjust not to count the six people that we know live there. And on the count and the issue——

Mr. MILLER. Let Dr. Prewitt answer that.

Mrs. MALONEY. May I continue? I do not believe I've interrupted you. May I continue?

Mr. MILLER. Yes.

Mrs. MALONEY. And on the issue of the double count, many of my friends, because I am a mother, happen to be the parents of daughters; and I can't tell you—and my daughter is in college. I did not count her. She is going to be counted at her university. But I can't tell you how many of my friends who have similar children my daughter's age at school either told me that they counted their child or literally called and asked me whether or not, because they know I'm working on the census, whether or not they should count their child. So I'm giving these as just practical examples of why we need this.

Now, I have a question that—Dr. Prewitt, you mentioned that you would use the 1.6 percent if you had to, but you also said that the tool that you had for the 2000 census was a better tool than 1990. And could you explain to us why it is better?

Mr. PREWITT. Well, obviously, we've drawn on our experience from 1990. We also have a sample of approximately twice the size of what we had in 1990. That was 150,000. Turned out to be, finally, 175,000 households. In 2000, 314,000 households. We do think the construction of our post-strata is drawn upon research of over 10 years about all of our matching procedures, how we're handling movers, our software development work. There's no end of ways in which we try to improve it. That's true of every census.

1990 was better than 1980, but 2000 is much, much superior operationally, just like the census itself is superior operationally to the 1990 census thanks to the U.S. Congress, that they allowed us to front load our recruitment staff. That's why we can say we're near 85 percent complete today. A lot of the improvements that we put into the census we've also put in to our ACE design.

I just—for a moment if I could refer us all to this—I brought this chart before—because I think it's important. Each of those peach boxes represent the moments in the census when we can miss people, but it also represents the moments in the census when we can erroneously enumerate, that is, double count such as the college students. So all the ACEs is nothing more complicated than this, all it is, is to go to try to find those persons who returned the form but didn't completely mail it back. They left some people off.

It's the people who—for whom we never got an address. We think there won't be many of those, but there are some. It's the people we got in nonresponse followup, but we didn't get the complete household. And it's the people we got in yet another process called coverage improvement followup.

All of those are processes to try to get everyone. Every one of those processes can leave someone out, and all the ACE is, is a way to go back and find out the percentage of people in those various boxes when we missed them, how we missed them and what their demographic characteristics are. It's not a very complicated thing. It's a very straightforward thing. If it works operationally, we think that we should give to the country the better data, the adjusted data, the more accurate data, the more corrected data.

That's all I can really say about it. It's an attempt to find the people we missed or to find the people we erroneously included, that is, the double count and make certain they are not represented in the final statistical records.

Mrs. MALONEY. Thank you very much.

Mr. MILLER. I just want to ask you to clarify one thing. Mrs. Maloney I think knows the answer to this. But when the lawyer spoke about not counting someone if the lights come on and off, that's not the way the Bureau would handle that; is that correct?

Mr. PREWITT. In that particular instance we would try to get a proxy interview.

Mr. MILLER. Correct. You would get proxy data. And, hopefully, you could find someone that would know who lives in there, and it would not necessarily be zero. You'd try to do everything you can to get some type of data from someone else nearby.

Mr. RYAN. Like a neighbor or something like that.

Mr. MILLER. Right. So I think the attorney was not as clear on the procedures as your process would show; is that correct?

Mr. PREWITT. That is correct. There are many things about that Supreme Court ruling that were not accurate.

Mr. MILLER. Mr. Ryan.

Mr. RYAN. I have a couple of procedural questions.

It's not always the inner city. We have to focus on the rural areas, too. So I'd just like to put in a word for Orfordville, WI, if I might.

Orfordville is a town of about 700 people. Hopefully, it will be a town of 700 people after the census is done, but the interesting thing about Orfordville is they all have post office boxes. That is just the way it works there. They all use P.O. Boxes. So when they didn't get the forms they were very much alarmed.

I think we followed up with your Chicago office, and I think we're doing a very good job of getting some enumerators over there to handle that situation. But what about the other Orfordvilles throughout America, small farming towns at the intersections of rural county trunk highways?

My question is, if we didn't intervene on the Orfordville situation and there are other towns like that who have P.O. Boxes who, because we don't have post census local review, didn't catch that and these are not included in the master address file, how do we catch these mistakes? Can the adjustment help a neighborhood where no people are counted essentially?

Mr. PREWITT. Yes, sir. But before I get to that question, we do have procedures to find those areas where there was this—where we thought it was city style, but it turned out to be——

Mr. RYAN. We're working on that right now. But what if it doesn't work?

Mr. PREWITT. Absolutely, we would find this in the ACE. Here's how we would find it. Since the ACE is a random sample of all blocks in the United States—as I say, about 12,000 blocks are in the ACE design—one of those blocks or some set of those 12,000 blocks would be exactly those areas by definition and proportionate to how many such areas they are.

When we go to that area in the ACE interviewing process, we'll knock on the door. They'll say, I never got counted; I got left out.

We will then determine how would that have happened, and we will then detect exactly that problem. Indeed, it will show up because our address file won't work.

We will have independently listed—as I say, we have independently listed every address in our ACE sample block; and we are now saying, my goodness, something has gone wrong. Because we have a listing of this household, but it's not in our master address file. How could that have happened?

Then we'll determine how that happened, and when we do the adjustment we will be able to adjust for exactly those population groups who fit into this top upper right peach box. This is missed housing units. Our ACE design is as focused on making sure that we account for missed housing units as missed people in known housing units.

Mr. RYAN. I'm going to go back to this Orfordville example, because I think it's an interesting one. Not only do they all use mostly P.O. Boxes—but let's take Footville, which is the town just up the road. Footville, for some reason, your master address file, even though the LUCA tried to change this, the change was not incorporated. You included everybody who lived in Footville, WI, a town of about 600 as if they were Janesville, WI, residents. So the names were correct. The addresses, however—the street addresses were right, but the cities were—were the larger city in that county. And they all had P.O. Boxes.

So when the enumerators came around to collect the data, they knocked on the door. The people would say, I never got a form; I was never counted; I was worried you wouldn't come by; glad you're here. And the enumerator then had a Janesville address.

Now it's up to the person who answered the door to change that address, I assume, from Janesville to Footville, but what if that didn't take place? What if an enumerator didn't make it to the house that was a P.O. Box and the address for the entire small town was lumped into another city and those people weren't counted? That means in a town of, say, 600 people you missed 200 people. That's a third of the city of the town. How does the adjustment fix that?

Mr. PREWITT. I want to make certain that we give you a full answer to this. So what I'm saying may not be completely responsive.

It's my understanding that what would happen—the important thing is the addresses are all geocoded, which means whatever the kind of the denominator is, what the town is called——

Mr. RYAN. Footville.

Mr. PREWITT. Footville is called——

Mr. RYAN. Janesville.

Mr. PREWITT. Janesville, that the important thing is to make sure that when we count the person in that household they occur on the block. From the census point of view, the name of the community is not what's important. The unit of analysis for us is the block, and they will be geocoded to that block. So they will appear in the right place.

Now, the process by which we make sure that all of our blocks get attached to the right place, but now we've got to make sure to connect to the right denominator. That wouldn't be a problem of the adjustment. That would be a problem of our geographic division

working with the local community. It would be easy to fix because we know where they are.

Mr. RYAN. I see my time running out. You mentioned in my earlier questioning that you're going to be giving the States the redistricting numbers kind of on a State-by-State basis starting maybe March 5 and then moving out, but your post-strata adjustment is based on a national scale, correct?

Mr. PREWITT. Yes, sir.

Mr. RYAN. How do you take that into consideration as you're releasing State redistricting data on a State-by-State basis—when your post-strata is national, how does that jibe or correspond with—say you put Vermont's out in March and then you put Wisconsin's out in April, then California's out in later April, how does that correspond and how does that take into account the fact that the post-strata is national but you're sending out individual States earlier?

Mr. PREWITT. All of the work that will have been used to create the correction numbers on a national basis will have already been done across all 50 States, the District of Columbia, and Puerto Rico, and the actual mechanical process of actually now creating the right products takes a while, and it takes a couple of days or whatever.

I better be careful, though, have to correct me, but it takes a period of time, and we will simply turn first to those States which have earlier redistricting deadlines as best we can. I should say the entire—I'm not making a promise about March 5. I'm only saying that, in principle, it will be a flow basis to try to respond to your question about January 20.

But, basically, we could wait until the last day of the month, March 31, to mail them all out the same day. But I think, as a courtesy of the States, we would want to get them out where we can get them out sooner where possible, but all of the work that has to be done in terms of making the correction numbers from the post stata will have already been done. Otherwise, the implication of your question is correct. We couldn't do one State and so forth.

Mr. RYAN. One more question. The apportionment data—and correct me if I'm wrong, the apportionment data will be done before the ACE adjustment is completed?

Mr. PREWITT. Correct.

Mr. RYAN. How will you be taking into consideration Orfordville and Footville, WI? If the apportionment data is done before the adjustment and if those towns aren't fixed and counted for, will they not be lost in apportionment but—may be caught up in redistricting but won't they be lost in apportionment if they are not fixed with the adjustment beforehand?

Mr. PREWITT. Not if they are in Wisconsin. The apportionment number is nothing but a State total.

Mr. RYAN. You're saying because it is this block, but what if an enumerator didn't hit a door and no one got answered?

Mr. PREWITT. No. If they are not captured in the census and we have not done the ACE, then the Wisconsin number will be deficient by that amount.

Mr. RYAN. Because we have a lot of reports about rural areas who are subsisting mainly of P.O. Boxes and the enumerators just

don't catch them—you know, we're planting right now in Wisconsin. People are in the field. They are not in their homes right now.

Wouldn't a post census local review make sense? Wouldn't a 1-month post census local review—let the county clerks, let the local county board supervisors take a look at the data and say, gosh, you missed half the town of Orfordville because, during the time you were coming around with enumerators, they were out in the field planting. Wouldn't post census local review make sense for these cases?

And these cases I appeal to you are not unique. They are all over the place. Our Governor, Tommy Thompson, is saying he's getting it from the entire State of Wisconsin. Why wouldn't we want to do post census local review for those kinds of instances?

Mr. PREWITT. Obviously, a post census local review would have to be done for 39,000 jurisdictions, not just yours, which means you're asking us to redo the census starting sometime in October or November. That's impractical. If you wanted an apportionment number by December 31, we can't start redoing the census based upon 39,000 different mayors or county commissioners saying we would like you to come back and count again because we don't think everybody got included.

Mr. RYAN. What about a voluntary post census local review, like localwise?

Mr. PREWITT. This gets into a very complicated thing having to do with the nature of distributive accuracy and numeric accuracy. It's really what the court case went to and so forth. And I can get into this if we have time, but any kind of voluntary process like that that was used in some places and not other places would have all kinds—at this late stage in the census would have all kinds of implications for the final quality of the data.

I can't imagine if we made this voluntary that the only State that would be interested would be Wisconsin. I think every State in the country would say come back and count us again. We may find a few more people. That is what the ACE does.

Congressman Ryan, I'm not trying—you are making a case for the ACE. You are making a case for why we have to do this quality process to go out and determine if we miss people, where they live, and then correct for that.

Mr. RYAN. Actually, I'm making a case for post census local review for apportionment and everything else because LUCA was designed to fix this. It didn't fix it, though, in some of these towns. Some of these towns did participate in LUCA, did send their data, and they still—we still have the problems.

So that's why I'm saying, why not exhaust every effort possible? I still contend that there may be a chance, there may be a small timeline, a small window to do a voluntary post census local review so these rural towns who are having these problems can make sure they are counted. There is a lot of anxiety out there over this. I just appeal to you to take a look at that.

Mr. MILLER. We have a vote coming up, but we have time for Mr. Davis.

Actually, I'm glad, Mr. Ryan, you're on this panel. Because rural areas, as we all know, have problems of their own. And we keep focused on large metropolitan areas and the migrant population,

immigrant population, but there are unique problems in rural America, so I'm glad you can bring them up.

I agree with you, by the way. It's too bad we don't have full census local review which the House of Representatives passed but was opposed by the Census Bureau and by the Democrats.

Mr. Davis.

Mr. DAVIS OF ILLINOIS. Dr. Prewitt, are there any post-strata groups that we've conclusively determined to absolutely be the most difficult ones to enumerate or count?

Mr. PREWITT. Well, based on our 1990 experience, we would expect a group that was made up of young African American males in inner cities who rent and who live in irregular housing, thats unrelated to each other and so forth in that housing, that's likely to be a particularly hard-to-count population group.

Also, based on 1990, though we think we've done a lot of work on this, Indians living on reservations with other sets of characteristics were more difficult to count.

Age is actually a big factor in how well we count people. That's also true in the rural areas.

By the way, the post-strata structure, of course, includes a special post-strata just for rural areas and especially for rural areas where they rent, which we know to be a hard-to-count population group, highly mobile and so forth. So we are—in that sense, the design takes care of—it sweeps across all of the problem situations in the country, not just fixed on the one.

But, yes, sir, we will have a particularly difficult time with that particular population group.

Mr. DAVIS OF ILLINOIS. I thank you very much. I think it's been a very productive hearing, and I certainly want to thank you for your responses.

And, Mr. Chairman, I want to thank you; and I'm going to yield to the ranking member here.

Mrs. MALONEY. I tell you I'm getting tired of all of this, of this constant effort really to disrupt the efforts of the professionals at the Census Bureau from doing their job and from correcting the undercount.

I would like to put in the record an article from the "Washington Post" written by David Broder entitled, Playing Hardball on the Census; and I think he clearly puts into focus what's going on.

He says, in preparing for the showdown on the census, Republicans reshuffled the leadership of the House Census Subcommittee and hired its new staff director, Thomas Hofeller—this was back in 1998—a Ph.D. Professor and battle-tested GOP strategist in redistricting. And he talks about meeting with Hofeller and how Hofeller goes to a blackboard analysis of the Census Bureau's plans stressing the risks they see of serious miscalculation with untested techniques and a tight timetable.

But as I was leaving, Broder says, Hofeller offered a decidedly non-academic comment; and he said, "someone, he said, should remind Bill Daley, the Secretary of Commerce and overseer of the Census Bureau, that if he counts people the way he wants to, his brother, Chicago Mayor Richard Daley, could find himself trying to run a majority-minority city." This is Hofeller talking. And Broder

then explains this blunt reference to racial ethnic realities is not uncommon on either side of the fight.

Among the thick file of scholarly papers Hofeller gave me was a memo entitled, Why Conservatives Should Be Opposed to Census Sampling; and it went on and said and warned in these papers—again a direct quote from the Republican papers—a census that uses sampling and statistical adjustment will be the biggest victory for big government, liberalism since the enactment of the Great Society. These statistical techniques will be used to add millions of virtual people to big-city population centers, thus increasing the political power and levels of Federal Government funding in those jurisdictions.

Then came two pages of answers of how this outrage can be stopped. And it outlines the courts, in Congress, the grassroots, and they are trying to do this.

We have been to the Supreme Court. I understand there has been another suit filed in Virginia against scientific sampling, and I remind my colleagues that two budgets have been held up. Anti-sampling language was attached to a disaster relief bill, and yet we have pages and pages of testimony that there is an undercount. We either correct it when the professionals have told us how to correct it or we deliberately don't count people. That's what this hearing is about, whether we correct for the undercount or whether we do not and therefore deliberately not count people.

Mr. MILLER. Thirty seconds.

Mr. RYAN. Mrs. Maloney, I just wanted to put for the record, where I come from it's not a Republican-Democrat issue. Democrat—liberal Democrat politicians from my home State—Senator Herb Kohl, Mayor Norquist of Milwaukee—are also opposed to sampling. We think it's bad for our State. We think the scientific community is out on this one. So I just wanted to say it's not a conservative-liberal thing, Republican-Democrat thing everywhere. In some places, it is. It's just wrong to paint that very broad brush.

With that, I yield.

Mrs. MALONEY. May I respond? Because my name was mentioned.

Mr. MILLER. Let me make my statement, please. We are running low on time. We can come back if you want.

I want to put in the record an editorial by Peter Skerry in last Sunday's Washington Post. It was titled, We're Overstating the Importance of the Undercount. I think it's a good explanation of the fact we really are overstating the undercount.

What this hearing was about was whether we were going to use statistical methods and adjustments to a census. There is real, legitimate concern that the method will not work at the block level and to use it for the redistricting purposes—I think it was a good hearing. There are still a lot of questions to be answered. We'll be discussing this, I'm sure, for the next months ahead, but there is a real debate within the statistical community that the method will not succeed, and that's the reason we've got to be careful. As Justice Scalia said, there is a potential political manipulation.

On behalf of the subcommittee, I would like to thank you for appearing here today.

Mrs. MALONEY. May I respond?

Mr. MILLER. We have a vote. We'll come—if you want to come back——

Mrs. MALONEY. I would like to respond to what was stated by Mr. Ryan.

Mr. MILLER. We will then recess and come back after we have a vote.

Mrs. MALONEY. I would like to respond now for 2 seconds.

Mr. MILLER. If you can do this in 15 seconds, go right ahead. Otherwise, I'm going to adjourn it here.

Mrs. MALONEY. My dear friend and colleague from Wisconsin mentioned that it was not a division between the Democratic and Republican party, and he mentioned names in Wisconsin that supported his point of view. But there is a clear distinction between the two parties on a national level, from the President who supports the use of modern scientific methods to the entire leadership on the Democratic side.

And I would like to put into the record statements that have been reported by the press quoting the Republican leadership that they will not let it go forward. Newt Gingrich called it a dagger in the heart of the Republican leadership, and Linder said even if the court approved it he will stop it. When you say it is not a division between the two parties, it is——

Mr. MILLER. Mrs. Maloney.

Mrs. MALONEY [continuing]. It is clear. It is in the record not from my lips but from the independent press. I would like to put those statements in the record.

Mr. MILLER. Mrs. Maloney, present the records right now. Come on. We're trying to get a vote together. You talk about all this away from the partisanship, and all you want to do is go back to Newt Gingrich who left Congress over 2 years ago, 3 years ago. This is 2000. We're in the middle of the census.

On behalf of the Census Subcommittee, I want to thank you for being here today.

I ask unanimous consent that all Members' and witnesses' opening statements be included in the record. Without objection, so ordered.

In case there are additional questions that Members have for our witnesses, I ask unanimous consent that the record remain open for 2 weeks for Members to submit questions for the record and that the witnesses submit written answers as soon as practicable. Without objection, so ordered.

Meeting adjourned.

Mrs. MALONEY. I put the quote in the record from John Linder, the head of the RNC, now.

[Whereupon, at 12:04 p.m., the subcommittee was adjourned.]

[Additional information submitted for the hearing record follows:]

Thomas B. Hofeller
7119 Marine Drive
Alexandria, VA 22307

October 11, 2000

Honorable Dan Miller
Chairman
U. S. House Subcommittee on the Census
H1-114 O'Neill HOB
Washington, DC 20515

Dear Mr. Miller:

It is unfortunate that Mr. Broder misinterpreted my comments concerning the census and redistricting in the City of Chicago in my interview with him in June of 1998 (Wash. Post 6/21/98). I have discussed this matter with him privately and am convinced that he now understands my position.

In the decade of the 1980's I was involved in redistricting litigation concerning congressional, legislative and city council districts in Chicago. On each occasion I offered testimony SUPPORTING the creation of new minority districts in Chicago.

Specifically I offered testimony on behalf of the African-American minority plaintiffs in the Chicago City Council case of *Ketchum v. Byrne*, 630 F.Supp 551 (N.D. Ill. 1985). In this case the minority plaintiffs were attempting to gain a fair number of the ward seats in direct opposition to the city council redistricting plan enacted by a Council control by aldermen who were part of the old Daley Machine. I offered the court an analysis that proved that the minority neighborhoods were severely underrepresented in the Council.

In all court cases in which I have offered testimony, I have always taken a position in favor of fair minority representation. In many cases, including *Ketchum* my testimony has been in favor of increased minority representation – often in opposition to the Democratic establishment.

My comment to Broder dealt with the fact that the 2000 Census, being presided over by Daley's son, whether adjusted OR NOT, would result in fair representation for the African-American and Hispanic residents of Chicago. Given the expected results of the 2000 Decennial Census (adjusted or not) the minority voters in Chicago could, in fact, achieve a majority in the Council at last. The irony I attempted to present was that this event would happen under the Mayorship of Daley's other son. So the wrongs of the past, perpetrated under the leadership of the father and his successors, would be made whole under the leadership of the sons.

This is all I meant by my statement to Broder. This is an observation, which is supported by my record IN COURT, under oath. I have always supported the creation of fair minority districts in my 35 years of redistricting work, and will continue to do so in the future.

I explained this position, over a year and a half ago, to both Ben Chevat, Ms. Maloney's Administrative Assistant and to David McMillen, the lead census person on the Democratic Staff of the Committee on Government Reform.

I was disappointed, but not surprised, that Ms. Maloney once again brought up this misconception at a Subcommittee hearing this last spring. She or members of her staff, who I assume prepare her remarks, know that this is NOT what I said to Mr. Broder.

If Ms. Maloney brings this issue up again, I would appreciate your assistance in encouraging her to do the HONORABLE thing and put this issue to rest. I would also appreciate it if you could make this letter a part of you official records for the Subcommittee.

Thank you for your assistance in this matter.

Sincerely,

Thomas B. Hofeller

UNITED STATES DEPARTMENT OF COMMERCE
Economics and Statistics Administration
U.S. Census Bureau
Washington, DC 20233-0001
OFFICE OF THE DIRECTOR

JUN 0 8 2000

The Honorable Dan Miller
Chairman, Subcommittee on the Census
Committee on Government Reform
U.S. House of Representatives
Washington, DC 20515-6143

Dear Mr. Chairman:

I think it will be helpful to provide the Subcommittee with information following up on several of the issues discussed at the May 19 hearing on the Accuracy and Coverage Evaluation (A.C.E.) program.

In our discussion of the post-strata developed for the A.C.E. survey, I mistakenly indicated that "relationship" would be used as a post-stratification variable. This is not the case. Decennial Statistical Studies Division Procedures and Operations Memorandum #Q-24, which has been forwarded to the Subcommittee, presents the final post-stratification plan for the A.C.E survey. You will see after reviewing that memorandum that "relationship" is not one of the variables we will be using.

At the hearing, you asked about the assessment of accuracy of decennial census counts, including block-level data. Two types of accuracy need to be considered: numeric accuracy and distributive accuracy. These terms were coined in the context of the evaluations of the 1990 census post-enumeration survey. Numeric accuracy refers to how close the overall count of a particular geographic area or demographic group is to the actual number of people who reside in that area or belong to that group. Distributive accuracy refers to how close the relative proportion or share of a demographic group or geographic area is to the true share.

Numeric and distributive accuracy are both critical goals in conducting the census. In planning census processes, the Census Bureau designs operations to correctly enumerate people and housing units. Considerations of numeric accuracy are key to census planning, because there can be reasoned predictions as to the effect that a proposed operation will have on numeric accuracy. It is difficult if not impossible to know *a priori* the effects of a particular census operation on distributive accuracy—distributive accuracy can only be assessed once the census is complete, and its assessment is an important component of census evaluations. An assessment of distributive accuracy should properly focus on uses of census data that require a "proportionate shares" or "fixed pie" allocation.

In designing census operations, the Census Bureau does not reject operations that would improve numeric accuracy and meet other criteria for inclusion, though such operations might affect distributive accuracy negatively (or indeterminately). For example, the Local Update of Census Addresses (LUCA) program most likely results in improvements in numeric accuracy for those participating jurisdictions, but because not all jurisdictions participate or invest the same level of resources even if they do participate, LUCA could reduce overall distributive accuracy.

With regard to the accuracy of block-level data, I inadvertently said that block-level data, based on a statistical correction of the census, would be more accurate. I cannot predict whether statistically corrected data will be, on average, any more or any less accurate than the uncorrected block data. Moreover, it is inevitable that the A.C.E. will result in the population in some blocks being overestimated and the population in other blocks being underestimated. However, this is not the important issue. The use of statistical methods to correct initial census results was never intended nor expected to produce unqualified improvement in the smallest geographic areas, like blocks. The uses of census data are based on aggregating blocks into larger geographic areas. It is the accuracy of these aggregate data that is critical to the use of census data, not the accuracy of any individual block. Therefore, accuracy should be discussed and analyzed at the aggregated levels for which census data are used and in consideration of how the data are used. The Census Bureau expects that the A.C.E. estimates will produce better data for aggregations of blocks that are basic to the uses of census data, including states, congressional and other legislative districts, and cities.

Finally, you asked about the internal committee that has been formed to develop and implement the A.C.E. program. The Census Bureau established the Executive Steering Committee for A.C.E. Policy (ESCAP) in late 1999 to advise the Census Bureau Director in determining policy for the A.C.E. and on the integration of the A.C.E. results into Census 2000. This Committee meets every two weeks to discuss technical and policy issues associated with the A.C.E. The Committee will provide recommendations, as required, on these issues to the Director. The ESCAP is chaired by Mr. John H. Thompson, Associate Director for Decennial Census, a career employee of 25 years with the Census Bureau. It includes the following senior Census Bureau staff:

Deputy Director and Chief Operating Officer
Principal Associate Director and Chief Financial Officer
Principal Associate Director for Programs
Associate Director for Methodology and Standards
Associate Director for Demographic Programs
Assistant Director for Decennial Census
Chief, Decennial Statistical Studies Division
Chief, Planning, Research and Evaluation Division
Chief, Population Division
Chief, Decennial Management Division
Senior Mathematical Statistician

Should you have any questions about this information or require any further information, I trust that you will have a member of your staff contact Ms. Robin Bachman, Chief of the Congressional Affairs Office, at (301) 457-2171.

Sincerely,

Kenneth Prewitt
Director

cc: The Honorable Carolyn B. Maloney

UNITED STATES DEPARTMENT OF COMMERCE
Economics and Statistics Administration
U.S. Census Bureau
Washington, DC 20233-0001
OFFICE OF THE DIRECTOR

JUN 21 2000

The Honorable Dan Miller
Chairman, Subcommittee on the Census
Committee on Government Reform
U.S. House of Representatives
Washington, DC 20515-6143

Dear Mr. Chairman:

Thank you for your letter of June 9, 2000. This responds to Representative Paul Ryan's questions, itemized below, following up on the U.S. Census Bureau's hearing before the Government Reform Subcommittee on the Census, which took place on May 19, 2000.

"Certain parts of the state [Wisconsin] are experiencing unusually low response rates in areas where the expected rates, according to your web site estimates, would indicate normally high response rates. Many in Wisconsin have expressed concern that these low response rates are related to some of the post office box problems.

"It is my understanding that many residents who rent PO boxes because they reside in rural areas received letters indicating the Census was coming and letters reminding them to fill out their forms, but never received forms. The forms were sent to PO boxes, apparently accidentally, then retrieved; however, some, but not all, were then redelivered by hand.

"In light of these errors, I am concerned about the accuracy of the Master Address File. Wasn't LUCA supposed to identify these types of things? Wouldn't many of these same people with PO Boxes also had PO Boxes ten years ago?"

In developing the Master Address File (MAF), the Census Bureau started with its 1990 census address list. This was updated with address information from the United States Postal Service for areas that have city-style (house number/street name) addresses. In areas with other types of addressing systems (for example, those typically found in small towns and predominantly rural areas), the address list was updated with addresses and location descriptions from a Census Bureau Address Listing operation. Local and tribal governments were then invited to review these updated lists for accuracy, adding and deleting addresses where appropriate. Temporary Census Bureau workers also canvassed areas to verify existing addresses and record new ones.

Post Office box addresses were not included in the MAF, because a post office box is associated with a person, not a specific housing unit. That is, a person rents a post office box and can retain that box even if they change their residence from one address to another, as long as they continue to receive their mail in the same post office area. In addition, all post office boxes are located in a specific post office, which is located in a single census block. Census 2000 requires that the MAF associate each response with the specific geographic location of the housing unit in which the responding person/family

lives. We can make that association for the households that use city-style addresses and for the households where temporary Census Bureau workers recorded location description addresses in areas without city-style addresses, but not for post office box addresses. Therefore, in preparing the address list for city-style areas, we developed a list of house number/street name addresses associated with a specific block, even if the mail is delivered to a post office box.

There are some mail-out areas where a large number of households receive their mail at a post office box, despite the fact that those households have city-style addresses and are located in mail-out/mail-back areas. In these instances, the USPS regulations did not allow for the delivery of the Census 2000 questionnaires. The Census Bureau, as part of the distribution plan, was prepared to have employees retrieve the forms from the USPS and redeliver them to the appropriate city-style address. In other situations, the households were not enumerated until they were contacted during the Nonresponse Follow-up (NRFU) operation. It is worth mentioning that the NRFU has gone very well in Wisconsin and that operation is virtually finished in all parts of the state.

The Local Update of Census Addresses (LUCA) program was successful, and many new addresses and street/road updates were provided as a result of the careful review undertaken by many local and tribal governments. We believe that the combination of LUCA, the New Construction program, and the block canvassing conducted by temporary Census Bureau workers resulted in a very accurate address list.

"How many communities in WI and nationwide were sent forms to PO boxes in error?"

The Census Bureau did not send forms to post office boxes. USPS regulations prevented the delivery of forms to the addresses of households that received their mail in post office boxes. After we complete the NRFU and process the results, we expect to be able to identify those mail-out/mail-back areas that had a significant number of post office box deliveries. We will provide this information to you at that time.

"If the Census Bureau knew ahead of time that they were not going to deliver forms to PO boxes, why were thousands of forms delivered to PO boxes and then retrieved?"

The Census Bureau did not know ahead of time exactly which city-style addresses had their mail delivered to post office boxes. If the information had been available, we would have expanded the update/leave or update/enumerate program in Wisconsin and not mailed the questionnaires. As we anticipated during our planning for Census 2000, the undelivered/undeliverable questionnaires the Census Bureau retrieved from many post offices were those mailed to city-style addresses that the USPS could not deliver because the household uses a post office box address for mail delivery. Understanding the

difficulties with this type of situation will lead to better planning for the 2010 census, provided there is adequate funding to conduct a thorough address list updating and improvement operation.

"How much extra money did it cost the department to first mail these, then retrieve the forms and hand-deliver them?"

Based on our experience in previous censuses, our plans included an operation whereby Census Bureau employees would redeliver questionnaires that the USPS could not deliver. A total of $25 million was budgeted for this operation.

"The Department of Administration of Wisconsin has indicated that not all of these forms were redelivered by hand; they are concerned that the LCOs may not be aware of which residences did not receive a form by mail.

"Additionally, the state is concerned, especially after reviewing the response rates to date, that the LCOs have not taken into consideration the need for additional enumerators in these PO box areas?

"Does the time frame for nonresponse followup include time to enumerate all of these PO box areas? How much time is set aside for nonresponse followup compared to statistical sampling?"

Assignment of personnel to conduct the NRFU operation is based on a number of factors, particularly the number of nonresponding households. Any form mailed by the Census Bureau, whether delivered or not, for which the Census Bureau did not receive a response, would automatically be included in the NRFU workload. The NRFU operation is scheduled to be conducted for up to ten weeks. The Accuracy and Coverage Evaluation interviewing is scheduled to be conducted for up to eight weeks. During its planning process, the Census Bureau identified an anticipated NRFU workload and hired sufficient staff for most parts of Wisconsin. The result is we are on schedule, and as mentioned above, the NRFU operation is virtually complete for the state of Wisconsin.

We hope this letter addresses Representative Ryan's concerns. Should you or he require further information, please have a member of your respective staffs contact Ms. Robin Bachman, Chief of the Congressional Affairs Office, at (301) 457-2171.

Sincerely,

Kenneth Prewitt

Kenneth Prewitt
Director

cc: The Honorable Carolyn B. Maloney

○