S. HRG. 107–1061

YEAR 2000 DECENNIAL CENSUS

HEARING

BEFORE THE

COMMITTEE ON COMMERCE, SCIENCE, AND TRANSPORTATION UNITED STATES SENATE

ONE HUNDRED SEVENTH CONGRESS

FIRST SESSION

MARCH 28, 2001

Printed for the use of the Committee on Commerce, Science, and Transportation

U.S. GOVERNMENT PRINTING OFFICE

88–017 PDF WASHINGTON : 2004

For sale by the Superintendent of Documents, U.S. Government Printing Office
Internet: bookstore.gpo.gov Phone: toll free (866) 512–1800; DC area (202) 512–1800
Fax: (202) 512–2250 Mail: Stop SSOP, Washington, DC 20402–0001

SENATE COMMITTEE ON COMMERCE, SCIENCE, AND TRANSPORTATION

ONE HUNDRED SEVENTH CONGRESS

FIRST SESSION

JOHN MCCAIN, Arizona, *Chairman*

TED STEVENS, Alaska	ERNEST F. HOLLINGS, South Carolina
CONRAD BURNS, Montana	DANIEL K. INOUYE, Hawaii
TRENT LOTT, Mississippi	JOHN D. ROCKEFELLER IV, West Virginia
KAY BAILEY HUTCHISON, Texas	JOHN F. KERRY, Massachusetts
OLYMPIA J. SNOWE, Maine	JOHN B. BREAUX, Louisiana
SAM BROWNBACK, Kansas	BYRON L. DORGAN, North Dakota
GORDON SMITH, Oregon	RON WYDEN, Oregon
PETER G. FITZGERALD, Illinois	MAX CLELAND, Georgia
JOHN ENSIGN, Nevada	BARBARA BOXER, California
GEORGE ALLEN, Virginia	JOHN EDWARDS, North Carolina
	JEAN CARNAHAN, Missouri

MARK BUSE, *Republican Staff Director*
ANN CHOINIERE, *Republican General Counsel*
KEVIN D. KAYES, *Democratic Staff Director*
MOSES BOYD, *Democratic Chief Counsel*

CONTENTS

WITNESSES

APPENDIX

YEAR 2000 DECENNIAL CENSUS

WEDNESDAY, MARCH 28, 2001

U.S. SENATE,
COMMITTEE ON COMMERCE, SCIENCE, AND TRANSPORTATION,
Washington, DC.

The Committee met, pursuant to notice, at 9:28 a.m., in room SR–253, Russell Senate Office Building, Hon. John F. Kerry, presiding

OPENING STATEMENT OF HON. JOHN F. KERRY, U.S. SENATOR FROM MASSACHUSETTS

Senator KERRY. The hearing will come to order. We are going to start just literally a moment early, because we are under the time gun here. A number of our colleagues from the House have requested to testify early, and I appreciate that.

I wanted to try to accommodate it. The problem is, we also have the Secretary under the same time pressure so we are going to try to move as expeditiously as possible and not lose part of the hearing in doing so.

Mr. Secretary, thank you very much for taking the time to come up. We appreciate it. As the Members of the Committee know, the census was something that a lot of us have been following very closely, and I am grateful to Secretary Evans, who said during his confirmation hearing in January that he would follow this closely and be happy to appear before the Committee again on that, and I appreciate his honoring that commitment today.

I am not going to go through my full statement.

Obviously, what is at stake here is the accuracy of the census, which really is at the heart of a number of issues, not just redistricting and the adequacy of representation in the country, but also the distribution of funds, and whether or not there is a sense of fairness and inclusivity in the process by which we not only represent America but see that America receives its fair share of the funds distributed federally.

There are roughly $185 billion a year that is distributed based on population counts that come from the census over a 10-year period. You are obviously, therefore, looking at over $1 trillion to $2 trillion. That makes a difference in the lives of our fellow Americans, and so it is particularly important for people to have a sense that this was done properly, fairly, and in a way that every American can have confidence that they are properly represented.

There are comments I would make with specificity about Demographic Analysis and other things, but I do not want to lose the

time now that we need desperately to move forward in terms of testimony and questions, and I will try to do it in the course of that.

[The prepared statement of Senator Kerry follows:]

PREPARED STATEMENT OF HON. JOHN F. KERRY,
U.S. SENATOR FROM MASSACHUSSETTS

Thank you, Mr. Chairman. I would like to thank the Chairman for calling this hearing, and especially for managing to schedule it during what we all know is an incredibly busy time for him. His willingness to arrange this hearing is reflective of his excellent leadership and of his understanding of the importance of this issue.

As the Committee knows, the census is an issue that I have been following very closely and that I am extremely interested in. I thank Secretary Evans for coming again before this Committee to address our questions and concerns. When Secretary Evans was before us for his confirmation hearing in January, the census and the possibility of using corrected data to compensate for people missed by the census were issues raised by several senators. During his confirmation hearing, the Secretary assured us that he would be happy to appear before this Committee again on this matter, and I appreciate his willingness to honor that commitment and testify here today.

I know that we have many distinguished witnesses here, and I don't want to delay the Secretary or the House Members, who I know have very busy schedules to maintain. But I would, just briefly, like to make a few comments about what we'll be discussing this morning. First of all, I believe very strongly that it is critically important to understand the consequences of missing people in the decennial census. I am quite pleased that the Census Bureau believes it has possibly achieved the most accurate census ever. I know that it took the tireless efforts of over 800,000 people and significant resources to reach this goal, and I commend all of those involved for their excellent work.

More good news: earlier this month, the Census Bureau reported that it had dramatically reduced both the net and the differential undercounts from the last census. Again, I commend all of those who dedicated themselves to improving the census. But in the midst of this progress, we must not lose sight of the fact that at least 6.4 million people were not counted in Census 2000. Blacks were missed at twice the rate of Whites, Hispanics were missed three times as often, and American Indians were missed five times as often.

The Bureau has spent $400 million on the Accuracy and Coverage Evaluation, or A.C.E., which evaluates accuracy and corrects for any undercount. As we know, earlier this month the Bureau did not recommend releasing the A.C.E. for the purpose of redistricting, and therefore, the official data remains the raw count, short over 6 million people.

It is my understanding that the Census Bureau has not yet released the complete data that would tell us who the people are who were not counted, but if the 1990 census is any indicator, then we can be sure that they were predominantly poor and predominantly minority. What can be gleaned from the 1990 data is that unless the 2000 census data is corrected to account for those missed by census takers, vast numbers of poor, minorities in this country will be denied hundreds of millions of dollars in federal assistance.

President Bush has often repeated that he is committed to leaving no child behind. I can think of no greater or clearer example of millions of children—literally—getting left behind than not being included in the census and potentially being denied federal funding. Roughly $185 billion per year in federal funding is distributed based on population counts derived from the 1990 census, to say nothing of the policy decisions that were made in states, cities, and counties around the nation based on this data. We must do everything in our power to ensure that this situation does not repeat itself. Indeed, the Census Bureau has worked hard to develop a methodology to respond to this situation. I sincerely hope that they are able to solve the mysteries of this issue before the fall to ensure that no child is left behind.

My hope for this hearing today is that Secretary Evans will help to clarify the Census Bureau's March 1 decision not to use the corrected number for redistricting. I am concerned that there is misinformation swirling about, there is an impression that the Bureau ruled against the use of modern statistical methods in any census data. I have read their recommendation, and I do not believe that to be the case. My interpretation of their report—which I think is relatively clear—is that the Census Bureau needed more time to verify the accuracy of A.C.E. Further, the Bureau stated that there is considerable evidence to support the use of corrected data and

that the majority of the evidence indicates the superior accuracy of the corrected data.

The reason that both clarifying this point is so important and understanding where the Secretary stands with respect to this process is so important, is that in the fall the Bureau will release numbers for use in determining federal funding allocations. It is unclear to me who will make the decision about which numbers to use and how that decision will be made. It is my sincere hope that the Secretary will elucidate this matter for us.

Senator KERRY. Mr. Ranking Member, do you have any comments?

STATEMENT OF HON. ERNEST F. HOLLINGS, U.S. SENATOR FROM SOUTH CAROLINA

Senator HOLLINGS. Mr. Chairman, I thank Senator McCain and yourself and, of course, the distinguished Secretary for being here. My statement will go in the record. I have another hearing, as we always do, so I will move to that and I will cut it short, because we have got House Members who are trying to make a roll call.

But Secretary Evans, it is sort of good to harken the history of the census count and the problems we have had in the 1990 census. President George Bush, Senior was in at that particular time and, as to the count, the lawsuits broke out like the measles with all of the undercounts and everything else. So we got the National Academy of Sciences to see how best to get the most complete, the most full, and the most accurate count possible. The Academy belabored it and came up with the sampling solution, which of course we have avoided thus far, but I feel very strongly in favor of that approach.

Everybody knows the politician stands in the well and says the Constitution says count everybody, which is totally impossible, yet on the other hand you want to make the best endeavor that you possibly can. Right to the point, what we want to avoid are headlines like today's headline: "Cities' Minorities Losing Census Undercount." But here we go again. In the famous words of our leader, Ronald Reagan. Here we go again.

So that is what concerns this Committee. We thought we had the right approach as developed, and that is why Secretary Mineta did not want to make the decision on sampling. He left it to the professionals because he did not think he had expertise to analyze the scientific studies which showed should be done, that is, use the sampling approach.

So I wish you would consider that background and history, and the present-day cries right now with respect to the problem we are having, because nobody's trying to get an overcount of anything. There is no chance of that, but the best system that I know of that has been devised so far, unless the Secretary has got a better one, is that sampling approach. We do appreciate your appearance here before the Committee, and I apologize for having to get to this other Committee meeting.

Thank you, Mr. Chairman.

[The prepared statement of Senator Hollings follows:]

PREPARED STATEMENT OF HON. ERNEST F. HOLLINGS, U.S. SENATOR FROM SOUTH CAROLINA

I would like to thank the Chairman for holding this hearing today. In addition, I would like to thank Secretary of Commerce Donald Evans for appearing before us

today to honor a commitment that he made during my all-too-brief Chairmanship of this Committee in January to return and have a frank discussion with us about the Census.

Mr. Secretary, as I have told you before, I am a strong supporter of statistical sampling. Even though some are calling the 2000 Census "the best Census ever," I am of the opinion that no Census is good enough if it leaves more than 3 million people uncounted.

I also believe that decisions regarding the use of sampling should be made by the professionals not by the politicians. It is convenient for the Administration that the professionals recommended against sampling, but I wonder whether the threat of being overruled influenced that recommendation.

However, I do not want to play the blame game that we so often play here in Washington. I want to look to the future. First of all, I would like to see the Census fully release its Accuracy and Coverage Evaluation (A.C.E.), upon which any adjustment to the Census numbers—any sampling—would be based. If the A.C.E. truly is flawed, what better way to know than to allow independent statisticians to take a look at the full data?

Second, I would like to encourage the Census to keep working to improve the 2000 data. Simply because the preliminary redistricting data has been released, the issue of accuracy in the Census is not resolved. We use these data, adjusted from year to year, as the basis for allocation of government funding. To that end, I applaud the Bureau for continuing to work on finding answers to the questions that they have about the adjusted data.

The issue of ensuring that the entire population is counted is very real, not just for large cities like New York and Los Angeles, but also for my State of South Carolina. Nearly 2 in 5 households in my State did not mail back a Census form—the Nation's lowest response rate except for Alaska. I am worried that thousands of South Carolinians were not counted. The State Budget and Control Board estimates that each uncounted resident could cost the State $2,000 per person in Federal funding. Overall, an inaccurate count could cost my State millions—and South Carolina is only the tip of the iceberg.

In conclusion, when groups and areas are undercounted, the strength of their vote is diluted, they do not get their fair share of Federal funding, and they suffer economically when banks and insurance companies base their decisions on demographic data. So I hope that the Census Bureau and the Department will continue to improve its data in order to achieve the most accurate count. That is the only way to ensure justice and fairness are the guiding principles in this vital matter.

Senator KERRY. Thank you very much, Senator Hollings.

Senator Inouye.

STATEMENT OF HON. DANIEL K. INOUYE, U.S. SENATOR FROM HAWAII

Senator INOUYE. Like my colleague, I will have to be leaving soon, Mr. Chairman. If I may, I would like to hear the Secretary's testimony.

Senator KERRY. Mr. Secretary, let us, if we can, invite your testimony, and then we will try to get into questions and see if we can incorporate some of the issues I talked about earlier.

STATEMENT OF HON. DONALD L. EVANS, SECRETARY OF COMMERCE; ACCOMPANIED BY MR. WILLIAM BARRON, ACTING CENSUS DIRECTOR

Secretary EVANS. Thank you very much. It is a pleasure to appear before this Committee again to discuss Census 2000. You have my prepared statement for the record, but I would like to just briefly touch on some of the points from that statement here today and then answer any questions you might have.

Senator KERRY. Thank you. Your full statement will appear in the record.

Secretary EVANS. Thank you very much, Senator. I appreciate that. I have said Census 2000 is the most accurate this nation con-

ducted in the past 60 years, perhaps ever. Its success is a direct result of the hard work and dedication of the employees of the Census Bureau and of hundreds of thousands of people, including thousands of your constituents who worked on Census 2000.

Census 2000 was an operational success. The Census Bureau met the deadline for releasing data, counted more residents than ever before, reduced the estimated undercounts, and reversed the trends toward smaller percentages of the population responding to the census.

One of the major reasons for the operational success of Census 2000 was Congress' commitment to provide full funding that brought a number of improvements to bear on Census 2000. The Census Bureau carried out an unprecedented outreach program specifically targeted to groups that historically have had the highest undercounts. They reached out to families, neighborhoods, and communities to encourage participation.

First, the Census Bureau developed an aggressive marketing plan and partnership program. Partnerships with a variety of governmental community and educational groups, 140,000 in all, were key to building support for census participation all the way down to the neighborhood level. These helped the Census Bureau to knock down obstacles that might get in the way of participation.

The bureau also greatly expanded its outreach into communities by using paid advertising for the first time, placing more than $100 million in ads to educate and motivate the public to respond, especially those who had been undercounted at higher rates in the past censuses.

Second, the Census Bureau took its educational outreach program directly into schools, providing materials to classroom teachers so that they could teach lessons on the census.

Third, the Census Bureau made questionnaires more user-friendly, so they were easier to read and to fill out. They published questionnaires in six languages and provided materials to help individuals fill them out in 49 languages.

Another improvement that went a long way to Census 2000's success was the hiring of highly skilled staff on a temporary basis to get the job done on time. Through its "Quality Counts" operation, the Census Bureau redoubled its efforts to ensure quality and completeness in the count, identifying homes that should be visited again to review and verify information or to fill in the blanks.

In using the advanced technology, the Census Bureau was able to process the data faster and to introduce quality assurance steps to be sure that data were captured accurately. Along with other results, due to these efforts the Census Bureau achieved a higher response rate through the mail than was expected. In fact, people in about 65 percent of the homes that were mailed questionnaires had already responded by mail before it came time for the Census Bureau to follow up with its door-to-door field operation.

As the Members of this Committee know, the Census Bureau conducted an independent survey of approximately 314,000 homes. The Accuracy and Coverage, Evaluation Survey, or A.C.E., that was designed to measure net census coverage according to the current estimates from the survey, Census 2000, achieved net cov-

erage rates for the total population of about 99 percent. The estimated coverage rates for individual ethnic groups were also extremely high. A.C.E. estimates support the conclusion that Census 2000 achieved both reduced net and differential undercoverage from the 1990 Census levels.

Of particular interest, the national undercount for race and Hispanic origin groups was reduced from the 1990 census rate of 1.61 percent to 1.18 percent for Census 2000, a reduction of about 25 percent.

Throughout the planning for Census 2000, a major issue of concern to the Census Bureau was whether the results of the Accuracy and Coverage Evaluation Survey could be used to make the census counts more accurate. A committee of Census Bureau professionals was formed to evaluate whether using the A.C.E. to adjust the census figures would improve the results for using redistricting.

The professionals serving on the committee collectively have 390 years experience as statisticians and demographers. Bill Barron, the Acting Director of the Census Bureau, presented me with that committee's report and recommendation, along with his own recommendation, on March 1, 2001.

Mr. Barron agreed with and approved the committee's recommendation that the unadjusted census data should be released as the Census Bureau's official redistricting data. The committee of Census Bureau professionals reached its recommendation because it was unable, based on the data and other information it had at the time, to conclude that the adjusted data were more accurate for use in redistricting.

The report of the committee is being made available for the record. In a nutshell, the A.C.E. results could not be squared with the population estimates derived from Demographic Analysis, a long-accepted method of judging census quality. There were other significant potential errors that might affect the accuracy of the A.C.E. estimates.

After receiving the Census Bureau's recommendation, I thoroughly reviewed the report and supporting materials, and I consulted with a diverse group of respected, non-Government statisticians and demographers in addition to Census Bureau professionals. On March 6, I announced my decision to release the unadjusted data for use in the redistricting process. In making my decision, I followed a process that was transparent, responsible, reasonable, and fair, and took full account of the view of experts, both career Government professionals, and professionals from outside Government.

As you know, the Census Bureau is obligated to complete release of the redistricting data to the states by April 1, and it will do so. I should emphasize that the committee of Census professionals could not have resolved critical questions about use of adjusted data prior to that deadline, or even soon thereafter, I am confident that the committee did all that it could, and that it reached the only reasonable conclusion.

Some have requested that the bureau release A.C.E. adjusted data notwithstanding the committee's recommendation to release unadjusted numbers as official redistricting data. The Department will not do so, because the A.C.E. numbers simply are not fit for

use. From everything we currently know, the actual head count was extremely accurate. The A.C.E. projection would not be an improvement, and those estimates very well may change as a result of the bureau's ongoing analysis.

It would be irresponsible for the Department to release redistricting data for which it cannot gauge the accuracy. Over the coming months, the bureau will gather additional data and continue its analysis of Census 2000 results. Before the end of 2001, the bureau will determine whether or not to recommend statistically adjusting the Census 2000 population estimates for use in other Census Bureau products, and as the basis for annual adjustments that will commence in December 2001. Any such adjustment data may also be used in the allocation of funds under Federal programs and for other purposes. At the moment, there is really no basis on which to predict what the outcome of the bureau's further work will be.

The Census Bureau is determined to build on the success of Census 2000, and the Bush Administration supports that effort, as reflected in the President's budget, details of which will be forthcoming. It funds continuing efforts for the Census Bureau to improve the census as we progress toward 2010.

I look forward to continuing to work with Members of this Committee and other interested committees and Members of Congress to define and provide appropriate support for a total Census 2010 effort.

Mr. Senator, that concludes my testimony, and I will be pleased to answer any questions you might have.

[The prepared statement of Secretary Evans follows:]

PREPARED STATEMENT OF HON. DONALD L. EVANS, SECRETARY OF COMMERCE

Mr. Chairman, Senator Hollings, and Members of the Committee:

It is a pleasure to appear before this Committee again and, on this occasion, to discuss with you Census 2000.

CENSUS 2000: AN INVALUABLE ACHIEVEMENT

I have said that the 2000 Census is the most accurate census this nation has ever conducted. A Constitutional mandate, censuses have been conducted every 10 years since 1790—22 times in all. So the success of Census 2000 is a remarkable achievement. It is attributable to the hard work and dedication of the professional staff at the Census Bureau and all the hundreds of thousands of people, including thousands of your constituents, who worked on Census 2000. I commend them all. We are indebted also to the American public, whose response exceeded expectations; to the thousands of Census partner organizations; and to the Congress, for your oversight, support, and vision in providing sufficient resources to conduct Census 2000.

This is an exciting period for those who want facts to bolster their understanding of our nation's people. The Census Bureau began to roll out the results of Census 2000 just three months ago, with the release of the state population totals used for apportionment. Again meeting the schedule set by the Congress, by the end of this week, the Census Bureau will have released for all 50 states, the District of Columbia, and Puerto Rico, population data—by age, race, and Hispanic Origin—that will be used to redraw legislative districts.

Not only do these current data releases allow the redistricting process to begin, but these are the first data from Census 2000 for counties, cities, towns, townships, and smaller geographic areas. They are the first race and ethnicity data from Census 2000, and the first to show the effect of multiple race reporting. Not a day has gone by in the last month without numerous news stories about the increasing diversity of our population and about which areas are growing in population and by how much. I share the fascination: These data tell us so much about ourselves, our neighbors, our great country. This gold mine of information will continue to yield a wealth of information as the Census Bureau prepares and releases much more

data over the next two years. All of this attention focused on population data reminds us what an important national resource we have in the census.

CENSUS 2000: AN OPERATIONAL SUCCESS

Census 2000 was an operational success. The Census Bureau met or exceeded its goals, including meeting the mandated deadlines for releasing data for use in apportionment and redistricting. This success can be attributed to the Congress' commitment to provide full funding for a number of improvements, including unprecedented outreach programs to groups that historically had the highest undercounts:

• **Marketing and Partnerships:** First, the Census Bureau implemented a multi-faceted, aggressive marketing and partnership program to encourage householders to include themselves in the census, by completing and mailing back their census forms. Based on the experience of declining response rates over the preceding three censuses, the Bureau had anticipated that fewer households would return forms by mail in Census 2000. Partnerships—140,000 in all—with state, local, and tribal governments; community and advocacy groups; the private sector; religious organizations; educational institutions; and the Congress were key to building support and removing obstacles to participation in the census. The Bureau successfully implemented paid advertising for the first time in Census 2000, placing over $100 million in media buys designed to educate and motivate the public to respond. Paid advertising also allowed the Census Bureau to target ads to groups that had been undercounted at higher rates in past censuses.

• **Educating Families:** As part of the Census in Schools program, the Census Bureau provided lesson plans, wall maps, and take-home materials to classroom teachers so they could teach lessons on the census.

• **User Friendly Questionnaires:** The Census Bureau designed the questionnaires so that they would be easier to read and fill out. The Bureau also sent advance letters and reminder cards before and after the questionnaires were mailed out to increase response. The Bureau further offered multiple ways to respond, to ensure everyone had a chance to include themselves in the census. These included printing questionnaires in six languages and making available upon request materials in 49 languages to assist people in completing the questionnaire.

These cumulative outreach efforts were successful. The expected mail response rate of 61 percent was significantly exceeded, reaching about 65 percent by the start of the field operation to follow up on homes for which a questionnaire was not returned.

• **Staffing:** The Census Bureau hired and retained enough highly skilled temporary staff, throughout the course of the census, to complete all operations on time. Because of a resourceful recruiting plan, research on pay rates and recruiting, and the attractive wages that the Census Bureau could offer because of the full census funding that the Congress provided, the Census Bureau was able to recruit some 3.7 million job candidates and eventually hire 960,000 people over the course of the census. Over 500,000 worked on the operation to follow up on those homes for which a questionnaire was not returned, and, through their hard work, the Census Bureau was able to complete the enormous task of personally visiting 42 million homes slightly ahead of schedule.

• **Quality Checks:** Because of the timely completion of the follow-up operation, the Census Bureau had the time and resources to conduct other operations designed to improve coverage, including additional re-enumeration efforts in selected areas. The Census Bureau called these operations "Quality Counts." Based on Census Bureau experience and using various quality indicators, the Census Bureau identified about 10 percent of the Nation's homes that it believed should be visited again in these review, verification, and clean-up operations. If it had not conducted these additional operations, the Census Bureau would have provided an incomplete enumeration of the population. The "Quality Counts" operations helped improve coverage and the census count.

• **Technology:** For Census 2000, the Census Bureau used digital imaging and optical-character recognition technology for the first time to recognize handwritten answers in addition to marked circles or boxes. This was a vast improvement over previous computer systems and allowed the Census Bureau to process the data faster and introduce quality assurance steps to be sure they had captured the data accurately. During the peak of questionnaire receipts, the Census Bureau's data capture centers processed 3.3 million forms a day. Each bit of information on the captured census forms was transmitted over secured lines to the Census Bureau headquarters, where staff performed quality control checks to ensure they had complete data. The improved data capture systems, with the ability to capture names, also

meant that the Census Bureau could offer multiple options for responding to the census with confidence that it could find and remove duplicate responses.

THE RESULT: A HIGHLY ACCURATE HEADCOUNT

The operational improvements not only contributed to the ability to meet legal deadlines, but more importantly they also produced an improved count. The Census Bureau conducted an independent survey of approximately 314,000 housing units—called the Accuracy and Coverage Evaluation (A.C.E.)—that was designed to measure net census coverage. It was also designed to measure differences in coverage rates for key groups.

The first chart attached to this testimony illustrates the remarkable job the Census Bureau did in counting people in Census 2000. According to current estimates from the A.C.E., Census 2000 achieved a net coverage rate for the total population of 98.82 percent. Even better, the estimated coverage rates for individual groups were also very high. The coverage rate for Non-Hispanic Blacks was 97.83 percent; for Hispanics, 97.15 percent; for American Indians and Alaska Natives on Reservations, 95.26 percent; for American Indians and Alaska Natives off Reservations, 96.72 percent; for Native Hawaiians and Other Pacific Islanders, 95.40 percent; for Non-Hispanic Asians, 99.04 percent; and for Non-Hispanic Whites, 99.33 percent. The A.C.E. results thus support the conclusion that Census 2000 achieved both reduced net and differential undercoverage from 1990 census levels.

Attached are two additional charts showing estimated net undercount rates for key groupings of the population for 1990 and 2000. Chart 2 shows net undercount rates for the total population and race and Hispanic-origin groups. Chart 3 shows net undercount rates for age and sex groups, owners, and renters.

The A.C.E. estimates that the net national undercount was reduced from the 1990 census rate of 1.61 percent to 1.18 percent for Census 2000, a reduction of about one-fourth. This reduction is substantial and reflects high census quality. The A.C.E. further found that not only was the net undercount reduced, but there was a substantial reduction in the undercount rates for certain groups and in the differential undercount. In 1990, minorities, renters, and children were differentially undercounted, that is, undercounted at higher rates than the population as a whole. While these groups still have higher undercount rates, the differential has dropped considerably.

The estimated undercount rate for Non-Hispanic Blacks was cut by about half—it dropped from 4.57 percent in 1990 to 2.17 percent in 2000; and the estimated undercount rate for Hispanics dropped by about 40 percent from 4.99 percent to 2.85 percent. The undercount rate for American Indians and Alaska Natives on Reservations in Census 2000 was 4.74 percent, a reduction of about 60 percent from the 12.22 percent published for 1990. For American Indians and Alaska Natives off Reservations, Native Hawaiians and Other Pacific Islanders, and Non-Hispanic Asians, Census 2000 showed undercount rates of 3.28 percent, 4.60 percent, and 0.96 percent, respectively. The undercount rate for renters has dropped from 4.51 percent to 2.75 percent and for children has been reduced by about half from 3.18 percent to 1.54 percent.

THE QUESTION OF STATISTICAL ADJUSTMENT

Throughout the planning for Census 2000, a major issue of concern to the Census Bureau was whether the results of the A.C.E. could be used to make the census counts more accurate. In June 2000, the Census Bureau Director preliminarily decided that using the A.C.E. for this purpose was generally feasible, but to reach a final decision, it would be necessary to consider operational data to validate the successful conduct of the A.C.E., to assess whether the A.C.E. measurements of undercount are consistent with historical patterns of undercount and independent demographic analysis benchmarks, and to review measures of quality.[1] The Bureau has long used demographic analysis as an independent check on the quality of the count. Unlike the A.C.E., which is a sample survey, demographic analysis uses records and estimates of births, deaths, legal immigration, and Medicare enrollments, and estimates of emigration and net undocumented immigration to estimate the national population, separately from the census.

A team of Census Bureau professionals—called the Executive Steering Committee for A.C.E. Policy or ESCAP-was formed to conduct the evaluation to determine whether using the A.C.E. to adjust the census figures would improve the results for use in redistricting. After extensive meetings and staff work and the review of many

[1] Accuracy and Coverage Evaluation: Statement on the Feasibility of Using Statistical Methods to Improve the Accuracy of Census 2000.

analytic reports, the ESCAP completed its report[2], and Acting Director Barron submitted that report and recommendation, along with his recommendation, to me on March 1, 2001.

As a member of the ESCAP and as Acting Director, Mr. Barron concurred with and approved the ESCAP's recommendation that unadjusted census data be released as the Census Bureau's official redistricting data. The ESCAP reached its recommendation because it was unable, based on the data and other information it had at the time, to conclude that the adjusted data were more accurate for use in redistricting.

The ESCAP found that both the census and the A.C.E. were of very high quality. The primary reason for arriving at its conclusion that unadjusted data should be released was the apparent inconsistency between A.C.E. and demographic analysis. The demographic analysis estimates are significantly lower than both Census 2000 and the A.C.E. estimates for important population groups. The ESCAP investigated this inconsistency extensively, but in the time available could not adequately explain it. The ESCAP noted that the inconsistency between the demographic analysis estimates and the A.C.E. estimates is most likely the result of one or more of three scenarios:

• First, that the 1990 census and the associated coverage measurement methodologies together undercounted the population by a significantly greater amount and degree than previously believed, but that Census 2000 included portions of this previously unenumerated population.

• Second, that demographic analysis may not have accounted for the full population growth between 1990 and 2000.

• Third, that Census 2000, as corrected by the A.C.E., overestimates the Nation's population.

The Census Bureau must further investigate these concerns before it can recommend that adjustment would improve accuracy of data for purposes other than redistricting. It is also investigating other potential errors that could affect the accuracy of the adjusted numbers. All of these issues are discussed in detail in the ESCAP's report, which we are making available for the record.

After receiving the Census Bureau's recommendation, I thoroughly reviewed the ESCAP's report and supporting materials, and I obtained advice from a diverse group of prominent, non-government statisticians and demographers, in addition to the advice of the Census Bureau professionals. On March 7, I announced my decision to release the unadjusted data for use in the redistricting process. In making my decision, I followed a process that was transparent, reasonable and fair, and took full account of the view of career professionals and outside experts.

I should emphasize that ESCAP could not have resolved the critical questions about use of adjusted data prior to the April 1 deadline for completing release of redistricting data to the states, or even soon thereafter. I am confident that the Committee did all that it could, and that it reached the only reasonable conclusion.

THE ROAD AHEAD TO CENSUS 2010

As I have said many times, Census 2000 is the most accurate in our nation's history. But we cannot rest on our laurels. The Census Bureau has already begun looking toward 2010.

While Census 2000 was an operational success and produced data of high quality, the process was costly, many people felt burdened by having to answer the long form questions, and the census was constantly at risk due to insufficient early planning and development, and disagreement on the design. If the Census Bureau has adequate resources early to build upon the successes of Census 2000, then it can reduce operational risks for the 2010 census and explore ways to further reduce the undercount.

In a letter of January 17, 2001, from Mr. Chris Mihm of the General Accounting Office to my predecessor, Secretary of Commerce Norman Mineta, Mr. Mihm announced that Census 2000 had been removed from the GAO's list of high-risk Federal government programs. That Census 2000 was on this list is a reminder of the great challenges the Census Bureau faced and overcame in conducting a successful census. In his letter, Mr. Mihm stated:

> As the Bureau plans for the 2010 Census, it will be important for the Department of Commerce to ensure that the Bureau completes its evaluations of key

[2] Report of the Executive Steering Committee for Accuracy and Coverage Evaluation Policy: Recommendation Concerning the Methodology to be Used in Producing the Tabulations of Population Reported to States and Localities Pursuant to 13 U.S.C. 141(c).

census operations as planned, and in a timely manner, explores innovative options that could help ensure a cost-effective headcount in 2010.

Completing Census 2000 evaluations will shed further light on what worked well or did not work in this census. To build on the success of Census 2000, to reduce operational risk, and to reduce the undercount even further, the Census Bureau must improve the accuracy of its geographic database and Master Address File, eliminate the long form from the decennial census by collecting those data in the American Community Survey, and reengineer the census process through early planning. The improved geographic systems will ensure that there is a complete and unduplicated address list and will facilitate automation and electronic data collection.

In this regard, the American Community Survey will provide more frequent detailed data for small geographic areas and allow the Federal statistical system to keep pace with ever increasing demands for timely and relevant data. And it will revolutionize the way we take the decennial census by simplifying the 2010 census requirements and allowing the Census Bureau to focus exclusively on the basic count. However, early 2010 planning and development is necessary for a re-engineered process for the 2010 census, taking advantage of opportunities provided by having improved geographic systems and the American Community Survey.

As reflected in the President's budget, details of which will be released shortly, the Administration supports the Census Bureau's 2010 efforts. I look forward to working with Members of this Committee, other interested Committees and Members of Congress, to define and provide appropriate support for the total Census 2010 effort. We cannot delay, as every day brings us closer to what will be an even greater challenge to capture our increasingly diverse, vibrant population.

Mr. Chairman, that concludes my testimony. I will be pleased to answer any questions you may have.

Chart 1: 2000 Accuracy and Coverage Evaluation (A.C.E.)
Estimated Net Coverage by Race and Hispanic Origin

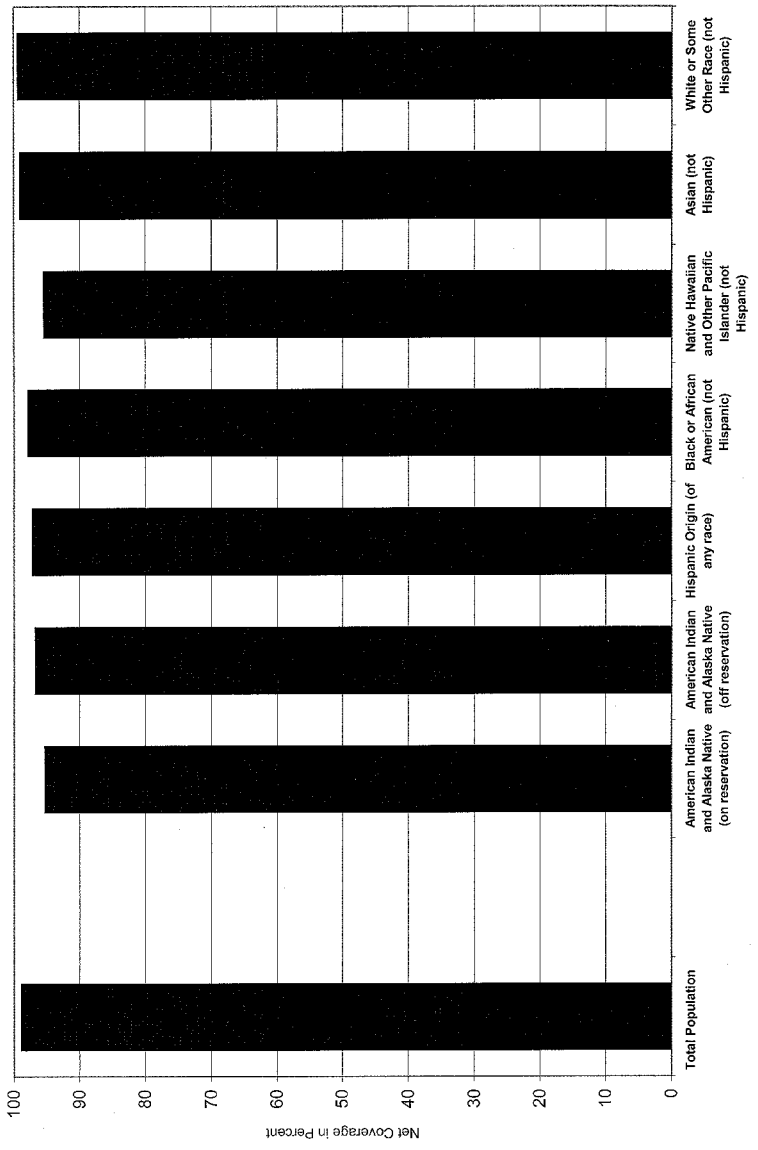

Chart 2.—Estimated Percent Net Undercount for Race and Hispanic Origin Groups: 2000 A.C.E. and 1990 PES

2000 A.C.E.			1990 PES		
Estimation grouping	Net Undercount (%)	Standard error (%)	Net Undercount (%)	Standard error (%)	Estimation grouping
Total Population in Households	1.18	0.13	1.61	0.20	Total population[1]
Race and Hispanic Origin:	.....	.....	.....	.....	Race and Hispanic Origin:
American Indian and Alaska Native (on reservation).	4.74	1.20	12.22	5.29	American Indian and Alaska Native (on reservation)
American Indian and Alaska Native (off reservation).	3.28	1.33	.....	.....	
Hispanic Origin (of any race)	2.85	0.38	4.99	0.82	Hispanic Origin[2]
Black or African American (not Hispanic).	2.17	0.35	4.57	0.55	Black or African American
Native Hawaiian and Other Pacific Islander (not Hispanic).	4.60	2.77	.....	.....	
Asian (not Hispanic)	0.96	0.64	2.36	1.39	Asian and Pacific Islander
White or Some Other Race (not Hispanic).	0.67	0.14	0.68	0.22	White or Some Other Race (not Hispanic)[3]

Notes: The race and Hispanic categories shown on the left side of this chart represent estimation groupings used in developing estimates based on the A.C.E. Survey and do not conform with race and Hispanic categories that will appear in the redistricting (P.L. 94–171) files and other Census 2000 data products. In developing the estimation groupings used to evaluate the coverage of Census 2000, the principal consideration was to combine people who were expected to have the same probability of being counted in Census 2000. Consequently, the race and Hispanic origin groupings used to create the A.C.E. estimates of coverage are exceedingly complex. For a complete description of the estimation groups, see DSSD Memorandum Q-37, which will be provided on request.

In general, American Indians and Alaska Natives (AIAN) are included in that category, regardless of whether they marked another race or are Hispanic. A few exceptions apply, especially for those who do not live on a reservation, on trust lands, or in an AIAN statistical area.

Similarly, Native Hawaiians and Other Pacific Islanders (NHPI) generally are included in that category, unless they lived outside of Hawaii and marked more than one race or marked Hispanic.

Hispanics are mostly in that category, unless they marked AIAN and lived on a reservation, on trust lands, or in an AIAN statistical area, or marked NHPI and lived in Hawaii.

People who marked Black or African American are generally in that category unless they fell in the categories described above; similarly those who marked Asian are generally in that category, unless they fell in the categories described above.

The final category includes most people who marked only White or only Some Other Race or marked three or more races but did not fall into the categories described above

The race and Hispanic categories shown on the right side of this chart represent selected population groupings used in conducting the PES and do not conform exactly with race and Hispanic tabulations that were released from the 1990 Census.

The data in this chart contain sampling and non-sampling error.

[1] Includes household population and some Group Quarters; excludes institutions, military group quarters

[2] Excludes Blacks or African Americans, Asian and Pacific Islanders, and American Indians on reservations.

[3] Includes American Indians off reservations.

Chart 3.—Estimated Percent Net Undercount for Age, Sex, and Tenure Groups: 2000 A.C.E. and 1990 PES

2000 A.C.E.			1990 PES		
Estimation Grouping (%)	Net Undercount (%)	Standard Error (%)	Net Undercount (%)	Standard Error (%)	Estimation Grouping
Total population in Households	1.18	0.13	1.61	0.20	Total Population[1]
Age and Sex:	.....	.....	.....	.....	Age and Sex:
Under 18 years	1.54	0.19	3.18	0.29	Under 18 years
18 to 29 years	.....	.....	.....	.....	18 to 29 years
Male	3.77	0.32	3.30	0.54	Male
Female	2.23	0.29	2.83	0.47	Female
30 to 49 years	.....	.....	.....	.....	30 to 49 years
Male	1.86	0.19	1.89	0.32	Male
Female	0.96	0.17	0.88	0.25	Female
50 years and over	.....	.....	.....	.....	50 years and over
Male	-0.25	0.18	-0.59	0.34	Male
Female	-0.79	0.17	-1.24	0.29	Female
Tenure:	.....	.....	.....	.....	Tenure:
In owner-occupied housing units	0.44	0.14	0.04	0.21	In owner-occupied housing units
In nonowner-occupied housing units.	2.75	0.26	4.51	0.43	In nonowner-occupied housing units

Notes: The data in this chart contain sampling and non-sampling error; a minus sign denotes a net overcount.
[1] Includes household population and some Group Quarters; excludes institutions, military group quarters.

Senator KERRY. Mr. Secretary, thank you very much. We appreciate the testimony. Your testimony, incidentally, I read completely last night just to get a sense of it and, indeed, you appropriately point out the degree to which the census this year improved in many ways. I mean, there was a very significant effort to do things better, and I certainly want to acknowledge that.

I don't think we should approach this in a contentious way that suggests there was not a bona fide effort to try to do the best census possible. The problem is that even those who engaged in the census, even those who undertook to do the best census possible, acknowledge, as I am sure you do, that there still is an overcount and an undercount. I mean, I assume you accept that even your current census is flawed to some degree. The question is, to what degree, is that correct?

Secretary EVANS. Right.

Senator KERRY. So we know there is some overcount, the overcount generally being, according to most historical analysis, individuals who are generally white and affluent, because it winds up being kids counted twice who are away at college, people who have second homes, and those are the kinds of situations that lead to double counting, whereas minorities, poor people, new immigrants, people in the inner cities, tend by and large to be those historically have been undercounted, and I assume you would agree with that.

Secretary EVANS. That is what the data seem to reflect.

Senator KERRY. Now, that said, Mr. Secretary, let me just kind of throw the sort of historical background of this on the table and ask you how we can proceed from here. This is a charged political issue, and I think the effort ought to be by all of us to try to defuse that tension, to uncharge it, but it has broken down largely along partisan lines. I mean, I just want to sort of lay it out on the table the way it is.

Democrats have tended to support the use of statistical sampling, because it appears to be a method of more accurately counting the minority population, those undercounted, and I guess if I were telling the truth about it, by and large, most people would say, "Well, those folks tend to be of sort of Democratic proclivity," whereas the people double-counted and the people on the upside who would just as soon see the formula not be as generous toward underserved areas have tended to be Republicans. Republicans oppose, Democrats have fought for it.

Now, this started in the Reagan-Bush administrations. The issue was precipitated by the Census Bureau decision of 1987 to use statistical sampling in the 1990 census in order to correct for the undercount, and that is when the A.C.E. was first supposed to be incorporated in the Census, but President Reagan's Department of Commerce overruled the bureau and attempted to cancel the funding for sampling. The funds were eventually restored, and plans for use of the method proceeded.

Then in 1991, President George Bush's administration appointed a Census Director, Barbara Bryant, with the support of a 7–2 decision by the Census board, recommended to then-Commerce Secretary Mosbacher that statistical sampling be used to correct the 1990 census. That recommendation was rejected, and the 1990 census was the first census to be less accurate, as a result, than the previous one.

The matter was then taken up by Congress, resulting in the 1991 passage of legislation directing the National Academy of Sciences and Census Bureau to design a plan with higher accuracy, and they came up with the extensive statistical sampling plan.

I might add that many Republicans, including Senator John McCain and Congressman Gingrich, strongly supported statistical sampling, and when the Republicans took over Congress after 1994, that is when you began to see a shift in their position.

Now, I say all of this as background, Mr. Secretary, simply to try to see if there is not some way in this new year for a new administration and this Congress to diffuse some of the politics of this, and to try to present America with a real bipartisan approach that has the confidence of everybody that it is not one side trying to protect its interest versus another, but rather, as Ranking Member Hollings said, we are trying to get the best count we can and the most available data that gives people the best confidence in the country, and I wonder if you might share with us, therefore, your sense of how you proceed from here.

While I am not going to argue with you that you received a recommendation from an Acting Director to proceed in a certain way, the rationale was that they felt they did not have time to resolve the issues, and I wonder if you would therefore now say to the Committee that you are prepared to give them the time, get this analysis and discrepancy between the data between the A.C.E. and the Demographic Analysis resolved so that we could see if we can find a bipartisan acceptable understanding of how we proceed from here.

Secretary EVANS. Senator, thanks for putting it in perspective, and your points are well-taken. We have the most accurate census in history, as I said in my remarks, vast improvement from the

1990 census, and I congratulate and salute all of those that were involved in the effort for their efforts. I look at—and as you acknowledged, I think, as I said in my remarks when I was here in January, I would look at this with an open mind, transparently, would listen to the bureau and what their recommendations would be, and that is what we did. We listened to the professionals, and the professionals made their recommendation to us and to me and I accepted that.

Senator KERRY. The recommendation is temporary.

Secretary EVANS. I have said, and I said in my testimony that they will continue to evaluate the data, and they will continue to make assessments as to whether or not statistical adjustment will make the data, could make the data more accurate.

My point was going to be that the lower the undercount is, the more difficult it is to say that you can statistically adjust the data to a more accurate level. But will we take more time to evaluate the data? Yes, we will. I have instructed Bill to do that, and he has indicated to me that it will be sometime this summer before they will be willing to be ready to come back to me with a recommendation as to whether or not the data should be statistically adjusted.

Senator KERRY. Well, let me try to put that in perspective. Number 1, you do agree that ESCAP did not decide that the uncorrected data is more accurate than the corrected data. That decision was not made, correct?

Secretary EVANS. Repeat that, Senator.

Senator KERRY. That the uncorrected data is more accurate than the corrected data. That decision has not been made.

Secretary EVANS. Well, what ESCAP said was that at the point that we made the decision, that they could not conclude that the A.C.E. data would be more accurate than the data.

Senator KERRY. Well, quoting from the ESCAP report, it says,

> "While the majority of the evidence indicates both a continued existence of a differential undercount of the population and the superior accuracy of the corrected numbers, the Committee has concerns, and until these concerns are fully investigated and addressed, we cannot recommend using it."

Secretary EVANS. They could not recommend to me that it would be more accurate to use the A.C.E. data at that point.

Senator KERRY. But at this moment we are, by everybody's understanding, living with an undercounted number of Americans. We know that. At this moment we know there are an undercounted number of Americans, and the question is, are we going to find a way to adjust for that, or are we prepared to live for the next 10 years knowing we have undercounted a number of Americans.

Secretary EVANS. Right, and that is what we are going to continue to work on and see if we can come up with a more accurate number.

Senator KERRY. Would you make the following commitment to the Committee: Would you make the commitment that we will have a resolution of whether or not your department will use corrected data for the purpose of allocating Federal funds? I mean, that is really what this fight is about, and the question is going to be whether or not we are going to have the opportunity for those Federal funds to be distributed in the most fair basis possible. I am not asking for an overcount.

Secretary EVANS. I understand.

Senator KERRY. No one here is asking you to overcount, but we simply know we cannot live with an undercount.

Now, can we have a commitment that this Committee will have a chance to review with you that decision before those Federal funds are allocated?

Secretary EVANS. Yes, you do.

Senator KERRY. And you would come back to us?

Secretary EVANS. Yes, I will.

Senator KERRY. I think that is very important. I appreciate that enormously. I mean, look, it is in all of our interests to do this as fairly as possible, and resolve the difference between the statistical results. There are some people arguing the Demographic Analysis may be in error, and the A.C.E. may, in fact, be more accurate, and the question is, if that is true, there may be a way to resolve that, and I suppose—my colleagues are here and I want to give them an opportunity also, but let me just ask you this. Would you be prepared to release the A.C.E. data down to the block level so that scientists across the country can help make this judgment in resolving any undetected problems?

Secretary EVANS. Let me turn to Bill and let him respond to that.

Mr. BARRON. Senator, we think at this juncture to release the data down to the block level would be misleading to folks. In fact, from the experts I have talked to, I do not believe they think they need data at the block level to make this decision, but if someone believes that to be the case we would be happy to entertain that request. We provide a lot of data to the National Academy of Sciences and others.

Senator KERRY. Well, I assume someone could FOIA it.

Mr. BARRON. I suppose they could, Senator.

Senator KERRY. Would it not be better to do this in a cooperative, open way? I mean, if the data in fact supports the conclusion that the quality was generally good, I would think you would want that to be thoroughly analyzed and support your conclusion.

Mr. BARRON. Block-level data are notoriously noisy data sets. Throughout most of this process I have had to defend every block-level adjustment, and we have had to say at that level the A.C.E. does not improve data.

You have to start aggregating data above blocks to see the improvement that dual system estimation, or sampling, as it has become known popularly, brings to the process, so by emphasizing block-level data we are really taking the worst part of the whole process and putting it out on the street, and at this juncture, since we have no confidence in those numbers, we do not want to do that.

Senator KERRY. Well, let me come back to that. I want to be fair to my colleagues.

Senator Brownback.

STATEMENT OF HON. SAM BROWNBACK,
U.S. SENATOR FROM KANSAS

Senator BROWNBACK. Thank you very much, Senator Kerry.

Mr. Secretary, thank you for being here. First, thanks for the great picture that you are showing America. It is on a regular basis the people who are reviewing this data, and it really helps us to look and see who we are today, and it provides the data and the background. I think that is a very valuable thing. We can discuss political issues here, and those are important to discuss, but the picture you are showing to America is far more important, what the people are seeing, and I think you have done a very nice job of that.

One thing I want to ask you about in looking forward on the census and looking at the next census, and on your American community survey, was this year, this census for the first time since I think we have conducted, or at least since 1880, we did not ask on the short form about marital status, the first time since 1880 that we have not asked that, and the reason I want to draw to your attention is that family structure has such an important impact on what takes place to our children, what environment our children are working in.

This is really very important and useful information for people to be able to have, to study, and to look forward to what is going to happen to our families and what is going to happen to our children. I am hopeful that as you look forward to the 2010 survey on the short form you will reinstate that marital question. There was a vote in the Senate last year unanimously in favor of doing that and reinstating that marital status question, and I am hopeful that you will do that.

Any thoughts? Have you had a chance to look at that?

Mr. BARRON. Senator, I was not at the Census Bureau when the final decisions on the short form were made, but certainly what you are asking for on a go-forward basis is exactly what we intend to do. We will be working with the Congress, and I think the Office of Management and Budget is coordinating that, but we have heard you, and we will be working with the Congress to decide what the questionnaire is going to look like for 2010 and in the A.C.E.

Senator BROWNBACK. Good. I hope you will look at that issue, because it really does have a huge impact. The second item is on your American community survey, which as I understand it is an ongoing, almost rolling survey that continues to be working. I have been particularly concerned that the census long form did not do an adequate job of gathering the family structure in the country, such that we have lost valuable data on what our families look like across the country.

I want to draw your attention—I would ask you again here what steps are being taken to ensure the American communities survey improves and enhances data collection on family structure in America, and I would not be so pointed on this, except for this impact on what is happening to our children and what happens on our school system and what happens in crime, within all of the various things that impact our society and impact our Government. We need that information in a solid form.

Any thoughts on what you are looking at on the family data either, Mr. Secretary, or Mr. Barron?

Mr. BARRON. The answer would be the same, Senator. In the 2000 census there was a very narrowly applied rule that if it was not explicitly required by law, or to support a legal purpose, then the questions were not asked, but this time we will be happy to work with the Congress to see if a broader approach could be taken, and we will obviously keep your thoughts in mind as we go through that.

Senator BROWNBACK. If you could. Senator Moynihan, before he left the Senate, and I worked closely on statistical gathering, and its impact on the country, and his point was that until we figured out a way to measure economic data like unemployment, inflation rates—we didn't know how to change things.

To be able to change something you have go to be able to measure it, and probably one of the most valuable things we can do at the Federal level is to be able to provide accurate measurements, whether it is on economic factors or social factors, and you probably have the lead tool for us on social factors of what is taking place across the country, and for us to be able to move in a positive direction socially we have got to have better numbers here and be able to provide it on a state-by-state, major geographic basis, and then let communities react to those numbers.

That is probably the most valuable thing we could do, and he and I were very strong on pushing that point. Thank you. I know we are under a time line, and we are under a vote, so I would yield to my colleague from California, or the chairman, and I will look forward to working with you on these two issues.

Senator KERRY. I am going to go vote, and I will be right back, Mr. Secretary.

Senator BOXER. Senator, do you want to put us in a recess?

Senator KERRY. If necessary.

Senator BOXER. Well, I do not want to miss the vote.

Mr. Secretary, welcome. It is nice to see you, and nice to see you, Mr. Barron, and I look forward to working with you on this and many other issues. I will be brief, because we do have this vote, and Senator Brownback, please do not feel obliged to listen if you have to leave. It may not be your favorite point of view here.

Well, my overriding principle has been from day 1 that we need the most accurate count, and I know we all feel that way. I have reasons to feel that way that go beyond just the fairness questions to real pragmatics, which is what happens to California when there is an undercount. With a population of 34 million people, we get the worst of it. We usually do.

I will explain what happened in the last census and how many dollars were shorted to my state for children, old people, transportation, and all the things we need. I do want to commend the bureau for producing a more accurate census in 2000 than was produced in 1990. However, I do share the concerns of many of my colleagues, Congresswoman Maloney and others, that the accuracy in coverage evaluation (A.C.E.) used to measure the accuracy of the raw counts shows there still remains a significant differential undercount for our minority groups compared to others.

That means that in my state, which has large minority populations, many urban communities will not receive the Federal funds they need desperately for basic programs, and their residents

will lack the proportionate representation that is their basic right under the Constitution.

The Census Bureau has admitted that in the 2000 census at least 6.4 million people were missed, mostly in the cities, and 3.1 million others were double-counted, mostly in the suburbs. According to a report in the New York Times, those numbers could be higher, with the bureau having missed as many as 7.6 million people, and double-counting 4.3 million others.

Now, for my state, that is just not an acceptable result. As a result of the 1990 undercount, California lost nearly $214 million in Federal funds, and the city of L.A. alone lost $120 million in Federal and state funds. Millions of dollars are at stake for educating our children, for title I, for health services, for the poor through Medicaid, and for a multitude of other services like child care and day care that allow parents to go to work and have peace of mind.

For the 2000 census, $400 million was spent on improving the A.C.E. and $700 million spent on improving the raw count. That is a lot of money, and Iam troubled by the fact, Mr. Barron, that you are not interested in releasing these numbers. They are not the secret property of the Census Bureau. They belong to the American people. The American people spent a lot of money, and I think these numbers belong to the American people. I would like to see you reconsider your decision. I think it is going to happen anyway. It is a democracy. It is going to happen.

Second, I would love to see the Census Bureau move with immediate haste to resolve the issues that have distorted the numbers and provide Congress with adjusted numbers well before the fall deadline for determining the distribution of Federal funds. As I say, these are not luxury programs. They are very necessary services to people who are trying to grab the American dream. They need these funds.

I certainly look forward to working with Secretary Evans. He and I have a nice working relationship. I want to keep it that way, and with you, Director Barron, and with my colleagues in Congress, and with organizations like the National Association of Latino Elected and Appointed Officials, the Mexican American Legal Defense Fund, so that we can reach a consensus on an adjustment to the census that reflects an accurate count.

I have questions. What I will do is place them in the record, because immediately following the vote, I am due at a conference on the environment. I have a few disagreements on little things like arsenic and lead, so I will not be returning. However, I will be submitting these questions for the record, and again, I make these requests in friendship, in a spirit of cooperation.

I think we still have time to work together to put everything out on the table and try to do our best to get an accurate count. Our people deserve it. I think you know that. USA Today reports that 350,000 people in Texas could have been undercounted. In New York, 300,000 may have been ubdercounted; in California a ½ million. These are not little side bars. These are living human beings who, for the most part, deserve the services that we provide.

So thank you very, very much, and we will stand in recess until Senator Kerry returns, which should be any minute now. Thank you very much.

[Recess.]

Senator KERRY. The Committee will come back to order, please.

Mr. Secretary, thank you. You have been very kind to hang in there. I am not going to keep you more than just a couple of minutes, because I know you are already late for your meeting.

Just one question. Mr. Secretary, I think we spent something like $7 billion on the census itself, and some $400 million was spent on the accuracy coverage evaluation. What does the bureau allocate for the Demographic Analysis model?

Secretary EVANS. I do not know what that number is, do you?

Mr. BARRON. I do not. It is clearly smaller.

Senator KERRY. It is considerably smaller.

Mr. BARRON. I think it is $250 million for the A.C.E.

Senator KERRY. Not that the accuracy is obviously measured by the amount of money we spend, is it not? I mean, is there some reason to believe that, in fact, we may have a more accurate A.C.E. model now than we do Demographic Analysis at this point? Is that conceivable?

Mr. BARRON. I think the value of Demographic Analysis, Senator, in part is that it is an independent check on the quality of the A.C.E. It is not a system that has changed a lot over time, but except for one very important omission that I can talk about, the base system of adding births and deaths is, we think, a pretty solid system. There is an issue with exactly what the solution is to improve that in terms of money. We are looking at that, but I do not know that on a go-forward basis we are going to have a proposal that we are seeing, if money would improve this problem. We are not sure it would.

Senator KERRY. Do you see any problems in terms of the timeframe we are looking at before we have to make judgments about the distribution of funds that would be able to come to a public conclusion that can be sort of put to the test with respect to trying to resolve the discrepancies between data sets?

Mr. BARRON. Senator, I think unlike the period we just went through, where in essence we had to make a decision by March 1, I think that by fall we are going to be able to go to the Secretary with an analysis of the issues we encountered, not just on Demographic Analysis, but also with the accuracy and coverage evaluation survey, and make a recommendation. I am confident we are going to be able to do that. If not, then I am going to have to come up and tell you that I am not, and why I am not.

Senator KERRY. Well, Mr. Secretary, can I make a suggestion that is made in completely good faith to try to resolve this? I can understand why you made the decision you made, given the input you had, given the time issue and you obviously had pressure on you with respect to redistricting, needless to say, you have got to make a decision. But on the other hand, I would hope you could also understand the measure of concern and even skepticism that a lot of people have about what may happen now, as we proceed down the road.

It seems to me it is very much in your interest—I mean, to look at the California potential lawsuit now, you look at the questions raised, it just does not serve any of us well to have the country at

odds over the methodology by which we determine how many folks are living here.

There ought to be some way of people acting in good faith to come up with an agreement that the data is trustworthy, that the methodology is trustworthy, and all we are talking about is a fair count. I can assure you there is no way for me to be advantaged or for the folks you are arguing for to be, quote, advantaged in this because we are not going to get an overcount.

The only issue is, will the overcount issue be addressed in a way that gives confidence to people this has been done fairly for the purpose of the next 10-years distribution. I would think you would want it to be done to the greatest capacity possible, and I sense you do, but I just want to encourage you to consider the notion that if the data is out there, subject to analysis, it will be subject to different people's analysis. It is better, probably, to suffer that than to have people saying you are hiding something, or you are unwilling to submit it to scrutiny, so I would strongly urge you to let folks kind of pore over it.

I do not think that is going to resolve the issue completely, but hopefully we can work together, Mr. Secretary, in the next months to come to some conclusion about how we can avoid any controversy over this.

Secretary EVANS. Senator, we will give that consideration and give it some thought. We share the same ultimate goal. We both want the American people to believe that the data that we have released are fair and are accurate. But there is another component to just the count, and that is the location, where these people live. When you get into the statistical analysis and that portion of adjustment, it gets complicated and complex, and those are some of the issues that I know that the professionals struggle with.

It is not only, having to say how many people, but you have to determine where they live, all the way down to the block level. So know that you and I share the same goal. We both want the American people to have the fairest and most accurate data that we can provide them, and that is what I intend to do.

Senator KERRY. Well, I appreciate that commitment, and obviously, I mean, this is something people have struggled with for long periods of time. I think the statistical sampling has become remarkably sophisticated. It is subject to huge amounts of scientific statistical analysis and tends to be objected to, frankly, on the partisan lines I have described That is its own message, and I think we need to be particularly sensitive to that message as we go forward here.

We have the scientific know-how about how to make up for this undercount without at all abusing the other side of the coin, if you will. There are ways even to weight that statistical sampling to guarantee that you don't abuse it, so it seems to me that the test here is the test of good faith, and the effort to try to not live another 10 years with an undercount that everybody acknowledges exists.

There must be some way to make up that undercount fairly and thoughtfully. That said, Mr. Secretary, we have taken more of your time than you had originally agreed to, and I am very, very appre-

ciative for your doing so, and I apologize to whoever it is you are keeping waiting.

Thank you very much.

Secretary EVANS. Thank you, Senator, very much.

Senator KERRY. What we will do is go to the third panel, and we will just announce ahead of time we will interrupt that panel the minute our Members of Congress come back. They have a vote, and obviously have a distance to travel, so if we could invite Dr. Ericksen, Dr. Murray, Mr. Vargas, and Dr. Wachter, we will commence, and then we will interrupt, as I said, at the appropriate time.

Gentlemen, thank you very much for coming. I appreciate your participation in this hearing very much. I notice the hearing room half-emptied as the Secretary left with members of the Commerce Department. I do not know if we had the whole Census Bureau here or what.

Dr. Murray, why don't we begin with you, and we will just run right across the table, and I appreciate your being here.

STATEMENT OF DR. DAVID W. MURRAY, DIRECTOR, STATISTICAL ASSESSMENT SERVICE AND CENSUS MONITORING BOARD, CONGRESSIONALLY APPOINTED MEMBER

Dr. MURRAY. Thank you, Senator. I am Dr. David Murray. It is a pleasure and honor to be in front of you, Senator Kerry and Members of the Committee. I am honored to appear before you to discuss the 2000 decennial census.

I serve as director of the Statistical Assessment Service here in town. I am also here in my capacity as a congressionally appointed member of the U.S. Census Monitoring Board. As you know, Congress created this bipartisan panel to observe and report on preparation and implementation of Census 2000, and I have written testimony which I will submit, and I also request one thing, Senator. In my written testimony, I alluded to a report that our monitoring board put out. I brought a copy of that, and if that is acceptable I would like to also include that as part of my testimony.

Senator KERRY. Absolutely. We will make the full report a part of the record.

Dr. MURRAY. Coming to grips with technical issues in the census is obviously a challenge to all of us, and I think in particular for Senators who have to quickly come up to speed on the issue. I tried to prepare in my mind an analogy as to where we stand, and I hope this is a reasonable and homiletic expression of what our difficulty is.

Every 10 years, as it were, the country has to balance its checkbook.

Senator KERRY. Pull the mike up a little bit.

Dr. MURRAY. Every 10 years, sir, the country has to balance its checkbook, has to bring its accounts into rectification. It has three ways of measuring and looking at what our balance might be, and we have to do that with the census. Our fundamental problem is, we cannot balance our checkbook this time. The three different measures we have give us a different set of results.

Now, we can count our dollars. That's the enumeration. That gives us a sense. We can go back and look at the bank records and our check register. That is what the Demographic Analysis (DA) is roughly comparable to. It gives us a very different picture of what our true balance is. Or we can take a sample of about 3/10ths of 1 percent of the households of our account and try to correct it based on that.

We have three fundamentally different measures that in some measure all of them have strengths and weaknesses. Each one has a weak point. Each one has relative strengths, but they differ from each other in a fundamental and inconsistent way that we found could not reconcile, and as such we do not necessarily have one number that is better than another.

We do not necessarily know that we can correct the census by using, for instance, the A.C.E. What we have is different numbers. The A.C.E. gives us a different number than the enumeration does. The DA gives us a different number than the enumeration does. Given that circumstance it seemed the most prudent course of action was to stick with the one, the devil we knew, rather than the devil that we don't, and it was a prudent and wise decision to go with the enumerated numbers as the most accurate that occasion the fewest number of costs to the system.

In some sense the A.C.E. began to be problematic, as is the DA, and the sense that the cure may have been worse than the disease, and let me just say a couple of formal things here. Where do we stand with an estimated undercount of only 1.18 percent? That is, we have a census that at the moment is 98.82 percent accurate. We should realize that the cost-benefit ratio of our respective choices begins to shift.

Given all the attendant legal and political difficulties that the A.C.E. engenders, its saving grace was that it might be a good technical fix, but now that is in question. The A.C.E. appears to have levels of statistical noise in its probability fluctuations that are greater in magnitude than this signal it was designed to detect and correct, that is the 1.18 percent undercount is smaller in magnitude than the error range that we use out of the statistical sample. The probability fluctuation is greater than that.

Applying the adjustment to a census as accurate as the one we have just completed begins to slide down the slope of diminishing returns, technically, as well as in terms of governance consequences. We face a genuine dilemma which the ESCAP report issued at the end of February well captured. The actual enumerated Census 2000 count placed us in a new and perplexing landscape, one unanticipated by the designers of the A.C.E.

The count of 281.4 million surpassed the best reckoning of the population given by the Demographic Analysis by nearly 2 million people. This is an anomalous outcome, since the DA has traditionally stood above the enumerated count and told us about the magnitude of the undercount in the enumeration.

Now, we have broken through the DA measurement. By the reckoning of the DA we actually have an overcount as the measurement that we get from the DA and the enumerated census. Perhaps, on the other hand, the enumerated census was, in fact, a

huge achievement, diminishing the undercount substantially, so much so that it is the most accurate number we have.

Various scenarios to explain the irreconcilability have been proposed by the ESCAP report and by others, but no scenario has achieved a reconciliation without introducing yet other anomalies. We are somewhat like the person who has a sheet over us. We pull the sheet up to cover our chin, and find our feet exposed. We pull the sheet down to cover our feet, we find our chin exposed. Every statistical adjustment that we try to bring to bear produces new anomalies and new problems.

The emerging and tentative adjustment from the A.C.E. took us even further away from solid ground, with an estimated result for approximately 284.7 million. The gap with the enumeration count is substantial, while the gap between A.C.E. and DA is fundamentally problematic, too large, in fact, to be reconciled by any scenario yet deployed. Moreover, the A.C.E. itself contains internal anomalies that are difficult to engage with.

A.C.E. may be a valuable tool that tells us things about the undercount, that tells us things about the quality of the census, but the A.C.E. data cannot be shown to be more accurate than the numbers which they would replace, the enumerated numbers. The fundamental requirement of the A.C.E. was that it be superior to the data which they would supplant. The deepest fear is that using the A.C.E. numbers introduces error and irreconcilability, particularly at the local level, where apportionment decisions must be made in redistricting.

As such, it may actually be a cure worse than the disease, producing more problems and more anomalies than we already have, hence the prudent idea that we go with the unadjusted data as the best course of action, and learn the lesson that basically an intensified census at the local level has a stronger chance to correct the fundamental problem we address, the undercount.

We all care about the undercounted. We want to attain incorporating them into the system. It looks as if the best alternative way of doing that is to intensify and strengthen the census count itself, and not resort to an unproven method of probability that introduces more problems than it solves.

Thank you.

[The prepared statement of Dr. Murray follows:]

PREPARED STATEMENT OF DR. DAVID W. MURRAY, DIRECTOR, STATISTICAL ASSESSMENT SERVICE AND CENSUS MONITORING BOARD, CONGRESSIONALLY APPOINTED MEMBER

Mr. Chairman, Senator Hollings, Members of the Committee, I am honored to appear before you today to discuss the 2000 decennial census. I serve as the Director of the Statistical Assessment Service and I am also here in my capacity as a Congressionally Appointed Member of the U.S. Census Monitoring Board. As you know, Congress created this bipartisan panel to observe and report on the preparation and implementation of Census 2000.

Rarely has a policy dispute generated a ratio of heat-to-light greater than did the just-concluded census dispute over the use of statistically adjusted data for apportionment of political power. Because the matter is dauntingly technical, the press faced a serious challenge in telling the story, no matter how many explanatory data charts and expert consultations were available. Nevertheless, even given the difficulty of the task, the media's performance over the last three years was (except for a couple of stories from the *Washington Post* and the *Los Angeles Times*) uniformly disappointing.

Most succumbed to the temptation to cast the story as purely political, with potential winners and losers resorting to raw clout as they disdained scientific accuracy. In general, those who favored adjusted numbers were characterized as seeking to "count every American" (and incidentally aid their party's representation), even though their proposal was in fact to estimate, rather than actually count, missing people. Alternately, those who expressed doubts were cast as opponents of "modern scientific methods" seeking to preserve political advantage by deliberately ignoring missing people, principally members of minority communities and children.

As the *Houston Chronicle* (Feb. 22) editorialized, "Some ideologues oppose correcting the numbers . . . The opponents of statistical analysis are mistaken, of course . . . Those who oppose adjustments . . . either do not understand arithmetic, or they understand it all too well." The "whose ox is gored" story line built around putative political motives nearly always won out over real engagement with the technical complexities. As a consequence, numerous myths about the census adjustment process and its supposed consequences were introduced into the media bloodstream.

In the resultant morality play, adjustment advocates usually came off as earnest advocates for the poor, who could be aided by a simple application of statistical justice. Those who favored an enumerated count, on the other hand, were often cast as stubbornly refusing to use a readily available technical means to solve a social problem—"correcting" the undercount by statistics. Lost in the fracas were genuine arguments about the feasibility and advisability of supplanting the standard enumeration with these technical means—a position ultimately validated not only by the Supreme Court decision of January, 1999, but as well on February 28, 2001 by the decision of the Census Bureau itself. The enumerated count prevailed for good technical, not political, reasons.

At first glance, the undercount problem should have a simple solution. In the 1990 census, the net undercount was roughly 4 million people, about 1.6 percent of a population of 248 million. That is, 98.4 percent were properly enumerated (in contrast, the 2000 census missed approximately 3 million, which represented just 1.18 percent of 281 million people—a 25 percent improvement). We knew about the undercount because we could compare the enumeration to higher figures from Demographic Analysis, which were regarded as more accurate. We could have saved a lot of money (the 2000 census cost over $6 billion, much of which goes to finding that last percent) by simply adding a 1.6 percent "correction" to the overall population and calling it quits. But the census, unlike other government data, needs to know two things about Americans—how many in the aggregate, and also, how many in geographical (and demographic) distribution, in the smallest geography/detail.

That is, accuracy means not only getting the total count right, but positioning people where they actually reside, so that apportionment of political power can be congruent with their actual presence. And now the social problem gets tricky, because the undercount (either the1990 1.6 percent or the 2000 1.18 percent) is not evenly distributed geographically (it tends to cluster in a handful of counties nationwide—mostly dense urban ones), nor is it evenly distributed demographically. This becomes the heart of the challenge.

In general, the likelihood of being undercounted is thought to be related to being identified in various racial/ethnic groups (among other factors, such as home ownership). Members of minority communities are more at risk for undercount, other things being equal, than are non-Hispanic Whites. This means that the 1990 undercount of 1.6 percent was actually composed of a 0.7 percent undercount rate for non-Hispanic Whites, a 4.6 percent rate for Blacks, a 5.0 percent rate for Hispanics, all the way up to an estimated 12.2 percent rate for American Indians on reservations. Because there are legal triggers involved in these disproportions—there has to be a differential adjustment distributed proportionally, not a uniform add-back. Doing this correctly, so that one actually improves accuracy rather than introducing more problems into the count, is an enormous mathematical challenge.

Here we encounter the first media-generated myth—that the statistical adjustment was based on a proper sampling methodology, like we find in political poll taking, which could then be used to "correct" the undercount. This is only partly true. The actual process of determining who was likely missed in the census derives from wildlife biology, where it is known as a "capture/re-capture" form of "dual system estimation" (DSE). Want to know the number (and the species proportions) of fish in a lake? One could drain the lake and count the bodies, but a more viable process is to cast a net, capture and count a sample of the fish, and then tag them. After the fish are released, one makes another cast and re-captures some of the tagged fish in another sample. By comparing the two catches, we can figure out the ratios of those caught in this "dual" system and make good estimates about the real population of the lake.

Of course, the system isn't perfect. Some types of fish are likely to be missed in both the first and the second net cast. These are termed the "wily trout" about which we can only make indirect guesses, if we already suspect that they're "really there." So why do we suspect that they are really there in the population? Largely by comparing our enumeration to another measurement, vital statistics records such as birth and death certificates, which tend to give us a higher count of the population than those enumerated. These record provide what was introduced above, the Demographic Analysis (DA—about which more in a moment), which serves to indicate our likely shortfall. (Some have wondered why we do not just rely on a Demographic Analysis-type census in the first place, based on a variety of administrative records. The thought deserves consideration, but there are problems. For instance, the DA itself depends upon an estimate, since the number of those immigrating and emigrating must be modeled.)

Granting that there is an undercount, what is the best response? Some advocated intensifying the enumeration, trying to reach all quarters (or at least substantially reducing those missed). They have been substantially vindicated by the 2000 outcome, which saw the straightforward enumeration actually cut each of the differential undercounts by half or more—a genuine triumph. Others decided to try an experiment. The DSE methodology was thought sound enough that it could be incorporated into the census design, which would first take one cast of the net (the actual count, which is in reality a sample of the population, since we know some were missed), and then return to take another sample of 314,000 households, the denizens of which were "caught" again. This process was called the Accuracy and Coverage Evaluation (A.C.E.). By comparing the records for an address on the two captures, one can find correct "matches" (a person found both in the enumeration and the re-capture), find overcounts (those found in the enumeration, not found in the follow-up), and "find" undercounts (those appearing in the more intensive search of the A.C.E., but not recorded in the enumeration). So far, so good. Matches and mismatches form the basis for a statistical model of how to adjust the whole population, both upward and downward, for various groups.

But what about those "wily trout" that evade being caught in either net? That remains a genuine dilemma, the technical name for which is "correlation bias." There are people who are nearly impossible to reach no matter the methodology, and they make up an unknown proportion of the undercount. Basically, you cannot know what you cannot find. An attempted solution is to "model" those people based on those you did find who were likewise hard to count, such as using anomalies in the sex ratios of those found in certain demographic groups. For instance, if we assume women are easier to catch than the men who correspond to them in age and race, and we find proportionally more women in our counts, we can estimate the number of men who "should" be there as well. We can only hope that the sex ratios provide a good model of the unknown; there is no way to demonstrate it.

Measurement errors in either of the two "samples" are a real threat (as is the fact that some data are not based on actually touching someone's nose, but are derived from information provided by proxy—a neighbor or even projected onto a household from the characteristics of nearby community members with "similar" characteristics—this is known as "imputation"). After all, even the perfect plan is being implemented on the ground by an army of recently hired part-time census workers who are as prone to mistakes and fatigue as any of us.

Even when all goes well in the field the greatest problem is the matching process. Remember that we tagged the fish, presumably on the fin. This is not a popular thing for the American government to do to the people who happen to be residing here (the census counts citizens and anyone else present as well), no matter how efficient it could make the census. Moreover, fish rarely all on their own suddenly pull up stakes, as it were, and depart for another lake, nearby or across the country, without letting us know. Hence, we're never really sure that we're catching and matching the same fish when we make our DSE comparison, which we will then project onto the whole population. (In an earlier incarnation of the A.C.E. design attempted in 1990, a single mis-matched family of five led to nearly 45,000 people being erroneously added into the adjusted population. This problem and other very consequential mistakes in the earlier version—dubbed the PES, for Post Enumeration Survey—were only discovered two years after the 1990 census by a panel of expert reviewers. Fortunately, the PES adjustment was not applied.)

But let's put all those concerns aside for a moment, and presume that all went well in the measurement process. Have we really taken a sample with the A.C.E. that can then be used to correct the count? The answer is still, um, not exactly. There are several remaining steps to go. There is first the problem of selecting the 314,000 A.C.E. households. Every pollster knows that a proper sample, which will be used to extrapolate opinions "upward," as it were, onto the entire population (not,

as has been noted, what the A.C.E. actually does), must guard against being skewed or biased in its selection of participants. The best guard is to have a systematically random probability process for the selection. But this can not be done exactly for the A.C.E. design, since there are constraints on the sample that result in trade-offs. For instance, there must be some households allocated to every one of the 50 states, and further, we must ensure that the households are "distributed" in such a way that they represent demographic groups of interest.

When the households are selected, some weighting formulas have to be devised to make sure that their members adequately represent the groups in question in the overall sample. Skeptics of the process wondered just how much the selection of the A.C.E. households and the weightings applied could pre-shape the kind of answers that the sample was inclined to provide. (Because of correlation bias, we believe that the A.C.E. adjustment is prone to report that certain demographic groups were "undercounted" virtually no matter how good the initial count turned out to be. That is, the A.C.E. process may itself be biased to "discover" an undercount for certain types of people, perhaps even in conditions when the initial count itself was, in the aggregate, already too high.) At any rate, thinking that the sample selection and weighting was based on scientific grounds alone became a matter of (by all indications, properly granted) trust.

But the most important step was to divide up the 314,000 households into what were termed "post-strata." That is, the population in the sample was stratified and assigned into multiple cells that, according to the A.C.E.'s sociology, represented appropriate "types" of Americans. The appropriateness of a type was related to the probability that an individual in one of those cells would be missed in the enumerated census. The post-strata (think of the cells in an Excel spreadsheet) represented the intersection of variables like race/ethnicity, sex, age, and tenure (homeownership or not), the whole apparatus further divided by four regions of the country and by type of community (a range of larger-to-smaller metropolitan areas continuing to rural). As it turns out, only the largest post-stratified type, non-Hispanic Whites, were subdivided into the full set of post-strata distinctions. For most other types, the cells had to be conflated because of small absolute numbers, meaning that Asian Americans, for instance, were placed into two national cells (owner and non-owner) without regional breakdowns.

That is, not all groups were post-stratified by the same criteria. Proportionally larger demographic types (whites) could be subdivided more finely without seriously affecting data quality, while other smaller demographic groups had to be treated as broad bands across the whole country (that is, the data were national in the first instance, and no effort was made to subdivide them by finer-grained distinctions). This decision would have later consequences, such as being forced to "adjust" the population of one state based on data actually derived from nearby, or even relatively distant, states. (This fact led some adjustment critics to argue that the A.C.E. design seemed more consequential in shifting demographic shares within the population rather than prioritizing the need for accurate state-by-state counts of all demographic groups.) Other post-strata subdivisions, such as the number of age-group breakdowns or the degrees of community density, were likewise collapsed for some demographic groups where numbers were small, while for those under age 18, the male-female distinction itself was dropped. (Representative examples of post-strata would be non-Hispanic white male homeowner between the ages of 18–29 living in the northeast in a large metropolitan area; Hispanic female renter between 30 and 49 living in a rural area anywhere in the country).

While at first glance a sample of 314,000 households (close to a million people) is a huge number, providing reassurances about likely margins of sample error (the larger the sample, the smaller the likely probability spread, ideally), the actual population of each post-stratum cell becomes mathematically problematic. There were initially 448 post-strata in the A.C.E. (later conflated to 416). Hence, the 314,000 households divided by 416 post-strata actually yields only a little over 700 households per cell. That's not a reassuring number for sampling margin of error purposes, especially when we realize that the total number of households were not evenly divided among the post-strata. Because there are so few Americans in some of the assumed demographic "types," (example: Hawaiian or Pacific Islander female aged 50+ renting a house-trailer in rural Wyoming), the cells representing them, even when distributed regionally or nationally, are dangerously sparse.

Further, all of these purely quantitative concerns must be coupled with the apparent arbitrariness and uncertainty about the sociological assumptions underlying the choice of American "types." Were the assumptions actually legitimate models of the probability of being enumerated? We simply don't know. Overall, we must realize that the census represents the intersection of sophisticated quantification (assumptions about numbers) with real human beings (assumptions about which are, unfor-

tunately, anything but a sophisticated science). The A.C.E. design represents the place where two sets of very complicated models of the world derived from two very different disciplines interact, with any errors (in theory or in implementation) compounding each other. The results are then magnified by becoming the basis for adjusting the data on 281 million other people, the A.C.E. being considered the last word in accuracy, and hence, the benchmark standard for calibrating the entire US data collection system. Suffice it to say, the stakes are high for such a probability mechanism of unproven reliability.

And all that has been discussed above transpires before the results are released to the public, and ultimately encounters the requirements of the legal and constitutional system, fundamental provisions of which contradict the A.C.E. activity on the face of it. Finally, in a development beyond the scope of this discussion, we must remember the indeterminacy added to this census by the first-ever multiple-race selection, cross-cutting the whole system with 126 possible choices of racial/ethnic self-identification (which choices have themselves been acknowledged to be completely arbitrary governmental categories with no basis in scientific fact; moreover, the choices are unstable even in single individuals at different times).

What are the particular quantitative dangers of the post-strata? Demographers realize that they are caught in their own statistical version of Heisenberg's "uncertainty principle" when it comes to dividing samples into strata. You can pursue one piece of information, but only at the expense of its counterpart. The twin problems that must be balanced are "variance" and "homogeneity." Let us start with the second one. If we were devising a statistical model to subdivide inanimate objects, such as steel washers coming from an assembly line which we wanted to quality-check, our sample need not worry too much about homogeneity. We can look for variation in defects, let us say, while being reassured that most fundamentals would remain relatively constant (the washers wouldn't suddenly form into quartets and start singing, for instance).

With humans (and somewhat less so for fish), that is not so clear. Homogeneity is assumed whenever we expect a given cluster of people to react the same way to some variable (in this case, getting counted). The larger the group of people chosen, the less assured we are that they are reliably homogenous. Let us say we were interested in the likelihood of being missed in the census and we treated as alike all Hispanic females nationwide older than 29 but younger than 50. Unfortunately, we would be led to believe thereby that a migrant worker who did not finish high school living in a colonia in rural New Mexico is as likely to have been missed in the census as a Member of Congress living in suburban New York City. That is bad sociology (moreover, the thinking is suspiciously akin to what in other contexts is termed "racial profiling").

Rather obviously, the way to avoid over-homogenizing is to have the group to which the assumptions apply be fairly narrow. The smaller the cluster of people, the greater the likelihood that they genuinely share characteristics of importance. But now we are settling onto the other horn of the dilemma. Groups small enough to be reliably similar are also small enough to produce large variance (the statistical "spread" of the data) when their results are applied beyond the group. Hence, the design problem for the A.C.E.: develop sufficient post-strata that every cell is composed of reasonably homogeneous members, but do not make so many post-strata that small cells produce inherently unreliable sample data. Once again, a trade-off is faced, trying to optimize a response to the twin challenges.

As it turns out, the National Academy of Sciences, which was routinely characterized by the press as having "endorsed" census adjustment, in reality only agreed to the *principle* of statistical adjustment as quantitatively sound. Some members were never enthusiastic about some particulars of the actual A.C.E. plan (much less its field implementation), especially given that it was hurriedly developed in response to a 1999 Supreme Court decision ruling against a much more ambitious version of statistical adjustment, an effort to create a "one number census" based on an Integrated Coverage Measurement.

Let us grant for the moment that the A.C.E. design was adequate for our purposes (and you must not forget that our purposes include political apportionment and redistricting, as well as the proper distribution of federal funds, over and above the need to tabulate the aggregate numbers in the census). For all 416 (collapsed) post-strata, the matching process between the enumeration and the A.C.E. begins to tell us about which types were overcounted, which were undercounted, and which are "just right" (once again, this must be a somewhat simplified description; there are other complicated process to cause concern such as the unduplication of records or imputations).

Now comes the adjustment activity. Based on the "signal" derived from the A.C.E., we develop another weighting, regarded as a "correction factor," which we

export back into the total population count after it has been likewise stratified to match the A.C.E. types. (Again, critics argue over terminology. To term the factor a "correction" appears to prejudice the case that the result is somehow more accurate than the original number to which it is applied. Accordingly, it may be more valid to simply term the factor an "adjustment," acknowledging that the A.C.E. doesn't necessarily produce a "better" number, just a different one.) Adjustment factors can be positive (we are adjusting an undercount by using a number higher than one) or negative (we are adjusting an overcount by using a number less than one). We then multiply the count for each group in the enumeration by their respective adjustment factor, the product being what we record as their actual (adjusted) count.

This latter process led to some surprise when it was realized that the effect was to "delete" from the census actual people who had bothered to do their civic duty and fill out a form. Being understandably sensitive about appearances, the Bureau denies that anyone is deleted. They prefer to note that what happens is merely that a negative record is imputed to the census count, in effect nullifying the count of a real person that chose to participate. Whatever terminology we accept in this issue, a study of the 1990 PES identified no fewer than 1.48 million such "nullifications" based on overcount assumptions. Preliminary data from the Census 2000 A.C.E. indicate approximately one million such nullifications would take place.

At any rate, we are now at a point in our analysis where we can adjust the census statistically to "correct" the count. If a post-stratum has a positive number, such as 1.08, that means that we found more people in the A.C.E. survey in that stratum than the enumeration had recorded. Rather than 100 people, let us say, the A.C.E. is telling us there likely are 108 people. Hence, every time we find a record back in the enumeration national census for some other one in that post-stratum, we do not record just a one; we instead write down 1.08 for each one found. That means for every group of 100 people we find anywhere in the country who fit this profile, we "add" 8 more people of that type (where, exactly, do we put them? More in a minute . . .). Now we are rolling at last. For every 10,000 found, we write down 10,800. For every ten million, well, let us see here, the model tells us we have got 800,000 more people just like that, which we have to place somewhere on the map, even though we've never actually met nor counted them directly.

Moreover, some actual post-strata receive corrections that are hefty indeed. Hispanic males aged 30–49 rural non-owners in low enumeration districts, for instance, receive a correction factor of 1.19; that's nearly twenty percent, 120 for every 100, 60,000 for every 50,000. These are, in an important sense, virtual people, who must nevertheless be awarded their "fair share" of very real political power and funding (which are, by the way, zero-sum entities; if I give this finite resource to someone, it can only happen at the expense of someone else in direct proportion). It follows, of course, that for those receiving a negative correction factor, we write down for every one we encounter a number less than one. So for the presumably overcounted types, whenever we find them, we write down .92, for instance, and then add them together. For every 10 million of these losers, of course, we only record 9,200,000.

Even more remarkable, we have just engaged in a process that is not really "sampling" at all, but rather another (and less supportable) statistical maneuver known as "synthetic estimation." Recall that the adjustment consists of comparing one sample (the enumeration) with another sample (the A.C.E. population) and seeking matches. Based on the assumption that the A.C.E. results are always to be considered superior to the actual count (which may not be true, especially if the enumeration, which made a greater effort to activate local community outreach, was more successful at coverage of the recalcitrant than was the more "professional" A.C.E. re-contact, undertaken without the intensified community efforts), a set of "adjustment factors" are computed for each post-stratum. So far, so good, as far as statistical probability goes.

But then the adjusted numbers are applied to the entire national population with each post-stratum receiving its proportional adjustment higher or lower. The overall effect is a movement that goes in two directions. First the sample adjustments are adduced upward, as it were, to the national totals, and then brought back down, as it were, to the local level when the count is adjusted block-by-block.

It is this second movement back down from the aggregated total and distributed onto the smallest components of the population groupings that causes statistical concern. When we bring the totals from the national level back into the local aggregations we are engaged in what is no longer "sampling extrapolation" by any means, but rather a different maneuver—the "synthetic estimation." The fundamental (and contested) assumption behind the "synthetic" part is that because a certain proportion or ratio of a population can be asserted about a whole group (the US popu-

lation), therefore each distinct component of the aggregated whole likewise must mirror those proportions or ratios in equal manner.

But what is true of a statistical whole is not necessarily true of each individual component (statistics, after all, representing a summed average of many measures). Imagine for a moment that I discover a ratio of females to males at a university of thirty thousand students—females are 55 percent, males 45 percent. At the aggregate level, that is, the whole university, this can be accurate, without necessarily implying that each classroom in the university replicates this exact proportion. French classes, for instance, may not show the same ratios of female to male as chemistry courses, even though when taken together they "average" the overall ratio. To likewise expect every table in the cafeteria to exactly mirror the overall ratio quickly leads to absurdity—we should expect, under the principles of synthetic estimation, to find exactly 5.5 females and 4.5 males at every table of ten. Clearly something is wrong.

And yet this is just what the census adjustment process leads us to formulate. The adjustment factors for each post-stratum population found in the A.C.E. sample are "nationalized," as it were, and then applied down to the local level of neighborhoods, expecting the same ratios of under- and over-counted to apply at every level of the population hierarchy—state, county, congressional district, census tract, local block. As it turns out, the A.C.E. plan did, in fact, run into difficulty with this "synthetic" assumption, which further reflects the problems noted above in the discussion of assumed "homogeneity" of the post-stratum.

In actuality, we begin to see many difficulties with the operation of the A.C.E. conceptually, over and above those concerns linked to problems of measurement error and implementation issues. One of the central conceptual difficulties is that the statistical estimation, incorporating, as it must, a certain probability margin of error that is ineradicable (it being inherent in the operation of probability), only begins to "even out" its errors at certain levels of aggregation. That is, for the gross level of the total population (the aggregate count of the total number of the population, roughly 281.4 million persons), the probability errors (the inevitable pluses and minuses wavering from the actual target) do "average out." For instance, for every 100,000 measurement "pluses" that are too high there will also occur about 100,000 corresponding measurement "minuses" that cancel each other out. But a lower levels of aggregation (state, county, district) the possibility starts to magnify that they do not all "average out," and we may well be left with residual error—a less accurate count than we began with in the enumeration.

There is considerable dispute as to what level, exactly, we begin to lose ground with that adjustment, and actually start introducing error by adjusting. It may happen at the state or congressional district level, especially for selected demographic groups in the post-stratum (and perhaps worsened by the realization that we are using gross regional or even national data to adjust populations within a state—that is, we are not directly adjusting a state's population based on data derived only from that state). Whatever the eventual resolution of that dispute about the accuracy/inaccuracy threshold, everyone now agrees that at smaller levels of aggregation (in counties with less than 100,000 people it becomes clearly problematic), on down to the block level, we can no longer assure ourselves that the adjustment is superior to the unadjusted numbers, and we begin to seriously suspect that the adjustment is actually distorting our understanding by introducing error into the count.

Yet it is at the block level that politically important decisions must be made—such as the boundary of an electoral district. Moreover, the hoped-for randomness of the pluses and minuses canceling each other out is further belied by the practice of the re-districters, who tend to accumulate together blocks of people who share certain demographic characteristics, if for no other reason than their physical propinquity. Hence, if a particular demographic post-stratum is off within the margin of error in one consistent direction, there will be no balancing out of the error because they will be grouped together with similar blocks likewise erroneous in the same direction. The effect is to amplify the error in the redistricting result, rather than having randomness producing a canceling-out effect.

Problems abound. By virtue of the adjustment design, we have generated estimated people (virtual people) who have never been contacted nor identified, yet must be placed in some concrete location in an actual census block. The principle for assigning them a "local habitation and a name" is arbitrary and based on unproven assumptions. Yet their presence can have consequence in the apportioning of political power and funding. Moreover, the Congressional-appointed members of the Board further demonstrated in our report of September, 1999 that the effect of adjustment is to fail to position the undercounted correctly and proportionally in the communities where they were actually missed. By applying a "blanket" adjustment to every sector across the country, the adjustment gives the illusion of a remedy,

because the actual undercounted are not uniformly distributed across the country. Those communities that "lose" in the undercount do not receive a commensurate adjustment.

Further, it could be argued that the adjustment design, an effort to statistically "model" the population and then reformulate it, could have the effect of introducing more political features into the census than are found in the actual enumeration. Let it be noted that the "political" aspect of the census adjustment does not have to necessarily imply the active intervention of partisan concerns. As with budgetary or income tax battles, any process that is "assumption-dependent" is thereby open for political debate. Whoever sets the assumptions, or establishes the criteria for which factors are considered important (and in which order), can largely constrain the possible outcomes of the strictly quantitative process. All census activities are, of course, assumption-dependent, in this sense—witness the dispute between Utah and North Carolina over the allocation of the last House seat based on population. The issue hinged on whether or not overseas missionaries were assumed to be equivalent to overseas military in terms of their state assignment. It follows that every aspect of the census has political implications, in that it constructs political definitions and quantifies what are properly political entities—human beings in groups. Nevertheless, even given these caveats, census enumeration is relatively more assumption-independent than is the alternative—modeling the population for statistical adjustment purposes, where changed assumptions have the power to radically alter the entire nature of our national self-portrait.

A corollary of this reasoning is that the enumeration count will likely prove more accountable to democratic processes in the long run, as well. Witness the difficulties already encountered by policy makers and courts trying to understand and evaluate the highly technical nature of the adjustment's probability models. Who can truly grasp them and interrogate them but a very restricted group of technical experts? In this sense, an enumeration process, being relatively more transparent in its assumptions and enactment, may be not only more accessible and hence accountable but also more prudently consistent with the spirit of self-government.

Finally, where do we now stand? In the first place, with an estimated undercount of only 1.18 percent (that is, a census that is 98.82 percent accurate), we should realize that the cost/benefit ratio of our respective choices begins to shift. Given all of the attendant legal and political difficulties that the A.C.E. engenders, its saving grace was that it might be a good technical fix. But now that is in question. The A.C.E. appears to have levels of "statistical noise" in its probability fluctuations that are greater in magnitude than the "signal" it was designed to detect and correct (the 1.18 undercount is smaller than the margins of error range of the A.C.E. at certain levels of application). Applying an adjustment to a census as accurate (by all the evidence to date) as the one just completed begins to slide down the slope of diminishing returns, technically as well as in terms of governance consequences.

We face a genuine dilemma, which the ESCAP report issued at the end of February well captured. The actual enumerated count from Census 2000 placed us in a new and perplexing landscape, one unanticipated by the designers of the A.C.E. (and many other parties as well). The count of 281.4 million surpassed the best reckoning of the population provided by the Demographic Analysis (DA) by nearly 2 million people. This is an anomalous outcome, since the DA has traditionally stood above the enumerated count and told us the magnitude of the undercount in the enumeration. But now we have broken through the DA measurement. By the reckoning of the DA, we actually have an OVERCOUNT in the enumerated census. Perhaps, on the other hand, the enumerated census is right on the mark, having itself nearly eliminated the heretofore undercount. Various scenarios to explain this have already been proposed by the ESCAP report and others, but none has achieved a reconciliation without introducing yet other anomalies.

Moreover, the emerging and tentative adjustment count from the A.C.E. took us even further away from solid ground, with an estimated result of approximately 284.7 million. The gap with the enumeration count is substantial, while the gap between the A.C.E. and the DA (about 5 million) is fundamentally problematic—too large, in fact, to be reconciled by any scenario yet deployed. Moreover, the A.C.E. results themselves contain internal anomalies and inconsistencies, in addition to the incapacity to be reconciled with either of the two other measurements.

It should be apparent to anyone seriously engaged with this problem that while many specific details remain to be resolved, and further while the A.C.E. design has a valuable contribution to make in helping us understand what transpired in the census count, the inadequacies in concept and in practice preclude use of the numbers derived from the A.C.E. for the critical purposes of apportionment and redistricting. The A.C.E. methodology simply cannot met its primary obligation—being demonstrably more accurate than the data which they might supplant.

We must acknowledge the wisdom in the ESCAP recommendation to the acting Bureau Director (a recommendation accepted by Secretary Evans, and which moreover was endorsed by the previous Bureau Director Dr. Kenneth Prewitt) to regard the unadjusted numbers from the Census 2000 enumeration as the accurate numbers, that are the most appropriate for the Constitutional uses to which they are put. It was a decision made on the merits of the case as it was examined.

Let us not shun the larger lesson from this overall undertaking. The undercount is a genuine American difficulty, to which we need genuine solutions. No one should in principle be uncounted, and we must develop more effective remedies to ensure that the principle of the census is fulfilled. By all that we now know about the enumeration process, we should recognize a striking achievement, which was to reduce the differential undercount. The promise of that outcome is that we can close it yet more by intensifying the enumeration, by forming local partnerships to accomplish it, and by motivating people to find their way into full participation in the American system. "All politics is local," was wisely said. All censuses may likewise be local. Let us properly invest in what works best.

Thank You.

Senator EDWARDS. Thank you, Dr. Murray.

Mr. Vargas.

STATEMENT OF ARTURO VARGAS, EXECUTIVE DIRECTOR, NATIONAL DIRECTOR, NATIONAL ASSOCIATION OF LATINO ELECTED AND APPOINTED OFFICIALS (NALEO) EDUCATIONAL FUND

Mr. VARGAS. Thank you, Senator. My name is Arturo Vargas. I am the executive director of the National Association of Latino Elected and Appointed Officials Educational Fund, and thank you for the invitation to appear before you at this Senate hearing on the decision to release the adjusted census data for Census 2000.

The NALEO Educational Fund is the nation's leading organization that enables Latinos to participate fully in the political process. We are a 501(c)(3) nonprofit bipartisan organization with a bipartisan board of directors that took a strong position in support of the most accurate census possible, and in support of the Bureau's decision to use statistical sampling.

We were one of the organizations that worked with the Census Bureau through a partnership to promote a full count, and we were particularly proud of the mail-back response rates that we saw in Latino communities throughout the country. We commend the Census Bureau for the many elements of the Census 2000 which made it such an operational success, and we will be happy to share with this Committee at an appropriate time our views on the elements of the census that were particularly successful as well as the areas where we believe that we could have improvement for the 2010 census.

Now, as we all know, the preliminary estimates of the undercount by the bureau indicate that there was, in fact, a differential undercount, and that differential undercount does not fall equally among all Americans. Latinos, African Americans, Asians, children, immigrants bear the brunt of the undercount.

When the career statisticians at the bureau announced the recommendation against the release of the adjusted data, they based their decision on three different methodologies that have been discussed by the Secretary and the Acting Director. These professionals unequivocally concluded that there is considerable evidence to support the use of adjusted data. However, as has been indi-

cated, they were troubled by discrepancies between adjusted data and the results obtained by the Demographic Analysis.

Senator, Members of the Committee, I think the issue before us is, we must permit the bureau to take the time it needs and the resources it needs to resolve this issue. The ESCAP committee did not conclude that the A.C.E. data were less accurate than the Census. The ESCAP committee merely concluded that at the time the recommendation was required, they could not resolve the discrepancies. They did not say that the census data are, in fact, more accurate than the A.C.E. They did not say that the A.C.E. was more accurate than the census data. They indicated that at the time of the decision of their deadline, that they could not make that conclusion yet.

We should not be forced to live with a 10-year error because of a 3-month deadline. As we heard Secretary Evans express here today, the Census Bureau intends to continue its evaluations. I think it is the major responsibility of this Committee to support the professionals at the Census Bureau and ensure that they have all the resources and all the time necessary to complete the analysis of the Demographic Analysis, Census 2000, and the A.C.E. so that in the fall they can make this recommendation as to whether or not, in fact, Census 2000 can be corrected to compensate for the 3.3 million Americans who have been excluded from Census 2000.

Now, as an organization that works hard to ensure that Latino Americans are able to fully participate in our political process and in our society, the differential undercount has a strong, harmful impact on that goal of Hispanic Americans. It has an impact on our political representation, on the ability of institutions such as schools to provide appropriate services and classrooms, resources to children.

We cannot live with a 10-year error, Mr. Chairman, and I would ask the Committee at this time to hold accountable Secretary Evans for the commitment he made here today to this Committee that in the fall Census Bureau professionals will be recommending to him a full recommendation as to whether or not to adjust the census data.

Again, I think the issue here is not that the A.C.E. was less accurate, or that the census itself was more accurate. The issue here is that the ESCAP committee at the time they were required to make their recommendations, did not have the time sufficient to conclude one way or the other.

I think it is most important for the interests of this country that we allow the Census Bureau professionals to do its job with the full support of Congress, the full support of the administration, to ensure that when they make this new recommendation as to whether or not the A.C.E. is more accurate or not, that this country can be allowed to go forward for the next decade with the most accurate decade available to it.

Thank you, and I have summarized my comments here, and we have submitted our comments for the record.

[The prepared statement of Mr. Vargas follows:]

PREPARED STATEMENT OF ARTURO VARGAS, EXECUTIVE DIRECTOR, NATIONAL DIRECTOR, NATIONAL ASSOCIATION OF LATINO ELECTED AND APPOINTED OFFICIALS (NALEO) EDUCATIONAL FUND

Chairman McCain, Ranking Member Senator Hollings and Members of the Committee: I am Arturo Vargas, Executive Director of the National Association of Latino Elected and Appointed Officials (NALEO) Educational Fund. Thank you for the invitation to appear before you today on behalf of the NALEO Educational Fund to discuss the full impact on the Latino community of the recent decision by Commerce Secretary Don Evans to release Census 2000 data for redistricting that has not been adjusted to correct for the differential undercount.

The NALEO Educational Fund is the leading national organization that empowers Latinos to participate fully in the American political process, from citizenship to public service. The NALEO Educational Fund carries out this mission by developing and implementing programs that promote the integration of Latino immigrants into American society, developing future leaders among Latino youth, providing assistance and training to the nation's Latino elected and appointed officials; and by conducting research on issues important to the Latino population. The NALEO Educational Fund is a 501(c)(3) non-profit, non-partisan organization. Our constituency includes the more than 5,400 Latino elected and appointed officials nationwide.

As a member of the Commerce Secretary's Decennial Census Advisory Committee, I am pleased to be able to discuss with you the decision to release unadjusted Census 2000 data as the official data for the purposes of redistricting.

The NALEO Educational Fund is committed to ensuring that our nation will be able to rely on the most accurate data possible from the 2000 Census. Our organization, like hundreds of others across the country, mobilized to encourage all U.S. residents to answer the census. We are particularly proud of the mail back response rates in several Latino majority communities which demonstrated the sincere desire among millions of Latinos to make themselves count in 2000. We commend the Census Bureau for the many elements of Census 2000 which made it such an operational success, including its partnership program and commitment to work closely with community institutions, its high quality outreach and advertising program, and its efforts to hire an enumeration force that had the skills and capacity to carry out this monumental task. We would be happy to share with this Committee at another appropriate time our views on the elements of the census which were particularly successful and those areas in which we would recommend improvements for 2010. Our focus today, however, concerns the most basic element of the census, the accuracy of the data on which we will rely upon for an entire decade.

As we all now know, the preliminary estimates released from the Bureau indicate that the differential undercount was not eliminated. While the Census 2000 was an operational success, there was a net undercount of 3.3 million Americans. And many of those missed were Latinos—over one million. That, Mr. Chairman, is more than the entire state of Wyoming.

When the career statisticians at the Bureau initially announced their recommendation against release of the adjusted data, they based that decision on their examination of three different methodologies used to determine our nation's population: the traditional "headcount," the statistically-adjusted data based on the Accuracy and Coverage Evaluation (A.C.E.), and the Bureau's separate demographic analysis. These professionals unequivocally concluded that "there is considerable evidence to support the use of adjusted data;" however, they were troubled by discrepancies between the adjusted data and results obtained by the demographic analysis. They had to meet a deadline to make a recommendation regarding the release of the adjusted data, and they simply ran out of time to examine and explain those inconsistencies.

It is critical that we permit the Bureau to take the time it needs to resolve this issue. We should not be forced to live with a 10-year error because of a three-month deadline. If the Bureau determines that the adjusted numbers are more accurate, the Bureau should release them for redistricting and other purposes. The connection between redistricting and the Census goes back to the founding of our nation. The redistricting process plays a key role in ensuring that our democratic process provides fair representation for our nation's residents. The use of unadjusted data for this process will result in inherently mal-apportioned districts. Because the undercount occurred predominately among minority populations, Congressional and state legislative districts with substantial numbers of minority residents will in fact contain a much larger population than what the unadjusted data indicate. Thus, those districts would in reality be comprised of a larger number of residents than districts which are predominately non-minority. The differences between the size of

the actual population in such districts could exceed the deviation permitted under the "one person, one vote" principles of current law.

We are also concerned about the negative impact unadjusted data could have on voter participation in communities with language barriers. Section 203 of the Voting Rights Act requires jurisdictions that meet certain criteria to provide language voting assistance to their residents. Jurisdictions qualify if (a) they include at least 10,000 voting-age citizens who belong to a single language community with limited English-language abilities, or (b) such citizens comprise more than 5% of their voting-age citizen population. This is determined by census data.

For the Latino community and the nation as a whole, the repercussions of not releasing data adjusted to correct the undercount will extend far beyond our political system. In general, accurate, corrected data are vital for all types of programs and services. As you know, Mr. Chairman, there has been much discussion about the dramatic growth of the Latino community, and its implications for this country's economic, social and political institutions. This is an important discussion, because as a result of this growth, our community and nation will face many challenges. Moreover, community providers, urban and rural planners and policy makers must be equipped with the most accurate baseline data available to make the comparisons and assessments that are critical for their work.

A census undercount also drastically undermines access to quality education, a particularly important issue for Latino families. The Census Bureau's most recent Current Population Survey data reveal that 36% of the Latino population is under the age of 18. Decisions about the allocation of resources in school districts are based on census data. We know very well who was actually missed in the 2000 census. In low-income communities, it was immigrants and children. What this means to many Latino communities across this nation is that when school administrators are determining where to build new facilities, the number of teachers they need, or the number of school books to buy, they may mistakenly plan for 10,000 children, instead of the 12,000 who actually reside and attend school in the district. Given the extraordinary crisis in our public schools today, and their inability to adequately educate the nation's Latino children, this is an extremely critical juncture for our nation's future success.

So there is much at stake for the Latino community, not just politically, but also economically. If the Latino population is not fully counted, the communities in which they reside will likely lose funding for schools, hospitals and other vital social programs. These communities will, in effect, be disenfranchised for the next ten years.

What is even more pressing now, Mr. Chairman, is the recent revelation that the Census Bureau has, using scientifically approved methods to correct the undercount, produced a corrected set of numbers down to the block level for the 2000 census in all 50 states. Today perhaps, Mr. Chairman, in this committee which has a tradition of openness and full disclosure, we can receive a commitment from the Census Bureau and the Commerce Secretary to release the corrected data if the Bureau determines they are indeed more accurate than the traditional "headcount." Our government may have spent as much as $400 million to pay for the A.C.E.

If the A.C.E. has produced the best numbers available, Congress and the American taxpayers should be entitled to this important information.

Mr. Chairman, we urge Secretary Evans to direct the Bureau to complete an analysis of the accuracy of the adjusted data as quickly as possible. If that analysis reveals that the adjusted numbers are more accurate than the unadjusted count, the Bureau should immediately release the data for redistricting and other purposes. If the analysis is completed after the data can be feasiblely used for redistricting, the adjusted numbers should still be released, for public policy planning purposes and to assist us in improving the way we conduct future census efforts. If the analysis reveals the adjusted numbers are less accurate than the unadjusted count, the adjusted data should still be released to enhance our understanding of census enumeration methodology.

I thank the Chairman, the Ranking Member, and the Committee once again for providing the NALEO Educational Fund with the opportunity to share our views today on the release of the Census 2000 data.

Senator EDWARDS. Thank you very much, Mr. Vargas. As the chairman indicated earlier, when the House Members got back from their vote, we were going to interrupt this panel so that they could present their testimony. Dr. Ericksen and Dr. Wachter, we need to interrupt you and proceed with the House Members.

I do not know who wants to go first. Mr. Clay, would you like to begin?

STATEMENT OF HON. WILLIAM CLAY,
U.S. REPRESENTATIVE FROM MISSOURI

Mr. CLAY. Mr. Chairman, thank you for inviting us to speak before your committee today. As you know, the census is an issue of great importance to every Member of the House of Representatives every 10 years, this country has peacefully redistributed congressional power between the states based upon the census and remarkably only once in our nation's history has this process failed us. In 1920, Congress would not accept that the population change had shifted the majority from rural areas to urban areas. This disbelief led to a decade of inactivity.

As a result, in 1929 Congress put an end to the decennial battles in Congress over how the census would be taken and how seats would be apportioned by approving legislation that established permanent authorization for the census, and a standing formula for the distribution of seats among the states. In other words, Congress put an end to the political wrangling over the census.

Unfortunately, that legislation enabled another unsavory tradition, congressional districts of unequal size. By the time the Supreme Court ruled that these districts should be equal, the largest congressional district was 10 times the size of the smallest. In this decision, the court said such variances were blatantly unfair. The Supreme Court, in righting the wrongs of redistricting, put this census clearly in the middle of the political battles once again.

Voting rights and the census are subjects near and dear to the hearts of African Americans. Historically, we have been shortchanged by both. The first census counted African Americans as three-fifths of a person. That ugly tradition of part person, part property, continued until the passage of the Fourteenth Amendment in 1868.

In 1940, African Americans first made us aware of the errors in the census when more African American young men registered for the draft than the Census Bureau thought existed in the entire nation.

The realization that African Americans are routinely overlooked at a much higher rate than whites has been the basis of fixing the errors in the census. Even the Census Bureau's own numbers show that over 2 million African Americans were missed in the 2000 census.

Today, we find ourselves at the end of a 5-year political dog fight over the census in that same unsavory tradition of the 1920's. It is time for this Congress to act as responsibly as the 71st Congress did in 1929.

There is a very easy way to put an end to the arguments over the quality of the 2000 census. Release all the numbers for public scrutiny. If the 2000 census is as accurate as the Secretary says it is, the numbers will show it. If there are flaws in the census, the numbers will show that, too, and if there are flaws in the dual system estimates created by the Census Bureau, public scrutiny will reveal those flaws as well. One thing is for sure, hiding the num-

bers will only increase the public suspicion that the process has been rigged.

Congress has a long and honorable tradition of standing on the side of public disclosure of government information. It was Congress that broadcast the first Census results to the public, and it was Congress that established the Government Printing Office and the depository library program to keep the public informed. It was also Congress that passed the Freedom of Information Act and the Presidential Records Act, and it was Congress that said groups advising the executive branch should do so in the light of day, and pass the Federal Advisory Committee Act.

Congress has stood again and again for the right of the public to know what its government is doing, and the 107th Congress is obligated to carry on that great tradition. That is why I urge you to join with me in calling for the release of all of the data from the 2000 census to do any less would deny the American public its right to an open and honest government.

As a note of interest, representing the state of Missouri, we see that according to Gene Ericksen, Missouri was undercounted by approximately 31,000 people. I think that it is only right that the Census Bureau release the adjusted numbers for the states, because it is going to have a negative impact on the state of Missouri as well as all of the other states where an undercount exists, and I represent an urban-suburban district, and I am certain that those 31,000, the majority of them come from the area that I represent.

So I thank you, Mr. Chairman, for this opportunity.

Senator EDWARDS. Thank you, Congressman. Thank you for taking the time to be with us and sharing your testimony with us.

Mr. Gonzalez.

STATEMENT OF HON. CHARLES A. GONZALEZ,
U.S. REPRESENTATIVE FROM TEXAS

Mr. GONZALEZ. Mr. Chairman, thank you very much, and to the other Members of the Committee for this opportunity to appear before you today. I am here today as chair of the Congressional Hispanic Caucus Census and Civil Rights Task Force to share our thoughts on recent actions and decisions affecting the 2000 census and census data to be utilized over the next 10 years.

As we are aware, the 2000 census has shown the nation's Hispanic population to be the largest and fastest-growing minority population in the entire country. Unfortunately, it is also among the highest undercounted populations. As such, I would argue the nation's Hispanic community has the most to lose if we continue down the road in utilizing census data that by all accounts can and should be greatly improved.

We can all agree that the 2000 census was vastly improved from the past decennial censuses and that it was an operational success. However, there remains much room for improvement, and we should support a concerted effort aimed at identifying shortcomings in those areas need of improvement.

On March 1, 2001, Acting Census Bureau Director William G. Barron, Jr., who was here today—I am not sure if he is still here—recommended to Secretary of Commerce Don Evans that unadjusted data be released to the Census Bureau's official redis-

tricting figures to the states. Mr. Barron's recommendation was based on the report of the Executive Steering Committee for Accuracy and Coverage Evaluation Policy, referred to as ESCAP.

While that report did, in fact, recommend that unadjusted numbers be released as the Census Bureau's official redistricting information, it also stated that the Committee believed that (1) unadjusted census totals would reflect a national undercount, which has been brought up by each speaker here today, (2) that there was considerable evidence to support the use of adjusted data, and (3) that further research may establish that adjusted data would result in improved accuracy.

It is critically important to note that the Committee cited time constraints and lack of sufficient information at that point in time as key to its recommendation. In addition, in its report, the Committee was clear that although its recommendation was to release unadjusted data, that decision did not mean or even suggest that unadjusted data was superior in any way.

In a meeting on March 2, 2001, with Secretary Evans, just days before the Secretary announced its decision to release the less accurate, unadjusted data as the official redistricting figures, the Congressional Hispanic Caucus requested that the Secretary require the Census Bureau to complete its work on Census 2000, and to subsequently make a recommendation regarding the use of adjusted data for other purposes, including the Intracensal Census.

Specifically, the Hispanic Caucus requested that the Census Bureau be afforded any additional time and resources necessary to conduct further research into inconsistencies between estimates derived from outmoded Demographic Analysis tools and the modern scientifically enhanced and supported accuracy coverage evaluation estimates.

We believe that such an evaluation is critical to establishing the modern day reliability of Demographic Analysis as a tool and subsequently to determine whether or not adjusted data would result in improved census accuracy.

We further requested, and it has been requested here today by the witnesses, in keeping with the Census Bureau's longstanding policy of openness, that the adjusted or corrected block level data be made available.

I am going to make reference now to the ESCAP report on page 1 of 28, where it states, this report is also being released to the public at the same time that it is being forwarded to the Secretary of Commerce. Footnote 2, in addition to the requirement to make the report public, the Census Bureau firmly believes that full disclosure in a vigorous and informed debate will improve both the Census Bureau's internal processes and the public's understanding of statistical adjustment.

So in keeping with the very spirit of the bureau and its policy, and with that I would imagine the Secretary of Commerce would also follow that spirit, we are asking today that that unadjusted data be made available to anyone that would request it, and not wait for a Freedom of Information Act or any other type of request. Just put the information out there.

The Secretary in his testimony earlier this morning, sir, indicated that he did not believe that the information would be fit for

use. I beg to differ. It is all in the eye of the beholder. There are statisticians and scientists outside of the bureau that may find that information very useful for whatever purpose they choose. If there is disagreement as to its utilization, then we will have an open debate in the marketplace of ideas, which is the best type of dialog.

So we are here today, and I am representing the caucus again asking this Committee to please join us in the request of Secretary Evans regarding that full resources immediately be applied to reaching some conclusion on the discrepancy between the democratic analysis and the A.C.E. so that we can have the adjusted figures for all other purposes.

In the state of Texas, Senator, this impacts the Hispanic community like no other segment of the population. In Texas, it is 350,000. In my district, it could be 35,000, it could be 50,000, but when it comes to education, when it comes to roads, when it comes to libraries, when it comes to health care, all of it, we put a human face on it, and all we are asking is for accuracy, with the tried and proven scientific method.

This fight is still not over. We have characterized it as a civil rights debate and issue of the decade, and we still believe it is, and again, thank you for your courtesies.

[The prepared statement of Mr. Gonzalez follows:]

PREPARED STATEMENT OF HON. CHARLES A. GONZALEZ,
U.S. REPRESENTATIVE FROM TEXAS

My thanks to Chairman McCain, Ranking Member Senator Hollings and the entire Senate Committee on Commerce, Science and Transportation for inviting me to testify today on an issue in which I, along with every person residing in the United States of America, have a vested interest—the Census.

I am here today as the Chair of the Congressional Hispanic Caucus Census and Civil Rights Task Force to share our thoughts on recent actions and decisions affecting the 2000 Census and census data to be utilized over the next decade. As we are aware, the 2000 Census has shown the nation's Hispanic population to be the largest and fastest growing minority population in the country. Unfortunately, it is also among the highest undercounted population. As such, I would argue that the nation's Hispanic community has the most to lose if we continue down this road of utilizing census data that by all accounts can and should be greatly improved.

We can all agree that the 2000 Census was vastly improved from past decennial censuses, and that it was an operational success with higher than expected numbers of people participating by returning their census forms. I particularly credit Census Bureau employees nationwide, the thousands of communities and organizations across the country that partnered with the Census Bureau to promote participation, and the Census Bureau for devising and implementing its comprehensive plan, all of which contributed to a better census.

However, there remains much room for improvement and we should support a concerted effort aimed at identifying shortcomings and those areas for improvement. It is only through this process that we will be able to ensure improved accuracy and greater success in future censuses. Through this process we will also be afforded the opportunity to establish the validity of the Accuracy and Coverage Evaluation (A.C.E.).

In fact, it is my understanding that using the A.C.E. population analysis tool the Census Bureau has now completed the process of producing adjusted census data down to the block level for all fifty states. That data may be more accurate than data derived using only traditional methods, but we simply do not know because that corrected data has not been made public.

On March 1, 2001, Acting Census Bureau Director, William G. Barron, Jr. recommended to the U.S. Secretary of Commerce, Donald Evans, that unadjusted data be released as the Census Bureau's official redistricting data. Mr. Barron's recommendation was based on the recommendation included in the Report of the Executive Steering Committee for Accuracy and Coverage Evaluation Policy (E.S.C.A.P.).

While that Report did in fact recommend that unadjusted data be released as the Census Bureau's official redistricting data, it also stated that the Committee be-

lieved that (1) unadjusted census totals would reflect a national undercount, (2) that there was considerable evidence to support the use of adjusted data, and (3) that further research may establish that adjusted data would result in improved accuracy. It is critically important to note that the Committee cited time constraints and a lack of sufficient information as key to its recommendation.

In addition, in its Report, the Committee was clear that although its recommendation was to release unadjusted data, that decision did not mean or even suggest that unadjusted data was superior in any way. In fact, I believe that given the appropriate time and necessary resources to thoroughly and completely investigate the inconsistencies that led to what I consider to be a preliminary decision on the accuracy of adjusted data, the Committee would conclude, as it has already intoned, that adjusting census data would in fact result in a more accurate, true, and reflective evaluation of the nation's population.

In a meeting with Secretary Evans, just days before the Secretary announced his decision to release less accurate unadjusted data as official redistricting data, the Congressional Hispanic Caucus requested that the Secretary require the Census Bureau to complete its work on Census 2000, and to subsequently make a recommendation regarding the use of adjusted data for other purposes, including the Intracensal Census.

Specifically, the Hispanic Caucus requested that the Census Bureau be afforded any additional time and resources necessary to conduct further research into inconsistencies between estimates derived from an outmoded demographic analysis tool and the modern, scientifically enhanced and supported Accuracy Coverage Evaluation estimates. We believe that such an evaluation is critical to establishing the modern-day reliability of the demographic analysis tool and subsequently to determine whether or not adjusted data would result in improved Census accuracy.

While Secretary Evans stated support for moving full steam ahead, it appears that little if anything has been done by the Department of Commerce or at the Census Bureau to continue that evaluation process. The Congressional Hispanic Caucus is deeply concerned by this inactivity, particularly given the importance of determining the final accuracy or inaccuracy of the census and the impact that will have on Hispanic and other minority communities that were once again disproportionally undercounted.

For them, the estimated 3.4 million Americans, including over a million Hispanics, who the 2000 census failed to count using traditional methods, this is the civil rights issue of the decade. If we fail to do all that we can to correct the problems, then those 3.4 million people will simply not count for the next 10 years.

We should all take a step back and consider what this would mean for people living in each and every community where minorities and children were disproportionally undercounted. It means a real and increased potential for another decade of unjustly lost federal dollars to build schools, roads, day care centers, and to fund countless other programs that rely on census data. It also means diminished political power for each of those communities. That is hardly the level of justice and fairness that we as a revered Democratic nation should aspire to.

Therefore, in the spirit of fairness to all, regardless of age, ethnicity, race, or economic status, the Congressional Hispanic Caucus fully supports and encourages the Chair, the Committee, and the entire Congress to seek the release of the Census Bureau's adjusted census data so that it may be reviewed and where deemed appropriate applied.

Senator EDWARDS. Thank you, Mr. Gonzalez. Thank you for your leadership of the Hispanic Caucus, and thank you for taking the time to be with us here today.

Mr. Miller.

STATEMENT OF HON. DAN MILLER, U.S. REPRESENTATIVE FROM FLORIDA

Mr. MILLER. Thank you, Senator, and thank the Committee for allowing me to testify today. As you may know, I have been the Chairman of the Subcommittee on the Census since 1988, and I have been deeply involved in the oversight of the 2000 Census from planning through its execution.

Our public position on adjustment should not be a surprise to anyone. I have always been concerned that adjustment introduces more error into the census than it has the ability to correct, par-

ticularly at lower levels of geography critical for the redistricting process.

However, even if sampling errors were not the issue, there are larger legal and public policy implications that should be giving all of us very serious concern. Acting Director Barron outlined many reasons for the success of the 2000 census enumeration when he testified before my subcommittee several weeks ago. Among those reasons, congressional support allowed the Census Bureau to hire ½ million enumerators at competitive wages during a time of record low unemployment. This support also allowed for a first-time-ever paid advertising campaign, more than 40,000 local partnerships, and an unprecedented effort to provide multilingual assistance.

In the end, it was emphasis on counting people, not making estimates, that made this census a success and reduced the differential undercount of minority communities. We need to give a great deal of gratitude for the great work the Census Bureau did on this, and we should commend the efforts of the Census Bureau.

Supporters of adjustment say that success is not enough. They will argue for the release of these inaccurate numbers in the name of fairness and justice. I share their desire for fairness, but good intentions do not justify bad public policy, and a statistically adjusted census is bad policy on many different levels.

Let me explain. Statistical adjustment is less accurate where it is critically important to be accurate, at small geographic levels. We have known for some time that statistically adjusted numbers do not give us a more accurate picture of the population when describing smaller aggregations of the population such as small towns, rural areas, and blocks. The bureau's recommendation against sampling reads, "analysis for counties with populations below 100,000 people indicated that the unadjusted census was more accurate,".

Since apportionment and redistricting go hand in hand, it is arguably illegal for adjusted numbers to be used for redistricting, based on the 1999 Supreme Court ruling. Use of adjustment would reduce people's incentives to participate in the census, which would degrade data quality and reduce accuracy. Why stand up and be counted when you can sit down and be sampled?

It is also disturbing to me that under adjustment, some people are counted as less than a whole person. Under adjustment, everyone is assigned to a category. One of these categories, for example, is Native Hawaiian and Pacific Islander women over 18 who own their own home. If adjustment were to go forward, everyone in that category would be assigned a value of .95. In other words, every Hawaiian adult woman owning a home anywhere in the United States who answered the census would be counted as little more than 9/10ths of a person.

Adjustment assumes that all people in a certain category act alike, or have the same likelihood of filling out their census forms. These types of assumptions are something we should be moving away from and not embracing.

Finally, I worry about the subjective assumptions that are inherent to the statistical adjustment process. A professor at Harvard

University put it very well in a commentary in the *Wall Street Journal* on February 15. He said,

> "Unfortunately, statistical adjustment also gives much greater discretion to the Census Bureau. The correction procedure is based on populations of groups, and choosing them is very subjective. Do we treat all young, urban black males as a subgroup, or do we separate them by region? How many ethnic groups do we want to treat as distinct? This leads to a general point. As you allow for more statistical sophistication, you put more discretion in the hands of the statistician."

This census and the Census Bureau have proven we can achieve nearly 100 percent accuracy through a strong congressional commitment, strong Bureau management, improved technology, and expanded local partnerships. By all means, there is still room for improvement, but this and future Congresses should put a priority on the methods that are legal, accurate, and ethical.

Finally, on the question of releasing adjusted numbers, census data for Federal funding and other purposes, because of the errors and the problems with their adjustment, I think it would be highly irresponsible to release adjusted data for any official purpose.

One of the things we have heard for the past several years in the Committee by my colleagues on the other side of the aisle is, listen to the experts. Let us trust the experts. Well, we have this ESCAP committee of Census Bureau experts, and the experts said there is a problem with the data, and so let us listen to them. The data should not be released at this time. Let them continue to work on that data and if at some stage they feel they have accurate data, then it can be released, but it would be irresponsible, in my opinion, to release that data now.

Thank you, Mr. Chairman.

[The prepared statement of Mr. Miller follows:]

PREPARED STATEMENT OF HON. DAN MILLER,
U.S. REPRESENTATIVE FROM FLORIDA

Thank you Mr. Chairman and Members of the Committee, for granting my request to testify before you today regarding the 2000 Census.

As you know, as Chairman of the Census Subcommittee since 1998, I have been deeply involved in the oversight of Census 2000—from planning through its execution. My public position on adjustment should not be a surprise to anyone. I have always been concerned that adjustment introduces more error into the census than it has the ability to correct, particularly at lower levels of geography critical for the redistricting process. However, even if the sampling errors were not the issue, there are larger legal and public policy implications that that should give all of us serious concern.

Acting Director Barron outlined the many reasons for the success of the 2000 census enumeration when he testified before my subcommittee several weeks ago. Among those reasons, Congressional support allowed the Bureau to hire a half million enumerators, at competitive wages, during a time of record low unemployment. This support also allowed for a first-time-ever paid advertising campaign, more than 140,000 local partnerships, and an unprecedented effort to provide multilingual assistance. In the end, it was the emphasis on counting people—not making estimates—that made this census a success, and reduced the differential undercount of minority communities.

Supporters of adjustment say this success is not enough and they will argue for the release of these inaccurate numbers in the name of fairness and justice.

I share their desire for fairness. But good intentions do not justify bad public policy. And a statistically adjusted census is bad policy on many different levels.

• Statistical adjustment is less accurate where it's critically important to be accurate—at small geographic levels. We have known for some time that statistically adjusted numbers do NOT give us a more accurate picture of the population when describing smaller aggregations of the population, such as small towns, rural areas

and blocks. The Bureau's recommendation against sampling reads, "The analysis for counties with populations below 100,000 people indicated that the *unadjusted* census was more accurate."

• Since apportionment and redistricting go hand-in-hand, it's arguably illegal for adjusted numbers to be used for redistricting, based on the Supreme Court's 1999 ruling.

• Use of adjustment would reduce people's incentive to participate in the actual census, which would degrade data quality and reduce accuracy. Why stand up and be counted when you can sit down and be sampled?

• It is also disturbing to me that under adjustment, some people are counted as less than a whole person. Under adjustment, everyone is assigned to a category. One of these categories is "Native Hawaiian and Pacific Islander Women over 18 who own their home." If adjustment were to go forward, everyone in that category would be assigned a value of 0.95. In other words, every Hawaiian adult woman owning a home anywhere in the United States, who answered the census, would be counted as little more than nine-tenths of a person. Adjustment assumes that all people in a certain category act alike, or have the same likelihood of filling out their census forms. These types of assumptions are something we should be moving away from, and not embracing.

• Finally, I worry about the subjective assumptions that are inherent to the statistical adjustment process.

A professor at Harvard University put it very well in his commentary in the *Wall Street Journal* on February 15. He said,

> "Unfortunately, statistical adjustment also gives much greater discretion to the Census Bureau. The correction procedure is based on population subgroups, and choosing them is very subjective. Do we treat all young urban black males as a subgroup or do we separate them by region? How many ethnic groups do we want to treat as distinct? This leads to a general point: As you allow for more statistical sophistication, you put more discretion in the hands of the statistician."

This census and this Census Bureau have proven that we can achieve nearly one hundred percent accuracy through a strong congressional commitment, strong Bureau management, improved technology, and expanded local partnerships. By all means, there is still room for improvement. But this and future Congresses should put a priority on the methods that are legal, accurate and ethical. We should not continue the pursuit of a costly, unlawful, inaccurate, and racially biased adjustment of our constitutionally mandated decennial census.

Finally, regarding the question of releasing adjusted census data for federal funding and other purposes, because of the errors and problems with the adjustment, I agree with Bureau officials that it would be highly irresponsible to release adjusted data for any official purpose.

Senator EDWARDS. Thank you, Mr. Miller.

Mr. Gonzalez, in the 1990 census, over 6 percent of the Hispanic community in North Carolina was under-counted. During the last 10 years, according to the 2000 census, we have had almost a 400-percent increase in the Hispanic population. I cannot imagine what the undercount must be in this census.

You and your family, your dad, has been involved in the Hispanic community and served as leaders in that community for a long time. I wonder if you would talk a little bit about what impact undercounting has, because these are real people. It affects their lives. They are not numbers on a piece of paper. I wonder if you would talk about the impact undercounting has on the lives of the people who are not being counted.

Mr. GONZALEZ. I think in terms of city planners, Senator and in terms of how you build, you know, smart growth and communities, your school administrators and trying to anticipate the number of students, and especially in a state like Texas, a city like San Antonio, 150 miles from the border, but all of the southwest states with some of the greatest growth, your four largest growing cities, really, it is because of the Hispanic populations, and that is not even to say what is going on in states such as yours.

But there is a practical application. These numbers are used to plan ahead, not just for funding purposes, but to get accurate head counts on what you need as far as medical facilities, our county hospital, which has a tremendous burden of taking care of many people that are not insured. Everything goes hand in hand here. The uninsured problem, the number of people, we have to get an accurate number.

How do you plan how many students are gong to be in class? How many need Head Start? How many need meals? But it goes beyond that, in roads and highways.

I was thinking the other day, even Meals on Wheels is predicated on this. I had a meeting the other day with the Small Business Development Center in San Antonio. They service 79 counties in South Texas with a heavy Hispanic population, and with the cuts on the SBA, you can imagine how that is basically almost—the effect is multiplied, when we do not have accurate numbers.

We have reduced amounts in the budget. We are all talking about keeping within budget caps, there is going to be reduced spending. Every dollar becomes more precious. How do we apply these dollars in the best way, the most economical and effective way, if we do not have accurate numbers, and that is what the cities are faced with. That is what my district is faced with.

In talking to the University of Texas in San Antonio and their community development center, they are telling me any cuts will be drastic in trying to take care of the small businessman and woman in 79 counties. It is important to have this information, and have it now, so that we can have our experts look at this information.

And getting back to a very important point that my colleague in Florida has pointed out, their big concern has been, is the scientific method loses some of its accuracy at the block level.

If you believe in ESCAP, which I do, on page 27, block level accuracy is not an important criterion to evaluate either Census 2000 or A.C.E., so they made that conclusion. They have made that finding. They have addressed that concern.

This is all really important. I sit down with my city council, this is what they are bringing to my attention.

Senator EDWARDS. Thank you, Mr. Gonzalez. It is so important that we put a human face on this. These are not just numbers.

Mr. Clay, I wonder if you could comment, because I suspect some these are some of the same concerns you might have in the African American community.

Mr. CLAY. Mr. Chairman having spent 17 years in the Missouri General Assembly, and having some experience with pass-through Federal dollars that we pass on to local communities in Missouri, it is so important, such as Head Start, we know how important it is to get an accurate count of people throughout the state of Missouri.

It seems to me that in urban areas, especially the area that I represent, that the needs are so great that you rob those communities of needed resources when you do not accurately count those communities, so what happens is that those communities where overcounts occur are usually the ones that get extra resources, and

the ones where the need is the most get the least of those resources.

Senator EDWARDS. For the reason that both of you have described so eloquently, I think it is critically important that we have the census be as accurate as possible. I also have great difficulty understanding why the adjusted numbers are not being released and not subject to public scrutiny, which I believe they should be.

Senator Kerry, did you have questions? Mr. Miller, I think, had something he wanted to say.

Mr. MILLER. Let me add a couple of comments. The question as to the relief is whether they are accurate enough, and there is very serious concern by the statisticians at the bureau, the professionals, not politicians, that at this stage it is not accurate enough, so if we are going to trust the professionals, we need to do that.

One of the questions comes up about money. First of all, everybody agrees we have had a great census this time, and we really should celebrate the census. The differential undercount for Hispanics and African Americans has been cut in half so we have really had a great effort, and Mr. Gonzalez has worked hard in his district, working with the Catholic Church, the Hispanic Church and all of that, so we have done a good job in executing a better census.

With respect to the money issue, most of the money we are talking about is a zero sum game, and so everybody is saying we are going to lose all this money. Well, if, for example, Meals on Wheels, it is important in my district, lots of seniors in my district. There is a set pot of money for Meals on Wheels.

The question is, you are not going to get more money. It is just the total amount of money. That is a budget fight rather than a census fight, so I just think sometimes it is distorting the thing to say we are missing all this money. Well, if you add more money in his district, you are taking money from my district, so it is really a zero sum game when we are talking about money.

Yes, we need to have the most accurate census possible, and I think we did a good one, and I think we should really be proud of what the bureau did under the leadership of Dr. Prewitt and John Thompson and the others at the bureau.

Thank you.

Senator EDWARDS. Mr. Gonzalez, Mr. Clay, did you have any comment in response to that?

Mr. GONZALEZ. Well, I understand we all have competing interests. We represent our constituencies. But what if I just happen to have a certain segment of the population that a certain program is attempting to address their needs more so than another colleague. I can make that argument with census figures. I cannot make that argument if I do not have accurate numbers.

And I know this debate really goes almost to a political philosophy, or whatever, and it should not be. It is really one about sound science, and we are willing to debate that. If you want to go to court we are willing to debate it in a courtroom. I can guarantee you now that this will pass legal muster when it comes to scientific methodology and what is accepted out there in the scientific community.

Thank you.

Mr. CLAY. Mr. Chairman, just to add to that, I would think that the Census Bureau as well as the U.S. Government would want to get it right, would want to be as accurate as possible.

I mean, as I stated in my testimony, we can go back to 1940 when the undercount was so drastic among African American males that they did not realize that they had made an error until these African American males showed up in greater numbers for the draft for World War II, and then the U.S. Government realized how severe the undercount was in that community, so that needs to be corrected, and that is something that happened more than 60 years ago.

Senator EDWARDS. Mr. Clay, Mr. Gonzalez, I share your concern. Mr. Kerry is now back and may have some questions for you.

Senator KERRY. Thank you very much, Senator Edwards. Thank you for relieving me for that period of time. I appreciate it.

Mr. Miller, let me just comment very quickly on your comment about sort of the zero sum game. In one sense, yes, there is a specific amount of money, and it depends where it goes, but it really runs a little deeper than that, because when you talk about where it may go from, you are talking about the difference between affluent communities and communities that are struggling with resources, and you are talking about the difference of a community that has no tax base for property taxes, and yet their schools depend on it, and here you have an allocation that may be dependent on population as you do in certain school issues, or seniors, as Congressman Gonzalez has said, so it really is a question of fundamental fairness.

Secretary Evans acknowledged there is an undercount. I assume if you accept the notion that there is an undercount, you accept the notion that there is something unfair. Do you accept that concept, if there is an undercount?

Mr. MILLER. The differential undercount is what we are all concerned about. Blacks, Hispanics, for example, are counted low, and not as accurately as whites.

Senator KERRY. Do you think they should live with that for 10 years?

Mr. MILLER. No. We should try to get the most accurate census we can, but if the adjustment is not accurate, if the statisticians and the professionals and the career people in the bureau say the numbers are not accurate, would we want to use inaccurate data.

Senator KERRY. But they are not saying that they are inaccurate. They are trying to figure out where the distinctions may be. Now, why would we not want competent professionals to view the data—I mean, are you suggesting that some of these people sitting at this table are not professionals?

Mr. MILLER. No. I think we need to wait for the bureau to continue doing their work. They want to release the numbers before the bureau has certified the numbers.

Senator KERRY. Well, the problem is, we could run out of time. We could run out of time to measure and come up with a solution in which we share some thinking about how you rectify some of these differences.

I mean, I would like to have outside professional help me to say, well, Senator, here is how you could really do this in a fair-minded,

thoughtful way, but if we allow this to be behind closed doors for the next months without any capacity to measure the data, then we are stuck with whatever judgment is made.

Mr. MILLER. But the professionals should have a chance to look at it, and they have reached an agreement between the bureau and the Republicans, the Democrats, and other outside groups to have a chance to look at it, and they have also contracted with the National Academy of Sciences, which is very independent, on doing an analysis of the same numbers.

Unfortunately their analysis is not going to be ready until probably early next year, so it is a real mistake to try to use faulty data until the bureau, until the National Academy of Sciences, until the professionals sitting at this table and others have said they are accurate numbers. If you do not have accurate numbers, you are making a real mistake playing around with potentially faulty data. We do not know yet.

Senator KERRY. Does it mean anything to you at all that this is so divided along partisan lines?

Mr. MILLER. It is unfortunate.

Senator KERRY. Is it more than unfortunate? Is there a message there?

Mr. MILLER. I used to teach statistics. That is how I got to be Chairman of this particular subcommittee, but we have to base it, as we said all along, on science, and if the science says it is not right, why would we want to do something wrong—it is like a medical test.

Senator KERRY. I do not want to overrule science, believe me. If I started doing that for this, I would be in trouble on everything else. None of us are asking to overrule science, but there are three different methodologies here for how you arrive at the ball park. One, you have the Demographic Analysis, which they acknowledge has a problem with respect to immigration, You have the A.C.E., which we do not yet understand completely what the differential is, then you have the actual count.

Now, if all of them are showing an undercount, it seems to me it is like a court of equity, where a judge sits there and makes some judgments. Well, how are we going to make up for that undercount level we have counted? You are shaking your head, Mr. Murray. You do not think we should do that? We are going to come back to you.

But let me just ask Congressman Gonzalez and Congressman Clay a question, sort of following up on what Senator Edwards said. Your constituents, or the people in your districts who are cognizant already, or who have knowledge that there is not a count that is not fair, I mean, do they come up to you and talk to you about it?

Mr. GONZALEZ. Senator, as my colleague was indicating, we all came together, the city council, commissioners, the courts, the church, every social and civic organization, we came together to get people to make sure they return their forms, answer the door, that anybody who is not documented did not have fear of being arrested. It was an incredible effort, and we did end up with a more accurate result, but nevertheless one that is still inaccurate and leaves many undercounted, especially in the minority Hispanic commu-

nities. There is a very high awareness in my city that we are going to be basically short-changed.

Senator KERRY. What is the impact of that? What is the effect of that on the body politic?

Mr. GONZALEZ. As I have indicated, of course just planning——

Senator KERRY. Leaving out the planning.

Mr. GONZALEZ. The funds. We lost a substantial amount of Federal funds in the past 10 years because we missed 40,000 people. We missed 16,000 children in San Antonio. We are part of a lawsuit The city council has authorized the city to go into the lawsuit—I guess it is out of Los Angeles, or whatever—regarding trying to get these numbers and having the adjusted figures utilized.

All of this impacts services to the neediest of all citizens, and we are talking about, and some are not citizens but residents, and part of this, Senator, is trying to figure out why we have this discrepancy. Everyone is shocked at the number of immigrants in this country, of the number of undocumenteds. It has repercussions for Immigration and Naturalization Service and Social Services, as I have already indicated.

Senator KERRY. Well, let me ask you a very hard-nosed question, if I may, and welcome, Congresswoman Maloney, thank you, and let me just say to all of you, I apologize profusely, we have had to go back and come back and vote, and I know how hard that is, and I apologize for the interruption. It is just the nature of the beast around here, and we all try to do the best we can.

I want to ask one last question, then I am going to try to——

Mr. CLAY. Mr. Chairman, if I could just follow what Mr Gonzalez said, that also in the St. Louis community people are wondering just when will they be counted. They have witnessed the fiasco in Florida, where votes were not counted for the Presidential election, and now they are witnessing a back-pedaling of sorts by the Census Bureau, and when will they count.

Senator KERRY. This brings me to my question, and I am going to be sort of devil's advocate here. You know, some people sit on the other side of the fence, and they would say to uncounted people and to you, "Well, wait a minute, they had a responsibility to fill out their census form". We had the most broadbased outreach ever in history, and others filled it out and they got counted. Why do we have to make up for the seemingly unwillingness of these folks to participate? That is the question you get, and what is your response?

Mr. CLAY. Well, the system was quit a bit better, as Mr Miller says. I agree, the Census Bureau did a better job than in 1990. However, there for systematic reasons African Americans were missed in a disproportionate number, compared to the rest of the population, as well as Hispanics. They were missed also, because of system flaws, in just raw data, in counting people, so when we know that there is an undercount it is incumbent upon us, as a government, to go back and try to be as accurate as possible.

Senator KERRY. Would you say that to a certain degree that some folks whose level of education may be higher, or whose communication skills may be better, or whose status and position in life are better for whatever reason, that there may be sort of a disadvan-

tage in life that at this point plays out in how the census in fact is counted?

Mr. CLAY. Well, I would not go that far. What I would say is that people earnestly made an effort to be counted. If some were missed, we need to figure out why they were missed, and we need to go back and get as accurate a count as possible.

Senator KERRY. Let me let your colleague have a chance.

STATEMENT OF HON. CAROLYN MALONEY,
U.S. REPRESENTATIVE FROM NEW YORK

Ms. T4Maloney. I just would like to add to the discussion that the census is one of the great civic ceremonies that we all participated in and saw our responsibility to fill out our forms and be counted, and it is important not only for our own families, but for our neighbors, because if we are not counted, and we are left out of Federal funding formulas to the tune of $185 billion a year that are distributed on census numbers representation, which is the true power in this country, and just plain good data to plan for the future.

But we know that the people missed overwhelmingly are in urban and rural areas, and that they are oftentimes minorities, children, and the poor, and given the changing dynamics of America, the fact that some families are working two, three jobs, some families are backed up in apartments where two or three families are living there possibly illegally, and they do not want people to know, sometimes there is a language barrier.

And given the extraordinary efforts of the Census Department and the many organizations that partnered with the Census to break through that barrier, it is still a barrier, and we know from modern scientific methods that there is an undercount, there's no question about it, so the question is, do we correct it or not, and we have the scientific ability to correct it, and why we would not go forward with the most accurate numbers is something that I cannot understand.

I congratulate you, Senator, for your interest in this, and for your leadership in it. I consider it really the civil rights issue of the decade. It is representation from which all power flows, and if you are not counted, then the school is not where it is supposed to be, the health center is not there, and I would merely like to place in the record a letter that has been signed by 107 colleagues of mine in the House of Representatives calling for the release of the numbers.

The American people paid for it, the Congress allocated—over $400 million was invested in this information, and it belongs to the American people. Why are we not releasing it? It is absolutely unprecedented that such a valuable information should be withheld.

Furthermore, no community in our country should be left shortchanged. Every state, every city, every town, and every neighborhood is entitled to information about their area, and I do not understand any logic why they would withhold it. I would like really to place into the record—we are all in different committee meetings and voting on the floor—my comments, but I would like to make them very short, but first of all the census has always been very contested.

The first census director was Thomas Jefferson, and he wrote Washington and said there was an undercount of a million people. Washington took him at his word and put a million more in the first census count. That is a true story.

Unfortunately, it has been much more contested in this Congress. It has regrettably been the most partisan of issues. It held up two budgets. It held up disaster relief.

But it is very important, because it goes to the absolute core of our democracy. If you are not counted, you are not represented, and I just want to say that the so-called sampling method is known in the scientific community as the dual system estimation, and it is a proven statistical method that has been perfected by scientists during the last half-century.

The 1990 sampling results have been the foundation of nearly every major economic statistic of this country since the middle of the last decade they are using modern scientific methods, and in this A.C.E. follow-up we paid for a sample of 314,000 households, which was then compared to the data collected from those same households during the census itself, and the so-called accuracy and evaluation, or A.C.E., is known as the census decennial report card, and we should really have these results.

The census professionals in their report stated that they favor the use of modern scientific methods, and I quote, "the committee believes it likely that further research may establish that a judgment based on A.C.E. would result in improved accuracy," and I just want to underscore that even the opponents of sampling have apparently blessed this methodology, although they may not really fully realize that they have done so.

The administration has pronounced over and over again, we heard it today, that the 2000 census is the most accurate in history, and it came to this conclusion because of the results of the very scientific tool they oppose, the A.C.E. The A.C.E. program demonstrated that the 2000 census reduced the undercount, as compared to the 1990 census. As a result, they have decided there is no need to use the A.C.E. for its full intended purpose, to correct for the errors that remain in the Census 2000, and for the administration the A.C.E. is sufficient only to buttress their own political argument against a more complete count.

Unless we have access to the A.C.E. data, we cannot prove that the Census 2000 is, in fact, more accurate, so they really, really should release it. Why the administration is keeping it under lock and key, I do not understand. It is harmful to government, and to government planners, and I might note, Mr. Chairman, that the Census Monitoring Board released the net undercount numbers, and I would like to know why they are releasing it before the bureau does, and they took this extraordinary step of presenting what they believe are the net numbers for each state that were missed in the census, because the Census bureau itself will not release them.

I would just give the numbers for my own state. You can get them for all the states from the monitoring board, and these numbers show that there was a net number of 291,000 New Yorkers were missed in the census, and 188,000 in New York City alone, merely because the Census Bureau ran out of time, and according

to an independent survey that was done by Pricewaterhouse this would result in approximately a loss of $2.3 billion over 10 years.

That is a lot of teachers, that is a lot of police officers, that is a lot of mass transit, and bridges, and roads, and to know that the accurate numbers are there and that they are not releasing them is just absolutely unconscionable. I, on behalf of my colleagues in the House, particularly the 107 who signed the letter appealing for the release and the many people in America who would like to be counted but have been left out, we thank you, Mr. Chairman, for your attention to this and for this hearing, and for the time and understanding you have put in it, and I congratulate you and thank you.

[The prepared statement of Ms. Maloney follows:]

PREPARED STATEMENT OF HON. CAROLYN MALONEY,
U.S. REPRESENTATIVE FROM NEW YORK

Thank you Chairman McCain and Ranking Member Hollings for inviting me to testify today on the Census. Mr. Chairman, welcome to the "sampling debate." By that, I mean the national dialogue which has been held during the last decade in search of ways to produce a better Census, one that would cure the national problem of the undercount and racial differential.

It is a problem that has existed in every Census, beginning with the very first census conducted by Thomas Jefferson. The Secretary of State wrote to President Washington and our ambassadors overseas saying that he felt that the just-concluded census may have undercounted the fledgling nation's population by as much as a million. Perhaps, because he was Jefferson, he was taken at his word.

In today's modern census, the Census Bureau has to do more than assert the quality of the census. It must demonstrate to Congress and the American public that it knows with some precision about the accuracy of the so-called "head count." The modern solution is sometimes derisively called "sampling," but is better known in the scientific community as "dual system estimation." It is a proven statistical method that has been perfected by scientists during the last half century. The 1990 sampling results have been the foundation of nearly every major economic statistic of this country since the middle of the last decade.

In Census 2000, the Congress funded the most robust scientific measurement of accuracy ever conducted. We paid for a follow-up sample of 314,000 households, which was then compared to the data collected from those same households during the Census itself. The Accuracy & Coverage Evaluation, or A.C.E., is known at the Census Bureau as their decennial report card. The results are in for Census 2000, and the bureau has announced that the A.C.E. worked extremely well and that the Census staff believes as they made clear in their report, that it still favored statistical sampling, "The Committee believes it likely that further research may establish that adjustment based on the A.C.E. would result in improved accuracy.".

Even the opponents of sampling have apparently blessed this methodology, although they may not fully realize what they have done. The Administration has pronounced the 2000 Census the most accurate in history. It came to this conclusion because the results of the very scientific tool they oppose, the A.C.E., demonstrate that the 2000 Census reduced the undercount as compared to the 1990 Census. As a result, they have decided that there is no need to use the A.C.E. for its full intended purpose, to correct for the errors that remain in Census 2000. For the Administration, the A.C.E. is sufficient only to buttress their own political argument against a more complete count.

Now that may be a legitimate call, but unless all of us—the Congress, outside experts, the scientific community, and the American public—unless we all have access to the A.C.E. data, we can not prove Census 2000 is in fact more accurate. And I tell you Mr. Chairman, I have yet to meet any Thomas Jefferson's in this Administration, so I am unwilling to just take their word for it.

So here we are today, holding this hearing to ask why the A.C.E. results are under lock and key, known only to a handful of government employees, and when even Congress itself has not been provided the data. Why is the Census Monitoring Board releasing net undercount numbers before the bureau? They took the extraordinary step of presenting what they believe are the net numbers for each state that were missed in the census because the Census won't release them. Those numbers

show us that a *net* of 291,000 New Yorkers were missed in the Census, 188,000 in New York City alone, missed because the Census ran out of time.

Apparently, the Administration fears that we will not understand the data, or it may be used to advance political arguments the Administration opposes, or that we should simply have to trust them to look out for our best interests.

Well Mr. Chairman, I think we understand the numbers release today very well—Americans were missed in the Census and we should not stop until we get them all included.

For another perspective on this debate Mr. Chairman, I direct you to remarks made just yesterday by Chairman Greenspan before a meeting of business economists. His presentation was titled, "The challenge of measuring and modeling the dynamic economy." His address was, in part, a call for more resources to enhance data collection methods. He said:

> ". . . the experience of the last 40 years underscores a fundamental dilemma of business economics. Should we endeavor to continue to refine our techniques of deriving maximum information from an existing body of data? Or should we find ways to augment our data library to gain better insight into how our economy is functioning? Obviously, we should do both, but I suspect greater payoffs will come from more data than from more technique."

[Remarks by Chairman Alan Greenspan "The challenge of measuring and modeling a dynamic economy" At the Washington Economic Policy Conference of the National Association for Business Economics, Washington, D.C. March 27, 2001]

What Mr. Greenspan knows and others often overlook, is that our key economic indicators, such as unemployment, poverty, inflation, and consumer confidence are all derived from "samples" of American households and businesses. This sample data is collected through interviews every month, conducted mostly by none other than the Census Bureau.

I agree with Chairman Greenspan's main thesis. There is a greater payoff for our country when we in Washington have more data, better data, complete data . . . data like that produced by the A.C.E. Statistically sound data allows policy makers and experts to calculate with more precision the true stories of economic or social policies or as former Census Director Prewitt might say, allows us to get closer to the truth. Mr. Chairman, it is time now for the Census Bureau to release all the A.C.E. results. I urge you and both Houses of Congress to call upon the Director of the Census to release those numbers today. I would like to put into the record a letter sent yesterday to Acting Director Barron from 107 House Members urging just that.

It is unprecedented that such valuable information should be withheld. Furthermore, no community in our country should be left shortchanged. Every state, every city, every town, and every neighborhood is entitled to information about the demography of our country.

Together we made a four hundred million dollar investment in this information. It belongs to the people and it should be given to the people.

Senator KERRY. Thank you very much, Congresswoman. None of the Members of Congress need to feel obligated—obviously, I know you will not, if you need to move on. I do not want to tie you down, because I know your time is critical and you have got other things going on. Is there any last comment any of you would like to make as we move back, because I want to get back to the panel, if we can.

Mr. MILLER. If I may make one last comment, Mrs. Maloney has been the Ranking Member for the past 3 years on the Subcommittee, and she has said rely on the professionals. Well, the professionals say the data is not accurate enough to be released at this time, and in 1990, when they tried to do an adjustment, there was a difference—it was not called A.C.E. back then, it had another name, but they went through several different iterations of different sets of numbers.

So if you release them before you finalize the numbers you are creating a problem. That is what the argument is over, whether you release them now, or let the Academy of Sciences and let the professionals like these gentlemen sitting here have the chance to

be sure they are accurate, so let us trust the professionals of the bureau. They are the ones who say we are not finished with this, we have got to keep working on it, and they have got the resources.

We certainly have thrown all the money that they need to continue that job, so we need to continue having them work, and once they have made that decision, then we can decide what to do with that data.

Ms. MALONEY. I just might, since money was mentioned, I certainly would like the opportunity to respond, if I could, Mr. Chairman.

Senator KERRY. It was mentioned with great credit.

Ms. MALONEY. The census professionals merely ran out of time, and they stated that they favored statistical sampling, and I quote, "the committee believes it likely that further research may establish that adjustment based on the A.C.E. would result in improved accuracy," and what I have heard today from the very studied questions of the chairman, I believe the chairman asked, when will you release this, and it was not said when it would be released.

They said they may get the numbers by the summertime, but never did I hear that they would be released, so what we are calling on is for the information to be released so that the scientists, the government planners, everyone can look at it.

And then secondly, the chairman and others questioned whether or not the numbers could be used for the distribution of federal funds, and I have not heard a response on that, and silence is usually no, and if we know that we have a more accurate number, then it is only fair and just that it be used for the important distribution of federal funds.

Senator KERRY. I think, Congresswoman, that—you may have just departed for the vote at that point. I did ask him if we would be able to get him back here to address their judgment in order that we be permitted to apply these to the allocation of funds, and he did consent to do that, and suggested they would be able to have those numbers by the early summer, and so hopefully—and we agreed to continue the discussion, so hopefully we will be able to have a dialog that will permit us to make those judgments.

Now, you are correct, there was no consent to release, despite the request to do so. There was a statement about why they saw difficulties in doing it now. I think we need to continue to press that issue. I certainly intend to, and others will, and we want to explore that a little bit with the panel of experts we have here in front of us, and I intend to do that now a little bit, and so I am very grateful to all four of you. You have been extraordinarily gracious with your time, and with the inconvenience, and I thank you very, very much for doing so. Thank you for being part of this.

Now, had you all had a chance to testify? Not yet, so you are waiting patiently. Dr. Ericksen, you are next in line, and then Dr. Wachter. Thank you, and thank you all for your patience, too.

STATEMENT OF EUGENE P. ERICKSEN, PROFESSOR OF SOCIOLOGY AND STATISTICS, DEPARTMENT OF SOCIOLOGY, TEMPLE UNIVERSITY

Dr. ERICKSEN. First of all, thank you for the opportunity to discuss this issue. I think first we need to be clear about what we

mean by census error, because that is what determines whether or not this is the most accurate census ever. The net undercount in the 2000 census is estimated to be 1.2 percent, or 3.3 million people. That is not a particularly good way to assess the census.

A better way to assess the census accuracy is in terms of what we call gross coverage error, which is the sum of the number of omissions and the number of erroneous conclusions. The Census Bureau, working with the A.C.E. data, has released estimates that vary between 9.5 and 11.9 million.

The third criterion, which may be even better, is a differential undercount, comparing the A.C.E. estimates in 2000 with the post enumeration survey estimates in 1990. It appears that the bureau was able to reduce the differential undercounts of Hispanics and African Americans by more than half. All of these indicators place great credit to the Census Bureau. They have definitely made an improvement, but what we must be most concerned about is not so much the error at the national level, but whether or not it creates an uneven situation locally.

In the 1990 census, the Census Bureau estimated there was a net undercount of nearly 8 percent among black renters living in New York City, and in the same census they estimated a small overcount of 0.23 percent among white homeowners also living in New York City. According to these estimates the most undercounted congressional district in the United States is District 16 in New York State, of 6.5 percent. The 3rd District in New York State had an overcount of 1.3 percent, and so it is the variations in the undercount and the overcount are what we should be concerned with.

Now, the Census Bureau actually had tremendous discretion, and made several important decisions in the manner of taking the census. One very important decision is the program known as local update of census addresses, LUCA. Millions of addresses were added to the census, and as far as the research I have been able to do, and the reports that I have read, lead me to believe, it could very well be that this is the most important thing that the Census Bureau did to improve the accuracy of the 2000 census.

The problem with LUCA is that it was not applied evenly. LUCA was more effective in some areas than it was in others, and it could have contributed to error, although it improved on the national level.

A second decision that is important is the decision about what we call whole person imputations. In the 2000 census the bureau ended data collection at a much earlier date than it did in the 2000 census, and it appears that one of the outcomes of this early decision is that there is a greater use made of the computer in filling out, I believe it is 5.5 million records. The comparable number in the 1990 census was about 2.1 million.

Whether or not the whole person imputation, virtual people, if you will, improved the accuracy of the census remains to be studied, but if it did improve the accuracy or degraded the accuracy, it probably did that unevenly, because some areas had a greater incidence of these kinds of additions, and others, about half of the races of the people are other than non-Hispanic white. In other words, they are minorities.

Now, I have tried, using the data that are available to me, to evaluate the differential undercount among areas. Among the states, it appears that the deviations are not particularly great. Those states where the minority percentages are higher tend to have a higher undercount. Those states where the minority percentages are lower tend to have a lower undercount, but the differences between the higher states and the lower states are not great, but they do exist.

There is another situation. I am from Philadelphia. Every year, particularly once we have had 5 or 6 years go by the census, I get phone calls from reporters. They want to know, why is the population of Philadelphia getting so small. Indeed, I believe that it was thought that Philadelphia had more population loss than any other city in the country.

Now, because the 2000 census did a better job of reducing the differential undercount, it appears that Philadelphia did not lose as many people, and so now we are in the anomalous situation that the very improvement in the 2000 census creates errors in understanding the rate of population growth.

In order to figure this out, we need block-level data, because we need not so much to estimate the growth of any particular block, but we need to have the flexibility to put different combinations of blocks together to understand and try to understand how much of the change in population is real growth, or real decline, and how much of it is simply changes due to the way they took the census.

So I think we are in a situation to say that the Census Bureau did an excellent job. They had a lot more money to spend. I think that the issue of Demographic Analysis is an important one. Unfortunately, while we spent so much money on the collection of the data, it appears we have underfunded the Demographic Analysis research at the Census Bureau.

My colleagues who have looked at this tell me that the Demographic Analysis agrees with the results of the A.C.E. pretty accurately for non-Hispanic blacks and non-Hispanic whites The difference lies in the Hispanic population, which is probably due to problems measuring immigration, which we know to be a very difficult thing to measure.

I will conclude at that point and answer questions, if there are any.

[The prepared statement of Dr. Ericksen follows:]

PREPARED STATEMENT OF EUGENE P. ERICKSEN, PROFESSOR OF SOCIOLOGY AND STATISTICS, DEPARTMENT OF SOCIOLOGY, TEMPLE UNIVERSITY

1. INTRODUCTION

I am a Professor of Sociology and Statistics at Temple University, where I have taught since 1971. I teach courses on survey design and methods, general statistics, and demography. Last year, I taught a graduate seminar on the United States census, with an emphasis on the nature, causes, and consequences of census error.

I completed my doctoral dissertation at the University of Michigan in 1971. In this dissertation, I developed a method for calculating local population estimates when census data are unavailable. The Census Bureau provided financial support for this work. Since completing the dissertation, I have done substantial research on methods of calculating population estimates, both in census and non-census years. Over the past two decades, I have extended my research into the area of estimating local undercounts.

In October 1980 I advised plaintiffs in a lawsuit filed by the City and State of New York attempting to compel an adjustment of the 1980 Census. Between October 1989 and July 1991, I served as Co-Chair of a Special Advisory Panel appointed by then-Secretary of Commerce Robert Mosbacher to advise him on the possibility of adjusting the 1990 Census. After completing my duties on the Advisory Panel, I again advised plaintiffs in a lawsuit seeking to compel the adjustment of the 1990 Census. For both the 1980 and the 1990 Census, I conducted substantial amounts of research on data relevant to the undercount.

In June 2000, the Census Monitoring Board, Presidential members, hired me as a statistical consultant. My duties have included reviewing Census Bureau plans and reports regarding the 2000 Census, and analyzing certain data from the census that the bureau has provided to us. In particular, I have had access to the individual level data records of persons included in the Accuracy and Coverage Evaluation (A.C.E.) samples used to estimate the levels of undercounting and overcounting in the 2000 Census.

2. BASIC CENSUS CONCEPTS

There have been three major controversies of the 2000 Census. The first, settled by a Supreme Court decision, was whether sampling could be used as part of Non-Response Follow-Up (NRFU). The second, decided by Secretary of Commerce Donald Evans, was whether the results of the census should be adjusted to correct for the differential undercount of minorities for redistricting purposes. The third, still undecided, is whether the adjusted results should be released for public use. I understand that the Census Bureau has already calculated block level adjusted counts for the entire nation.

These controversies are intimately related. To discuss them, it helps to define two terms, "net undercount" and "gross coverage error." The net undercount is the difference between the number of people counted in the census and the bureau's independently calculated estimate of the national population. The net undercount was 1.6 percent in 1990. The bureau estimates that it fell to 1.2 percent, or 3.3 million people, in 2000.

These 3.3 million people are not the total number of people missed from the census. The net undercount is the difference between two quantities, omissions and erroneous inclusions. Omissions are people who should have been counted, but were not. Erroneous inclusions are counts that should not have occurred. They are frequently duplications of the same person counted in the same place. They may also be double counts at separate addresses. For example, college students may be counted not only at the dormitory where they actually live, but also by their parents at home. A family with a second home, perhaps used for weekends and vacations, may get counted at each address. It is conceivable that there could be large, but equal numbers of omissions and erroneous inclusions. Should this occur, the net undercount would be zero, and useless for the evaluation of census error. If the omissions and erroneous inclusions occur at different locations, some local areas would have overcounts and others would have undercounts even though the national net undercount was zero. For example, in 1990, the Census Bureau estimated a 7.76 rate of undercount for Black renters living in New York City. In that same census, there was an overcount of 0.23 percent among "non-Hispanic White and Other" homeowners also living in New York City[1].

There were several million erroneous inclusions in both the 1990 and the 2000 Censuses. Since these must be subtracted from the number of omissions to derive the net undercount, the actual number of people missed from the census is much greater than the net undercount. The Census Bureau has asserted that 8.4 million persons were omitted from the 1990 Census. The comparable number for the 2000 Census is between 6.4 and 7.6 million persons (see Table 1).

The "gross coverage error" is the sum of omissions and erroneous inclusions. In 1990, the Bureau told us that there were 8.4 million omissions and 4.4 million erroneous inclusions. The net undercount was 4.0 million and gross coverage error was 12.8 million. In 2000, the Bureau indicates that the net undercount was 3.3 million and the gross coverage error was between 9.5 and 11.9 million people.

The national net undercount is not a good indicator of overall census quality. If omissions and erroneous inclusions occurred in the same places, many of them would offset each other. Because the geographic distributions of omissions and erroneous inclusions differ, they offset each other only partially. The gross coverage error is a better indicator of census quality.

---

[1] Howard Hogan, "The 1990 Post-Enumeration Survey: Operations and Results," *Journal of the American Statistical Association*, 88: 1047–1060, Table A.1.

Omissions tend to be concentrated among poor, typically minority, populations. They occur at especially high rates among the urban poor. Erroneous inclusions, while present among the poor, frequently occur among the affluent. In 1990, there were net overcounts among Whites living in owner-occupied housing units in the Northeast. In 2000, there were net overcounts among White homeowners in several post-strata located in the Northeast and Midwest.

Census taking is controversial due to the perception that some groups are disadvantaged because of counting errors. The best statistic for evaluating this possibility is the differential undercount. This is the difference in the net undercount between White and minority populations (see Table 2). In 1990, the net undercount for Hispanics was 5.0 percent, for non-Hispanic Whites it was 0.7 percent, so the differential undercount between Hispanics and non-Hispanic Whites was 4.3 percentage points. The net undercount for non-Hispanic Blacks was 4.6 percent and the differential between non-Hispanic Blacks and Whites was 3.9 percentage points.

The results of the 2000 Census show considerable improvement in this statistic. The net undercount for Hispanics was 2.8 percent, for non-Hispanic Blacks it was 2.2 percent, and for non-Hispanic Whites it was 0.7 percent. The differential between Hispanics and non-Hispanic Whites was 2.1 percent and between non-Hispanic Blacks and Whites it was 1.5 percent. These differentials for the 2000 Census are less than half of the comparable numbers for the 1990 Census. This is real progress, and I congratulate the Census Bureau.

However, from the data I have seen to date, it is clear to me that statistical correction of the census would improve the accuracy. There were a considerable number of errors in this census—at least 9.5 million by the Bureau's own account. The Bureau reduced the numbers of omissions more than the numbers of erroneous enumerations, and therefore reduced the net undercount. Although it reduced the differential undercount as well, important racial differences remain[2], and these could be removed by statistical correction.

3. THE STATUS OF THE CONTROVERSIES

The Supreme Court has settled the question of whether sampling should be used for Non-response Follow-up. We do not need to return to it now. The Secretary of Commerce, following the recommendation of the Census Bureau, has declared the unadjusted count to be the official result of the 2000 Census. We similarly do not need to return to this question. The remaining controversy concerns the release of adjusted block counts.

I believe that these should be released. There are three main reasons for this. One is that while the Census Bureau did reduce all of the net undercount, gross coverage error, and differential undercount rates—important differentials remain. The American Indian undercount, while less than it was 1990, is at 5 percent. Both the American public and the scientific community need to understand the effects of the undercounts on the census results they rely upon.

The second reason is that there are important ways in which the published census results appear to be incorrect. Ironically, the very improvement in census coverage that the Bureau accomplished creates error. There is confusion between the amount of actual growth and the amount of coverage improvements. Many localities, especially those with large minority populations, may have inflated impressions of the 1990–2000 growth rates. The best way to fix this problem is to compare adjusted 1990 to adjusted 2000 census estimates. The use of adjusted data will reduce the inconsistencies created by improved coverage in the 2000 Census.

The third reason is that both the American public and the scientific community need to have access to the adjusted counts to make their own evaluations on the nature of census error. There will be some purposes for which statisticians and other data analysts may deem it best to use the unadjusted counts, and other purposes for which they will prefer the adjusted results. The Bureau needs to release the adjusted data, along with their evaluations of these data, to permit these considerations.

We also need to have a better understanding of the geographic distributions of gross error. Study of the gross error will help us to learn where there were important problems of census taking, and where improvements might have taken place in the 2000 compared to the 1990 Census. For example, did the addition of addresses through LUCA reduce omissions? Finally, the scientific community needs to make its own evaluation of the bureau's estimates of the gross error rates.

---

[2] These differentials are found whether we rely upon the A.C.E. or demographic analysis.

4. STATE AND LOCAL UNDERCOUNTS

To illustrate the kinds of calculations and evaluations scientists need to make, I have calculated undercount estimates for each of the 50 states and the District of Columbia. I have also done this for five large cities, Atlanta, Chicago, New York City, Philadelphia, and Houston, and the remainders of the states in which they are located. I relied upon the A.C.E. data provided by the Census Bureau to the Census Monitoring Board. Because I do not have access to the P.L.94–171 census count data, I cannot match the estimates of these quantities that the Census Bureau would calculate. I should be close, however.

The state estimates vary across a narrow range. All but 10 states have estimates within one-half percentage point of the national average, 1.2 percent, i.e., and they are in the range of 0.7 to 1.7 percent. States with larger shares of minority population tend to have higher rates of undercount, and the opposite occurs in states with smaller minority shares. Holding the minority shares constant, rates of undercount were higher in the West and lower in the Midwest. The main result, though, is that between-state variations are moderate.

Turning to the cities, I have calculated rates of undercount for five of them. Although each of the cities includes large minority populations, the rates of undercount are lower than in 1990. Each of these cities has a higher rate of undercount than the remainder of the state in which it is located.

These calculations are possible with the limited amounts of data that the Census Bureau provided to the Census Monitoring Board. To calculate comparable estimates for smaller areas, and to calculate the state and large city estimates with greater certainty, we need to have the adjusted block level data.

5. ESTIMATES OF POPULATION GROWTH

With the reduction in the net undercount, especially of minority populations, we have difficulty interpreting rates of local population growth. Use of the currently unadjusted data leads to comparison problems. Because the level of undercount, especially in minority areas, was so much greater in 1990 than it is in 2000, use of the unadjusted results of the two censuses leads us to overestimate population growth. The problem is well illustrated by the case of New York City. Since 1990, Census Bureau population estimates have indicated small increases, with a total growth of 1.4 percent occurring between 1990 and 1999. Extrapolating to 2000, the expected population count was 7,452,184, an increase of 1.8 percent since 1990.

The actual 2000 Census count was 8,008,278, an increase of 9.4 percent since 1990. This amount is nearly 7 percentage points greater than the expectation. The result is implausible. Local experts believe that New York City has grown, but probably not by this amount. We cannot tell how much of the change is due to improved census coverage, and how much is real growth. In addition, because New York City added several hundred thousand addresses during the LUCA, this operation alone could account for much of the apparent growth. LUCA was not part of the 1990 Census.

The New York City story was repeated consistently across the nation. I have calculated 2000 population projections simply by extrapolating the 1998—99 population change forward to 2000. In Philadelphia, for example, the Bureau estimated a loss of 17,367 people between 1998 and 1999. I simply assumed a comparable loss between 1999 and 2000. The resulting projection of 1,400,234 is 7.7 percentage points below the 2000 Census count of 1,517,550.

I have made comparable calculations for all counties that had at least a 500,000 population in 1990 for which the Census Bureau has published 2000 Census tabulations of P.L.94—171 data (see Table 3). There are 66 of these counties and 60 of them have counts that are higher than the population projections than I calculated. Only 6 have counts that are below the projections. This skewed pattern of difference reflects the fact that the 2000 Census counts are more complete than those of the 1990 Census.

I have divided the counties into four categories, depending on the percentage minority in the 2000 Census. Among those counties where this percentage was less than 20, the average difference between the projected estimates and actual counts was moderate, 1.72 percent. Fourteen of the 16 population estimates were too low.

Turning to the second category, where the percentage minority was between 20 and 29.9 percent, there were 22 counties, and 19 of them had estimates that were too low. The average difference between the population projection and the actual count was 2.35 percent.

The third category included 16 counties where the percentage minority was between 30 and 49.9 percent. Fifteen of them had projections that were too low. The

average difference between the population projection and the actual count was 3.78 percent.

The fourth and final category included 12 counties where the percentage minority was greater than 50 percent. All of these counties had projections that were too low, and the average difference between the projection and the actual count was 6.66 percent. These counties were spread across the country, with 5 of them located in the Northeast, 6 of them in the South, and 1 in the West.

Overall, we see that the discrepancies between projections and counts are greater where the percentage minority is greater. To understand these confusing patterns of population growth and coverage improvement, we need to have both the adjusted and unadjusted block counts. With these in hand, we can create our own combinations of local jurisdictions to study the extent to which the apparent growth is real or the product of changed and improved methods of census taking.

6. SUMMARY

The 2000 Census is now complete, and its surprising results require analysis. The Census Bureau must complete its evaluation of the demographic estimate, and assess its consistency with the A.C.E. results. Demographers and statisticians across the country recognize this as a crucial project, and they would like to have their own looks at the data. In addition to understanding the predictors of undercount in the 2000 Census, the changing patterns of undercount between the 1990 and 2000 Censuses, we need to figure out why the demographic and A.C.E. results are so discrepant.

We also need to understand the patterns of gross and net error. Although I have not been able to carry out intensive research on the subject, much of the improvement in census coverage appears to be due to improvements in the census address list. The Local Update to Census Addresses program was a big contributor to this. If this were true, we would expect a substantial reduction in the number of "whole household omissions" in the 2000 Census relative to 1990. If the address list is improved, then the number of entire households missed should go down.

On the other hand, many omissions occur in households where others are counted. These "within household omissions" occur when people filling out census forms misunderstand the instructions, and such omissions are especially prevalent among persons distantly or unrelated to the heads of household. Because such errors are generally unrelated to the completeness of address lists, or even to the proficiency of advertising and other outreach programs, we would expect the numbers of such omissions to be similar to those observed in 1990.

Finally, we need good estimates of the numbers of omissions and erroneous inclusions. The Census Bureau has estimated that there were over 20 million nonmatches, but only 6.4 to 7.6 true omissions in the A.C.E. The Bureau also estimates that only 3.1 to 4.3 million of the 12 million apparent erroneous inclusions are real. Many of the apparent errors are due to missing data and other similar problems, and the numbers of apparent errors overstate the problems. However, the Census Bureau has not yet shared with us the logic behind these estimates.

There is a great deal of research that needs to be done to understand the 2000 Census. Some of it must necessarily be carried out at the Census Bureau, but the bureau must make public all data relevant to this understanding so that researchers and academicians can draw their own conclusions.

Table 1.—Estimated Error Counts from 1990 and 2000

[In Millions]

	1990	2000
Estimated total population	252.7	284.7
Census count	248.7	281.4
Measured net undercount	4.0	3.3
Measured Gross Erroneous Enumerations	4.4	3.1
Implied gross omissions	8.4	6.4
''Gross coverage error''	12.8	9.5
Assumed Errors in Reinstated ″Potential Duplicates″	.....	1.2
Total assumed gross erroneous enumerations	.....	4.3
Implied gross omissions	.....	7.6
Implied ''Gross coverage error''	.....	11.9

Source: Obtained from the Census Bureau on March 21, 2001.

Table 2.—Reduction in the Differential Undercount, 1990 to 2000

Group[1]	Net Undercount Rate		Difference from non-Hispanic Whites		Ratio, 2000 to 1990 Difference
	(1990)	(200)	(1990)	(2000)	
Hispanics	5.0	2.9	4.3	2.1	48.8%
Non-Hispanic Blacks	4.6	2.2	3.9	1.5	38.5%
Non-Hispanic Whites	0.7	0.7	.....	.....	.....

[1] Asians, Pacific Islanders, American Indians and Alaska Natives were not included in this table as these racial categories were treated differently in 1990 and 2000.

Source: United States Commerce Release Wednesday, February 14, 2001.

Table 3

Errors in Population Estimates for 2000 by Counties and Percent Minority

County	State	Percent Minority	1990 counted	1998 estimated	1999 estimated	2000 projected	2000 counted	Difference, projected and counted
Honolulu	Hawaii	78.72%	836,231	871,768	864,571	857,374	876,156	2.14%
Washington	District of Columbia	73.32%	606,900	521,426	519,000	516,574	572,059	9.70%
Prince Georges	Maryland	72.96%	729,268	776,907	781,781	786,655	801,515	1.85%
Bronx	New York	70.13%	1,203,789	1,191,319	1,194,099	1,196,879	1,332,650	10.19%
Baltimore City	Maryland	68.37%	736,014	645,664	632,681	619,698	651,154	4.83%
DeKalb County	Georgia	63.40%	545,837	592,870	596,853	600,836	665,865	9.77%
Kings	New York	58.80%	2,300,664	2,266,242	2,268,297	2,270,352	2,465,326	7.91%
Queens	New York	55.92%	1,951,598	1,993,172	2,000,642	2,008,112	2,229,379	9.93%
Essex	New Jersey	55.54%	778,206	748,322	747,355	746,388	793,633	5.95%
Philadelphia	Pennsylvania	54.98%	1,585,577	1,434,968	1,417,601	1,400,234	1,517,550	7.73%
Shelby	Tennessee	52.17%	826,330	867,804	873,000	878,196	897,472	2.15%
Fulton	Georgia	51.18%	648,951	737,222	744,827	752,432	816,006	7.79%
							Average	**6.66%**
New York	New York	45.64%	1,487,536	1,546,508	1,551,844	1,557,180	1,537,195	-1.30%
Hudson	New Jersey	44.42%	553,099	553,030	552,819	552,608	608,975	9.26%
Cook	Illinois	43.73%	5,105,067	5,192,396	5,192,326	5,192,256	5,376,741	3.43%
Jefferson	Alabama	41.90%	651,525	660,039	657,422	654,805	662,047	1.09%
Dallas	Texas	41.65%	1,852,810	2,045,309	2,062,100	2,078,891	2,218,899	6.31%
Harris	Texas	41.27%	2,818,199	3,202,021	3,250,404	3,298,787	3,400,578	2.99%
Suffolk	Massachusetts	39.60%	663,906	641,333	641,695	642,057	689,807	6.92%
Montgomery	Maryland	35.22%	757,027	839,158	852,174	865,190	873,341	0.93%
Mecklenburg	North Carolina	34.97%	511,433	630,813	648,400	665,987	695,454	4.24%
Milwaukee	Wisconsin	34.38%	959,275	911,536	906,248	900,960	940,164	4.17%
Cuyahoga	Ohio	34.10%	1,412,140	1,380,428	1,371,717	1,363,006	1,393,978	2.22%
Travis	Texas	31.79%	576,407	709,182	727,022	744,862	812,280	8.30%
Davidson	Tennessee	31.66%	510,784	533,258	530,050	526,842	569,891	7.55%
Middlesex	New Jersey	31.58%	671,780	712,638	717,949	723,260	750,162	3.59%
Bexar	Texas	31.14%	1,185,394	1,354,837	1,372,867	1,390,897	1,392,391	0.11%
Fairfax	Virginia	30.09%	818,584	927,895	945,717	963,539	969,749	0.64%
							Average	**3.78%**
Jackson	Missouri	29.90%	633,232	655,055	654,484	653,913	654,880	0.15%
Oklahoma	Oklahoma	29.56%	599,611	632,865	636,539	640,213	660,448	3.06%
Marion	Indiana	29.51%	797,159	812,662	810,946	809,230	860,454	5.95%
Camden	New Jersey	29.12%	502,824	504,268	503,093	501,918	508,932	1.38%
Tarrant	Texas	28.77%	1,170,103	1,354,040	1,382,442	1,410,844	1,446,219	2.45%
Westchester	New York	28.65%	874,866	900,861	905,572	910,283	923,459	1.43%
Clark	Nevada	28.42%	741,459	1,161,259	1,217,155	1,273,051	1,375,765	7.47%
Hamilton	Ohio	27.63%	866,228	847,202	840,443	833,684	845,303	1.37%
King	Washington	26.59%	1,507,319	1,654,329	1,664,846	1,675,363	1,737,034	3.55%
El Paso	Texas	26.05%	591,610	694,603	701,908	709,213	679,622	-4.35%
Baltimore	Maryland	25.61%	692,134	721,556	723,914	726,272	754,292	3.71%
Franklin	Ohio	25.57%	961,437	1,021,578	1,027,821	1,034,064	1,068,978	3.27%
Tulsa	Oklahoma	24.98%	599,611	543,417	548,296	553,175	563,299	1.80%
Montgomery	Ohio	24.13%	573,809	570,141	565,866	561,591	559,062	-0.45%
Pierce	Washington	23.96%	586,203	675,962	688,807	701,652	700,820	-0.12%
St. Louis	Missouri	23.17%	993,529	997,347	996,181	995,015	1,016,315	2.10%
Bergen	New Jersey	21.59%	825,380	854,428	857,052	859,676	884,118	2.76%
Jefferson	Kentucky	21.50%	664,937	671,595	672,900	674,205	693,604	2.80%
Fairfield	Connecticut	21.21%	827,645	837,476	841,334	845,192	882,567	4.23%
Monroe	New York	20.86%	713,968	714,936	712,419	709,902	735,343	3.46%
Multnomah	Oregon	20.84%	583,887	630,573	633,224	635,875	660,486	3.73%
Nassau	New York	20.70%	1,287,348	1,300,995	1,305,057	1,309,119	1,334,544	1.91%
							Average	**2.35%**

Table 3

County	State	Percent Minority	1990 counted	1998 estimated	1999 estimated	2000 projected	2000 counted	Difference, projected and counted
Lake	Illinois	19.89%	516,418	608,348	617,975	627,602	644,356	2.60%
Delaware	Pennsylvania	19.68%	547,651	542,592	541,502	540,412	550,864	1.90%
Erie	New York	17.82%	968,532	933,702	925,957	918,212	950,265	3.37%
Summit	Ohio	17.00%	514,990	537,160	537,856	538,552	542,899	0.80%
DuPage	Illinois	15.95%	781,666	880,996	892,547	904,098	904,161	0.01%
Allegheny	Pennsylvania	15.67%	1,336,449	1,267,963	1,256,806	1,245,649	1,281,666	2.81%
Monmouth	New Jersey	15.61%	553,124	603,214	611,444	619,674	615,301	-0.71%
Suffolk	New York	15.40%	1,321,864	1,370,549	1,383,847	1,397,145	1,419,369	1.57%
Montgomery	Pennsylvania	13.54%	678,111	719,569	724,087	728,605	750,097	2.87%
Middlesex	Massachusetts	12.15%	1,398,468	1,422,465	1,426,606	1,430,747	1,465,396	2.36%
Essex	Massachusetts	11.66%	670,080	700,370	704,407	708,444	723,419	2.07%
Salt Lake	Utah	11.38%	725,956	845,913	850,243	854,573	898,387	4.88%
Norfolk	Massachusetts	9.74%	616,087	642,089	643,580	645,071	650,308	0.81%
Worcester	Massachusetts	8.74%	709,705	730,769	738,629	746,489	750,963	0.60%
Bucks	Pennsylvania	7.54%	541,174	587,863	594,047	600,231	597,635	-0.43%
Bristol	Massachusetts	6.84%	506,325	516,975	520,258	523,541	534,678	2.08%
							Average	**1.72%**

Source: www.census.gov/People estimates/County Population Estimates/ County Population Estimates for July 1, 1999 and the Population Change from July 1, 1998 to July 1, 1999/American FactFinder/Race Hispanic or Latino (2000)

Senator KERRY. Thank you very much.
Dr. Wachter.

STATEMENT OF KENNETH W. WACHTER, PROFESSOR, CHAIR, DEPARTMENT OF DEMOGRAPHY, UNIVERSITY OF CALIFORNIA

Dr. WACHTER. I am joined in this testimony by my colleague, David Freedman. We think there is widespread agreement on two chief points: First, Census 2000 succeeded in reducing differential undercounts from the 1990 levels. Second, there were serious questions about the accuracy of proposed statistical adjustments. The bureau advised the Secretary to certify the unadjusted counts, and we concurred, as did the Secretary's other outside advisors.

As you have heard, Demographic Analysis and A.C.E. point in opposite directions. Demographic Analysis is precious, because it is independent of the census and A.C.E., tested over time, and rests on fewer and simpler assumptions than A.C.E. While DA is hardly perfect, analysis already shows that it would be a stretch to blame Demographic Analysis for the whole or most of the discrepancy with A.C.E.

Mistakes in statistical adjustments in the census are nothing new. Our studies, David Freedman's and my own, on 1980 and 1990 data, have described three kinds of error, processing error, correlation bias, and heterogeneity. In the face of these errors, it is hard for adjustments to improve on the accuracy of census numbers for states, for counties, legislative districts, and smaller areas.

Statistical adjustments can easily put in more error than they take out, because the census is already very accurate. The changes, as Gene Ericksen mentioned, to state counts from adjustment are very small, measured in dozens of parts per million when you look at those changes, so the chief question is, what went wrong with A.C.E. in 2000?

Errors in responses to the survey, or in the statistical operations, may from some perspectives have been under better control than in 1990, but it appears that processing errors must have been worse in other respects. Research is underway to pinpoint the difficulties.

The bureau is investigating a form of error called balancing error. We suspect that troubles also occurred in a new treatment of movers, and in detection of duplicates, which were especially numerous in 2000, and which occur for minority populations as well as for nonminorities.

Correlation bias and heterogeneity are endemic problems that make it extremely difficult for adjustment to improve on the census. Correlation bias is the tendency for people missed in the census to be missed by A.C.E. as well. Correlation bias in 2000 seems to have amounted, as it did in 1990, to millions of persons. These people cannot be evenly distributed across the country, so statistical adjustments create a distorted picture of census undercounts.

Heterogeneity means that undercount rates differ from place to place within population groups treated as homogenous by adjustment. Heterogeneity puts limits on the accuracy of adjustments for areas like states, counties, or legislative districts. Our studies,

along with recent work at the bureau, show that heterogeneity remains a serious concern.

We must keep in mind that the census is used to distribute representation and resources directly among places, not ethnic groups. For A.C.E. to find an undercount of Hispanics or African Americans or Asian Americans, does not help if A.C.E. puts the missing people in the wrong places, and that is what A.C.E. is likely to do.

Census 2000 achieved a high level of accuracy. Given that, and given the problems with statistical adjustments, the Secretary s decision to certify the census counts was the right decision.

Thank you.

[The prepared statement of Dr. Wachter follows:]

PREPARED STATEMENT OF KENNETH W. WACHTER, PROFESSOR, CHAIR, DEPARTMENT OF DEMOGRAPHY, UNIVERSITY OF CALIFORNIA

I am joined in this testimony by my colleague David Freedman. We think there is widespread—although by no means universal—agreement on two chief points. First, Census 2000 succeeded in reducing differential undercounts from their 1990 levels. Second, there are serious questions about the accuracy of proposed statistical adjustments. The bureau advised the Secretary to certify the unadjusted counts and we concurred, as did the Secretary's other advisors.

Statistical adjustment faced a new problem in Census 2000. Independent population estimates are derived by Demographic Analysis from administrative records, including birth and death certificates and Medicare files.[1] These estimates show the Census overcounted the population by perhaps 2 million people. Proposed statistical adjustments would have added another 3 million people, making the overcounts even worse. Demographic Analysis is independent of the Census and the A.C.E. survey which underlies proposed adjustments. Demographic Analysis and A.C.E. point in opposite directions. While Demographic Analysis is hardly perfect, it is a stretch to blame Demographic Analysis for the whole of the discrepancy with A.C.E.[2]

Mistakes in statistical adjustments to the Census are nothing new. Our studies of the 1980 and 1990 data have described three kinds of error: processing error, correlation bias, and heterogeneity.[3] In the face of these errors, it is hard for adjustments to improve on the accuracy of Census numbers for states, counties, legislative districts, and smaller areas. Statistical adjustments could easily put in more error than they take out, because the Census is already very accurate.

What went wrong with A.C.E. in 2000? Errors in responses to the survey or in the statistical operations may from some perspectives have been under better control than they were in 1990. But, it appears, processing errors must have been worse in other respects. Research is underway to pinpoint the difficulty. The Bureau is investigating a form of error called balancing error.[4] We suspect that troubles also occurred in a new treatment of movers and in the detection of duplicates, which were especially numerous in 2000.

In July 1991, the Bureau recommended adjusting Census 1990 by adding 5.3 million people: processing errors were estimated at 1.7 million (these figures are net, nationwide). The figure for processing error later increased to 3.0 million, although independent estimates range up to 4.2 million. Our estimate is 3.6 million (Wachter and Freedman, 2000a).

In 2000, A.C.E. would add 3.3 million persons to the census count. The preliminary estimate of processing error is 2 million, as may be seen by doing some arithmetic on the percentages in Table 24 of the B–19 Report. That table allows 1 million for correlation bias. However, the underlying model (B–12) repeats the error discussed in Wachter and Freedman (2000a). Compare page 16 of B–12, although page 46 of the B–1 Report acknowledges the problem.

Thus, the estimate for correlation bias needs to be increased, perhaps by another million or so. If the population estimates from Demographic Analysis are approximately correct, the estimates for processing error will need to increase by several million, as in 1990.

---

[1] The bureau's estimates from Demographic Analysis are presented in the B–4 Report; see especially Appendix Table 2. The estimated total national populations are as follows: Demographic Analysis: 279.598 million; Census 2000: 281.422 million; A.C.E.: 284.684 million.

[2] The Demographic Analysis estimates for net undocumented immigrants and some other categories of non-citizen residents may be somewhat low. But, as the B–4 Report spells out on page 11, implausibly large revisions would be required to bring the totals into agreement with A.C.E.

[3] See, e.g., Freedman and Navidi (1992), Freedman and Wachter (1994), and Brown et al. (1999).

[4] The Bureau's research on balancing error is summarized on pages 24-25 of the ESCAP report.

Correlation bias and heterogeneity are endemic problems that make it extremely difficult for adjustment to improve on the Census. Correlation bias is the tendency for people missed in the Census to be missed by A.C.E. as well. Correlation bias in 2000 may have amounted, as it did in 1990, to millions of persons. These people cannot be evenly distributed across the country. If so, statistical adjustments create a distorted picture of census undercounts.[5]

Heterogeneity means that undercount rates differ from place to place within population groups treated as homogeneous by adjustment. Heterogeneity puts limits on the accuracy of adjustments for areas like states, counties, or legislative districts. Our studies, along with recent work at the Bureau, show that heterogeneity remains a serious concern.[6]

Census 2000 achieved a high level of accuracy. Given that, and given the problems with statistical adjustments, the Secretary's decision to certify the census counts was the right decision.

REFERENCES

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

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

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.

Cohen, M. L., A. A. White, and K. F. Rust, eds. (1999). *Measuring a Changing Nation: Modern Methods for the 2000 Census*. National Academy Press, Washington, D. C.

Fay, R. E. and J. H. Thompson (1993). The 1990 Post Enumeration Survey: statistical lessons, in hindsight. *Proceedings, Bureau of the Census Annual Research Conference*. Bureau of the Census, Washington, D. C.

Freedman, D. A. and W. C. Navidi (1992). Should we have adjusted the U. S. census of 1980? *Survey Methodology* 18: 3-74 (with discussion).

Freedman, D. A., P. B. Stark, and K.W.Wachter (2000). Aprobability model for census adjustment. http://www.stat.berkeley.edu/census/557.pdf To appear in *Mathematical Population Studies.*

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

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

Prewitt, K. (2000). Accuracy and Coverage Evaluation: statement on the feasibility of using statistical methods to improve the accuracy of Census 2000. *Federal Register* 65: 38373-38398.

Skerry, P. (2000). *Counting on the Census*. Brookings, Washington, D. C.

U.S. Census Bureau (2001). Report of the Executive Steering Committee for Accuracy and Coverage Evaluation Policy. Cited as the ESCAP report. The supporting documentation consist of reports B–1 through B–19. All these reports are available on the web. http://www.census.gov

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

Wachter, K.W. and D.A. Freedman (2000b). Measuring local heterogeneity with 1990 U. S. Census data. *Demographic Research*, an on-line journal of the Max Planck Institute. Volume 3, Article 10. http://www.demographic-research.org/Volumes/Vol3/10/

Senator KERRY. Thank you very much, doctor. Doctor, in your testimony I do not think you addressed the question of the immigration issue, in the DA. I mean, it is my understanding DA does not measure immigration, it measures births and deaths.

Dr. WACHTER. That is right. The immigration component is an assumption that is built in.

---

[5] Brown et al. (1999),Wachter and Freedman (2000a).

[6] Adjustment assumes that coverage rates (i.e., rates of census undercount or overcount) are constant within population groups called "post strata," across wide stretches of geography. Failures in this assumption are called heterogeneity, or called "synthetic error" in the B-studies. The Bureau's research on heterogeneity is summarized on pages 22-24 of the ESCAP report. Our work is described in Freedman andWachter (1994) as well asWachter and Freedman (2000b).

Senator KERRY. But isn't it critical? I mean, we have read article after article in the papers about increased numbers of immigrants, particularly in urban areas, and more even in rural areas than we previously thought, so would you not acknowledge that there is an impact on the DA by immigration?

Dr. WACHTER. On the Demographic Analysis?

Senator KERRY. Yes.

Dr. WACHTER. Yes. In my complete testimony in footnote 2 we take up this question. We believe that Demographic Analysis estimates for net undocumented immigrants and some other categories of noncitizen residents may be somewhat low. But there are constraints from data on the percent foreign-born from the Current Population Survey and elsewhere, which put bounds on the plausible adjustments that you can make.

We think you might be able to bring the Demographic Analysis estimates up closer to the census counts to make the apparent overcount in the census go away. But to make those estimates come up anywhere near the A.C.E., to suggest that there is not some set of major errors in A.C.E., seems really a great stretch.

There is not enough scope in plausible estimates for undocumented immigration. The apparent overcounts of some groups of African Americans, younger African Americans, according to some constructions of the multiple race issue also put constraints on how much leverage you can get. So I do not know of anyone who believes the Demographic Analysis can be plausibly brought up to agree with A.C.E. It may be brought up to agree with the census. To bring it up to agree with A.C.E. is too much.

Senator KERRY. If the Demographic Analysis had reported a higher number greater than A.C.E., would you have supported using corrected data?

Dr. WACHTER. No, I would not. The major uses of the census are in distributional matters. The important issue is in accuracy of distribution, and there are these endemic errors, including correlation bias, which argue that A.C.E. or suggest that A.C.E. would be putting people systematically in the wrong places.

Representative Maloney has departed, but New York is a good example. By our estimates the population share of New York State would be brought down by adjustment, despite the strong presence of urban central city minority groups. And this would have happened in 1990 also. We believe that the culprit is likely correlation bias, that this adjustment in New York State would be a mistake. It would carry the population shares in the wrong direction, because of distributional errors in the A.C.E. survey.

Senator KERRY. Dr. Ericksen, how do you respond to both answers with respect to the DA?

Dr. ERICKSEN. Well, first of all I think that there is a very serious problem in the level of research that has been done on Demographic Analysis. There simply has been enough done, but we may be facing a situation where a good demographic estimate of the population simply is not possible, because after all, undocumented immigration is just that, and we can try to make assumptions after assumptions after assumptions to change the number.

Senator KERRY. Is it a mistake to make assumptions because it is undocumented?

Dr. ERICKSEN. It depends on the amount of ancillary data you have. Now, back in 1980 and 1990, demographers were able to make good estimates of the amount of undocumented immigration by looking at things like births and deaths in the areas where undocumented people were concentrated.

I am told by demographers looking at this problem that it is more difficult to do now, and so I think there has to be a very large question mark to the use of Demographic Analysis for anything having to do with an immigrant population. I just do not think we know.

As to the other issue, I think that I go back to my written testimony. In the state of New York you have both extremes. You have minority areas in New York City where the undercounts are very high. You also have non-Hispanic white areas outside of New York City where you have just the opposite situation.

Now, how those two factors balance out is an interesting question. I do not know the answer to that question at this moment, but it is not implausible that they should be canceling out to bring New York State very close to, or just a little bit above, just a little bit below the national average.

Senator KERRY. Now, you have heard a number of people today cite the difficulty and inadvisability of releasing the A.C.E. data in order to permit you and others to make some calculations, would you comment on that question of advisability and necessity?

Dr. ERICKSEN. Well, I think there are two different points to make. First, in terms of general policy, we actually have had some debate in here today about how do we know the census is better. Well, we know the census is better because of the A.C.E., but we do not want to use the A.C.E. because we think it has errors in it. We go around and around on that.

I think the way to get out of that vicious circle is to do an analysis. If the census is better in 2000 than it was in 1990, why did that happen? What were the crucial decisions that people made?

I think that I have identified two of those decisions. One decision was the use of the local update of the census address program. We need to evaluate that. The only way that we can evaluate that is to have local data to compare places where that program was successful, and where it was not put into place.

Secondly, there was a decision to close data collection earlier and have a broader use of imputation. That also occurred on a local basis. The only way we could evaluate the result of that is on a local basis, so we need to have the block level data so we can understand the effects of that, and on the undercount, but we have to be able to aggregate the blocks in different ways.

The other part of my answer is that there are a tremendous number of localities across the United States who do not understand the census data, and they cannot understand their census data because they do not know how much of the change is due to real change and how much of it is due to the fact that the net undercount of this census is smaller than the net undercount of the last one.

Senator KERRY. What do you say to the notion that it is not ready, that it is premature, that they need to do more work with it?

Dr. ERICKSEN. Well, my understanding is the block level data have already been calculated, and so they could be given out tomorrow.

Senator KERRY. So you are saying that that is just not a legitimate answer?

Dr. ERICKSEN. Well, I think that there might be a misunderstanding here, because there are two issues. One is, could the block level data be released? The answer—I understand they have all been calculated. They could be released. Two, are the block level data more accurate? My understanding there is, the Census Bureau is contending that the research on the question needs to be done, and we will see what their answer is.

Senator KERRY. Is there any harm done by letting other people review that data now?

Dr. ERICKSEN. I think every one understands that the final decision on accuracy will be made sometime in the future, so as long as people are cognizant of that, I do not see that there could be any harm.

Senator KERRY. Do you have any way to make a judgment whether additional research will show that the A.C.E. is more accurate as a way of measuring?

Dr. ERICKSEN. Well, I am very cognizant of my colleague to my left, Ken Wachter. I am very cognizant of the opinions of other colleagues who feel differently. I think it is an open question.

Senator KERRY. Dr. Wachter.

Dr. WACHTER. It is an open question, and we need to research it. I would presume that one issue in the minds of the Census Bureau in terms of the early release of the block level adjusted data is the possibility that there is something like a computer coding error, which we had in 1990, in the A.C.E. estimates, and if there is, they would like to track it down before they put all these numbers out for the many uses that they would have. I do not have a strong view myself on the release or nonrelease.

Senator KERRY. What is a fair amount of time within which one ought to be able to do that? When would it be fair for people to have an expectation that this data could be released so we could avoid this controversy?

Dr. WACHTER. I should say I am not opposing the immediate release. I think there are scientists who——

Senator KERRY. Dr. Murray, you are opposing immediate release?

Dr. MURRAY. Senator, I want to be very careful here. For research purposes there are great uses and values to having A.C.E. data available. You have to ask the question, for which purpose, and it is important to recognize that there is some sense in which the A.C.E. data are really not finished, finished in the sense of, they are attached, each of them, each of the numbers that could be provided has error bars attached to it, has a standard error that is associated with it that is a fluctuation.

We have been, I think, provided a couple of misapprehensions in this discussion about the standing of the A.C.E. data, and if I might, just for a moment, explain, it has been treated as if it was only a matter of time, and if only we had more time, things would have been resolved.

The actual ESCAP report, the Executive Steering Committee on A.C.E. Policy, recommended to the bureau that they not use the adjusted A.C.E. data, and the bureau Director, Mr. Barron, also said very clearly we did not run out of time, we ran out of data. What we have is an irreconcilability, and we may in the future be able to resolve that.

Senator KERRY. I am puzzled by that.

Dr. MURRAY. Let me find a cite for you that says that.

Senator KERRY. I thought they specifically said that they needed to more fully investigate. They just needed more time.

Dr. MURRAY. That is part of the issues, Senator. I appreciate you reading it directly. Let me point to errors in the way we have had this entered into the testimony. We have heard so far today that the committee believes it likely that further research may establish that the adjustment based on the A.C.E. would result in improved accuracy. However, the very next sentence, which was not quoted, says, however, the uncertainty due to these concerns is too large at this time to allow for recommendation on adjustment.

Now, let us take that line over into the last page of the A.C.E. statement where they say, the ESCAP is unable to conclude at this time that the adjusted data are superior because—and here is the critical concern—further research on these concerns could, in fact, reverse the finding of the adjusted data's superior accuracy. It is indeterminate.

Senator KERRY. I understand that. Nobody questions that, but the question is, why can't other people be engaged in that similar research so that there is a sort of countercheck?

Dr. MURRAY. What our concern, Senator, is that—I think the concern here is that we cannot presume that the adjusted data are somehow superior, or will be found to be superior.

Senator KERRY. I am not making that presumption.

Dr. MURRAY. What we might find is, the data are actually less reliable than we expected, and here, as Dr. Wachter has said very clearly, redistribution, the issue of distributive accuracy, of getting the data correct at the levels of the local governments where they are actually applied, is the census' primary responsibility, and the issue with the A.C.E., the concern we have is that the A.C.E. does not put back the undercounted where they were actually lost. The A.C.E. is indeterminate.

Senator KERRY. We have a confidence question here about whether or not the undercount will be accurately reflected, put back by whatever methodology is arrived at, and therefore the question is, when is the appropriate time for that data to be judged by people outside who do not have the same interests, or sort of historical investment, if you will, in the process, who will look at it independently and help give confidence in the numbers.

I would assume—it is a question I would ask of both you and Dr. Wachter, that if you have 6.4 million considered to be undercounted, do you consider that very accurate, leaving out 6.4 million people?

Dr. MURRAY. Mercifully, Senator, it is an improvement.

Senator KERRY. That is not what I asked you.

Dr. MURRAY. It is a gross error, and it is troubling to me, and let me explain. It was my privilege, Senator, to live 9 years in Massachusetts, and you represented me.

Senator KERRY. Well, being represented by me and living there means you can say whatever you want, and I will forgive you.

[Laughter.]

Dr. MURRAY. Thank you, sir. I learned something very directly about you, which was your own passion and commitment to the problems and needs of the undercounted communities, the underserved and underprivileged communities. Senator, let me assure you, I share that, and have tried in every way possible to address the problem that we face, the social problem of the undercounted people. The heart of the question is, what is the best mechanism, the best instrument that will help us reach and encourage and incorporate those people and give them a fair share and bring them into the system?

Senator KERRY. I completely agree.

Dr. MURRAY. One of the difficulties is that the A.C.E.—oh, would that it did work according to its promise, but it represents the triumph of hope over experience to think that A.C.E. actually solves the problem. It has the tendency to introduce more error, to actually add people back into already overcounted communities and worsen the problem, and to fail to put people back in the undercount. It does not seem to solve our issue.

Senator KERRY. You are eloquently and forcefully making a case, but you are simultaneously, it seem to me, making a case for having this data analyzed by more people in order to come up with a fair methodology.

Dr. MURRAY. There is no principled or inherent reason not to evaluate the data.

Senator KERRY. My question is, when can we anticipate fairly that whatever adjustments need to be made, or judgments about the data, so that it could be released in order to permit that review before we all wind up with a decision that has been made and there is no recourse? Could I anticipate—it is now the end of March, or April. Could we have it by May?

Dr. MURRAY. In my discussions with bureau professionals who are doing this task, they anticipate the need for more data from the long form that will help them understand the foreign-born population, for readdressing this problem about whether or not Demographic Analysis is accurately capturing the undocumented, and that they have spoken about a target of potentially this fall having a better grasp of the problem one way or the other, and it seems to me to be reasonable.

Senator KERRY. This fall is too late, which is why we have engaged the statistical sampling.

Mr. Vargas, what would that mean to you?

Mr. VARGAS. Mr. Chairman, let me first thank you for your leadership on this issue.

It would have a tremendous impact on the Latino population. Let me just say that as a representative of the 33 million Latinos out there who have been counted in the census, that we are speaking for the 1 million that were missed, and these are 1 million who would not have fair representation in congressional districts and in

state legislative districts. It would result in malapportioned districts, districts in the Central City of Los Angeles, for example, that would be inherently larger than other districts in other parts of the state.

Let me also bring to your attention, Mr. Chairman, another issue. I did not mention in my oral testimony that I am a member of the Secretary of Commerce's Advisory Committee on the Decennial Census.

With respect to this subsequent evaluation that the Census Bureau professionals will be doing with respect to the A.C.E., I asked that when they make that recommendation to the Acting Director, who will make the decision as to whether or not to further release adjusted data? The Acting Director of the Census Bureau? Mr. Lee Price, who is the Acting Under Secretary of Commerce, plainly said that they did not know whose authority it will be to decide whether or not in the fall to release adjusted data.

They indicated that Secretary Evans' decision to rescind the rule that his predecessor put in place, Secretary Mineta, to allow the Department of Commerce—I'm sorry, the Census Bureau Director to make the decision on the release of the data only pertained to data for redistricting purposes, that if the professionals at the ESCAP committee recommend to release adjusted data, that they did not know whose authority it would be to approve that release. They did not know if it would be Acting Director Barron's decision or the Secretary's decision.

I think this is an issue that certainly this Committee should investigate, as to who ultimately has the authority to decide on whether or not——

Senator KERRY. If I had had more time to be here with the Secretary, that is one of the questions I will ask, but we will submit that question to him to try to determine that as a part of the record of this hearing.

Before we wrap this up, Dr. Ericksen, Dr. Wachter, what is your sense of what the timing might be able to be that would be able to be effective here so we could try to not wind up—I mean, look, my goal here is not to—I do not want an inaccurate count, but I do not want to live with the estimate for 10 years that 6.4 million of our citizens were left out.

President Bush has said he wants to leave no child behind. If you do not count them, you do not have a prayer of not leaving them behind.

Now, it seems to me there is just a fundamental judgment to make here. If everybody has agreed there is an undercount, and I have heard nobody suggest otherwise, there has to be a way through smart application. I mean, you can improve the census any number of ways, but statistical sampling has been the most accepted scientific method of trying to do it, so you are not just playing Solomon and trying to cut it down the middle.

Now, what could we anticipate, fairly speaking, so that the transparency of this is sufficient that people will not feel that a Republican administration and a Republican-run Commerce Department is trying to do what people saw a Republican administration do previously in the Commerce Department when they denied the use

of statistical sampling? Everybody knows what the impact of that is.

Now, I want to try to get away from the political gamesmanship here. I do not want this to be political. I am not trying to have a Democratic outcome, I am not trying to have a Republican outcome. I would like an American outcome, which is to count American citizens as fairly as we know how.

Now, is there some fair way for both of you, as sort of representing—I mean, one is a Republican Member, one is a Democratic appointee, but is there a way to come at this and have some agreement about when the statistics could be subject to scrutiny at large, so that people could make some judgments about the numbers that the American people have paid for to be gathered by an American agency? What do you think?

Dr. WACHTER. I do not think either Gene nor I really comes from a partisan political viewpoint on this, and we have often worked together in the Special Advisory Panel of 1990.

Senator KERRY. Do you know when that data might be able to be subject to scrutiny?

Dr. WACHTER. For the 1990 census, by the beginning of the summer of 1991 most of the relevant data that we worked with was there in good form. Isn't that right, Gene?

There are apparently certain kinds of larger problems with A.C.E. than there were with the Post Enumeration Survey in 1990, so I have heard people talk about the summer. But as I said, from the point of view of scientists like ourselves working with the data, I would have no opposition to seeing the immediate release.

Senator KERRY. Dr. Ericksen.

Dr. ERICKSEN. I think it is a basic principle of science that we make our data publicly available. Just to give you an example, the Journal of Public Opinion Quarterly has a policy that we not publish an article unless the author of the article makes a questionnaire available.

In my view, the data should have been made publicly available the day the monitoring board got them back in March. There is only going to be improvement for having more people enter this debate. The more I analyze the data, the more I see questions to ask the Census Bureau. For example, in every state that I have looked at that has sizable numbers of blacks and Hispanics, the net undercount there of those groups appear to be greater than they are of non-Hispanic whites.

That is a very important question to know, because we want to know about consistency, so the data in my view should have been released in March or February, whenever they became available. There is no harm whatsoever in engaging in the scrutiny. I also think that people in city planning offices across the country would be very interested in looking at the effects of a possible adjustment on their localities.

That is a different issue than making the decision. I think the mistake is to think nobody can look at the data until the decision has been made, and we are arguing the opposite.

Senator KERRY. Well, I am arguing the opposite very strongly. It seems completely inconsistent to me to be making the statements that this is the most accurate census we have ever had, and to base

your judgment that it is the most accurate census you have ever had because it is supported by the A.C.E., but you are not wiling to let the information on which you make that judgment be scrutinized.

So I am going to press very hard with colleagues and others to see if we can open this debate up. I mean, this is an important debate, and I have never known legitimate information, statistics to be available to people. I know they can distort them, and I know they can twist them, but there are a lot of folks out there who can analyze this and begin to make some judgment about accuracy and about where we go.

That is the nature of the American process, is to allow more information, not less, particularly information that number 1, that the American people paid for and pulled together, and number 2, that the American people are being asked to live with for 10 years that has a profound impact on their communities and their planning, and lots of other public officials who would like to, I think, have independent analyses of what happened in their community. So I think it is very, very important to try to get that information available, and to continue to have this debate.

I assume if we do not attempt t resolve this issue now, it will be with us for 10 years, with a lot of acrimony and divisiveness that is really unnecessary that should not accompany something like this effort to simply count the people who are here living in this country.

So with that said, unless there is an urgent need to add something, at this point I will leave the record open for 7 days for any colleagues who may have additional questions, or for any additional testimony we want to gather in writing, and I thank you all very, very much for taking time and for participating today.

Thank you. We stand adjourned.

[Whereupon, at 11:55 a.m., the hearing was adjourned.]

APPENDIX

RESPONSE TO WRITTEN QUESTIONS SUBMITTED BY HON. JOHN MCCAIN TO DONALD L. EVANS

Question 1. Mr. Secretary, your testimony highlights a large reduction in the undercount of minorities by the Year 2000 Census. However, some minority groups still complain that the Census does not adequately count them. Could you please describe what steps the Census Bureau took to reduce the minority undercount, and how the Census Bureau intends to improve upon its 2000 results?

Answer. The Department of Commerce is gratified by recognition of the strides made by the Census Bureau in reducing the undercount of minorities and others in Census 2000. This was achieved by implementing many new and expanded procedures to improve participation in the census, including: expanded partnerships; paid promotion, allowing both a visible national campaign and promotion targeted at specific population groups; increased availability of census forms and ways to respond; involvement of local/tribal governments in updating the address listing to improve housing unit coverage; tailoring special enumeration procedures to population groups and geographic locations; enhanced language program; expanded recruitment of census workers.

Further reductions in the undercount need to be addressed on several fronts. While the Census Bureau is proud of the results of Census 2000, it intends to improve upon them. This will be increasingly challenging because the population is becoming progressively more diverse.

The Census Bureau plans to reengineer the entire census process, with improved coverage as a major objective. The reengineering will be based on three strategies: A significantly enhanced and improved Master Address File and the geographic data base supporting it; Long form data collected and tabulated every year by the American Community Survey, and therefore not included as part of the 2010 Census; Early planning and design of the 2010 Census to fully take advantage of opportunities made possible by the other two strategies.

To achieve these goals we will need the support and endorsement of Congress, in addition to the necessary funding.

Question 2. The next major decision that you will have to take regarding the Census is whether or not to release adjusted data for federal fund allocation. I understand that you intend to make this decision in the middle of October. What processes have you set up and criteria you will consider as you make this decision?

Answer. The Census Bureau has prepared a plan for continuing the analysis of the Census 2000 data over the next several months, a process which will lead to a recommendation next fall concerning the use of adjusted data for use the intercensal population estimates program and for sample controls in demographic surveys. These data may also be used to allocate federal funds under various grant programs.

The plan contemplates that the Census Bureau will continue with its investigation into the quality of the Accuracy and Coverage Evaluation (A.C.E.), the demographic analysis, and Census 2000. That investigation will rely on new data collected from the long form, as well as more in-depth reviews of the data and methodologies that supported the recommendation of the ESCAP Committee last March to use the unadjusted data for redistricting purposes.

RESPONSE TO WRITTEN QUESTIONS SUBMITTED BY HON. SAM BROWNBACK TO DONALD L. EVANS

Question 1. For the first time since 1880, the 2000 Census short form did not gather data on marital status. As a result, important information on family structure, which is a vital indicator of child well-being, was lost. Last year, the Senate voted unanimously in favor of a resolution expressing support for restoring the mar-

ital status question to the short form. Does your office plan to resume the gathering of marital status data on the short form of 2010?

Answer. For Census 2000, in compliance with directions from Congress to reduce the number of questions asked on the decennial questionnaires, the Census Bureau undertook a comprehensive review of each question contained on both the short and long form questionnaires, eliminating those which were not specifically required by law. Through this process, the Census Bureau moved five questions from the short form to the long form because these questions produced data that were not needed at the block level. One of those questions was the item on marital status. The Census Bureau submitted its proposed questionnaire content for Census 2000 to Congress on March 31, 1998. Following this submission, the Census Bureau fully considered all expressed concerns about question topics before finalizing the questionnaire content. The Census Bureau received no comments regarding the content of the marital category, and thus left that question on the long form only. When the Senate's resolution on the marital status question was passed, it was too late to consider changes to the content of census questionnaires.

The formal planning work for the 2010 census again will undertake a comparable large-scale effort to elicit information from Congress and the public on what data should be collected in the next decennial census. In addition, the American Community Survey (discussed in more detail in a later question), if funded, will provide important new and more relevant data on family structure in America.

Question 2. It is my understanding that collection of some vital statistics data from the states can be spotty, and that the marital status question on the short form of the Census was used to flesh out those numbers. Some demographers have claimed that removing the question from the short form may impair the ability to make accurate state and local estimates of the married population. Is this a concern?

Answer. The Census Bureau expects no data quality deficiencies resulting from having the marital status question on the long form. The Census Bureau received over 17 million long forms and believes that marital status information from the long form, which is provided down to the census tract level, will be sufficiently accurate to provide demographers concerned with state and local estimates with the data necessary for their statistical needs.

Question 3. What specific steps are being taken to ensure that the American Community Survey (ACS) improves and enhances data collection on family structure in America?

Answer. Several questions in the ACS measure family structure, including questions regarding marital status, relationship, and births in the past year. The ACS has many advantages over a single day snapshot taken every ten years because the ACS provides more detailed measures of change. In addition, the quality of data collection, coding, and processing will improve through the use of a permanent, well-trained staff and continuous data collection.

Through the richness of the ACS content and the ability to create estimates for small geographic areas, America will he able to study a variety of family issues, such as matters related to children living in homes with a disabled veteran parent, children relying on the financial support of grandparents, characteristics of children living with immigrant parents who are not proficient in English, and other matters. The Census Bureau expects the ACS to produce quality data to answer these questions at all important levels, including the national, state, city and town levels.

RESPONSE TO WRITTEN QUESTIONS SUBMITTED BY HON. JOHN MCCAIN TO KENNETH W. WACHTER AND DAVID A. FREEDMAN

Question 1. You have just given testimony concerning possible causes of problems in the A.C.E. methodology. How would you design a method for adjusting the Census data to account for correlation bias, heterogeneity, and processing errors?

Answer. Given the small level of error in the 2000 Census, as presently estimated, we do not believe that adjustments can be designed which would improve on the Census counts.

Question 2. The ESCAP cited the discrepancy between the A.C.E. data, actual Census data, and Demographic Analysis as a major reason for releasing unadjusted data to the states. What processes would you recommend that the Census Bureau use to examine this discrepancy?

Answer. The most urgent need is for the Bureau to tabulate the results from A.C.E., the Census, and the Demographic Analysis with a consistent set of categories of age, race, and sex. These tabulations will be helpful for deciding on the next stages of analysis.

We recommend that the Bureau publish additional data for independent analysis by the scientific community, including (a) the targets used in their loss function analyses, and (b) the Census populations of Congressional districts by A.C.E. poststrata and the residual population.

We thank the Chair and the Committee for their interest and attention.

RESPONSE TO WRITTEN QUESTIONS SUBMITTED BY HON. JOHN MCCAIN TO DAVID W. MURRAY

Question 1. Some observers argue that it is important to release data adjusted with the A.C.E. results to more accurately account for population growth and enhance our understanding of Census methodology. Given your concerns about errors in the A.C.E. methodology, what would be the effect of releasing A.C.E.-adjusted data?

Answer. The A.C.E. process does indeed have the capacity to "enhance our understanding of Census methodology," and hence, studying the results of the A.C.E. in comparison to Census enumerated data can make an important contribution to our overall understanding of the country's population growth during the last decade. Similarly, studying the data captured in the Demographic Analysis could likewise contribute to our understanding of both the A.C.E. process and the Census enumeration. The comparison of the three measures of the population is important particularly because the respective portraits of the population that they offer can not be reconciled with each other, at this time and given currently available data. That is, there is an important research role for the use of A.C.E. data (which role does not necessarily require their "release" to the respective states).

Moreover, not only are the three data sets internally irreconcilable, they place us in an historically unprecedented situation, in that the Census enumeration shows a larger population than does the Demographic Analysis, a reversal of the expected relationship. That is, from the perspective of the Demographic Analysis, the Census enumeration data show a net overcount of the population, and not an undercount. The A.C.E. data, at least at this preliminary juncture in the analysis, take us even further away from an ability to reconcile the Demographic Analysis with the Census enumeration, in that the A.C.E. data suggest a population count even higher than either of the other two measures. Determining which measurement is giving us the most accurate data is not self-evident, at this stage.

These remarks characterizing the respective population portraits refer, however, only to the overall summary totals of the population, and do not address the even larger challenge of determining distributive accuracy (the proportionally correct allocation of the population accurately positioned in correct geographical areas), which is the essential function of the Census numbers. Given inherent uncertainties of the A.C.E. adjustment process, it seems very likely that the Census enumeration data, the unadjusted data gathered by a direct census, will prove to be the most correct and accurate data for distributive accuracy, a consideration that appeared to prevail in forming the recommendation by the ESCAP committee not to release adjusted data as the official redistricting data file.

In addition, there appear to be misconceptions on the part of the public and some public officials about the quantitative meaning of A.C.E. adjusted numbers. While some people take the A.C.E. data as concrete, specific population counts, in reality all numbers derived from a sampling process are subject to probability fluctuations. That is, rather than deriving a single, stable adjusted count number (for a particular post-stratum), what a sampling adjusted process really gives you is the value drawn from within a spread of values that varies in its range for various post-strata. Importantly, each of those estimated numbers must be accompanied by a Standard Error that expresses the confidence intervals for the value in question (confidence intervals providing the likely upper and lower bounds within which the mean figure is believed to fall a high probability of the time).

For instance, for the Major Group "Native Hawaiian/Pacific Islander," we might hear that the A.C.E. data for Census 2000 show a net undercount of 4.6 percent, while the A.C.E. undercount figure for Asian Americans (non-Hispanic) is pegged at .96 percent undercount (source: Census Bureau B–1 document, table 2a, "Percent Net Undercount for Major Groups: 2000 A.C.E."). Some people improperly regard those numbers as specific and concrete "findings" from the A.C.E. But in reality, the A.C.E. process actually reports those undercount estimates along with a Standard Error (SE) estimate, a plus/minus spread within which the true mean is felt to lie. Hence, for Native Hawaiians/Pacific Islanders, the actual A.C.E. data show an undercount of 4.6 percent with an SE of 2.77 percent. That means that the actual value for that post-stratum's undercount could range from 1.83 percent to 7.37 per-

cent. Likewise, for Asian Americans, the .96 percent undercount has an SE of .64, producing a range of values going from .32 percent to 1.6 percent. Which number represents the "true" undercount adjustment which must be applied to any local population?

Notice further that in these two cases, the SE is larger than half the value of the estimate mean figure. In such circumstances, the probability fluctuation possible is proportionately very large indeed, rendering our portrait of the population of those post-strata relatively indeterminate within the overall range. With a strikingly accurate Census enumeration that materially reduced the undercount, including dramatic reductions in minority post-strata, we rapidly come to the point where adjustment begins to give us diminishing returns; that is, we begin to replace one relatively accurate enumerated figure (which albeit has an estimated undercount percentage) with a second figure, the A.C.E. adjusted figure, that is not better or more accurate, just different. Moreover, its margins of error are in many instances as large or larger than the undercount percentage that it seeks to remedy.

Yet many observers continue to think that the adjusted A.C.E. number is somehow a straightforward percentage correction that can simply be added in to the respective post-strata populations. This is not so. Releasing adjusted data sets to the states could well abet this misapprehension, and even lead to ill-founded legal actions. Finally, the margins of error problem only increases as one approaches smaller and smaller geographical units, which it is the primary responsibility of the Census data to specify with accuracy for apportionment purposes.

Question 2. Some observers have suggested that a "sampling" method, such as the A.C.E., would be helpful in accurately measuring previously undercounted areas, such as parts of Arizona. Could you please discuss why the A.C.E. might lead to an incorrect adjustment to an area that is largely rural and has a large minority population?

Answer. Once again, we must keep foremost the importance of local level accuracy for the Census to accomplish its purposes of accurately apportioning political power and funding to actual people where they actually live. A serious difficulty encountered by the Census in rural areas, particularly those with a low density of a minority population such as found in Arizona, is the problem of inaccurate or incomplete address files for residences. Getting an accurate enumeration is a challenge in such circumstances, and, as the lessons of the Census 2000 amply demonstrate, requires an active mobilization at the local level of officials, agencies, and field personnel who are familiar with the geographical area, who have indigenous command of the respective languages and cultures of the area, who are trusted and accepted by local people, and who have a stake in achieving an accurate count.

It appears to be an unfortunate by-product of the sampling adjustment process that just such effective mobilization is not activated, and may actually be undermined by the process of statistical adjustment based on estimation assumptions for an area that is sometimes derisively referred to as "fly-over country." Rather obviously, "flying over" an area and adjusting it with estimated data derived from nationally-distributed samples cannot provide the kind of local knowledge, and community engagement, that a direct enumeration can accomplish.

The problem of an accurate adjustment is particularly acute when address files are incomplete or inaccurate, since the A.C.E. design depends upon a matching process, comparing the count from the enumeration with the count for the A.C.E. sample for each address covered by the sampling method. Errors in addresses lead to errors in matching, and these together lead to erroneous adjustments.

As our September, 1999 Report to Congress ("Unkept Promise") argued, the adjustment process has a tendency, particularly at local levels of geography, to misallocate the adjustment, adding people erroneously where they were not really missed and, sadly, failing to provide a commensurate correction for areas where striking undercounts actually occurred. Many undercounted communities remain undercounted after adjustment, the missing persons erroneously having been allocated by the adjustment process to other areas, or spread across wide geographical dispersions where they were not undercounted in the first instance. (Realizations such as these may well have motivated the language of the ESCAP committee's recommendation not to use adjusted data where they noted that, under any scenario yet devised, adjusted data were less accurate than enumerated data for counties of less than 100,000 people.)

Somewhat perversely, however, officials acquire the illusion that they have addressed their community's undercount by adjustment, when in fact they have not proportionally done so. Perversely, as well, it appears that the "unkept promise" of adjustment substitutes for the very community based actions that have a far better chance of actually locating the undercounted minorities, and even engaging them in

civic participation by the encounters and community mobilization required for a direct Census enumeration.

The difficulties of the address mis-matching are further amplified by the fact that the post-strata adjustment factor applied to many rural minority communities suffer from heterogeneity problems that the A.C.E. design could never sufficiently resolve. This means that sociological assumptions about a minority post-stratum being homogenous throughout the country (in their probability of being counted or missed in the census) were not born out in practice in actual communities.

As I noted in my earlier written testimony submission, one should not apply the same adjustment factor to a recently-arrived Puerto Rican immigrant to New York City as one applies to an Hispanic ranch owner in Arizona whose family has been in the community for two centuries simply because both parties identify themselves as Hispanic; yet this is exactly what the A.C.E. designs tries to do, using a nationwide post-stratum of Hispanics wherein Arizona state data are adjusted using data actually derived from adjacent or even distant states. As a further example, it would probably surprise many observers to learn that a post-stratum category that experienced one of the highest overcount rates was Native American male homeowners not on reservations. Simply put, the Native American community, ranging from the densely-clustered Hopi villages to the remotely distributed Navajo "outfits" to the remarkably wealthy and modernized Connecticut community of the casino-owing Mashantucket Pequot, has enormous heterogeneity in its actual social and economic make up, a factor that introduces indeterminacy into the A.C.E. assumptions.

It was considerations such as these that were reflected in the ESCAP committee's recommendation not to use adjusted data when they discussed remaining A.C.E. design difficulties termed "synthetic error." A second set of concerns in the ESCAP report, referred to as "balancing error" and addressing the problems of differential search areas in the Census and the A.C.E. for determining matches, represent an additional problem for rural areas of low density and household dispersion, as do the tendencies of some minority groups (such as Native Americans) to live in household groupings that do not coincide with traditional census assumptions about residence patterns, and the noted transience of workers in many rural minority communities.

For these reasons (among many), the A.C.E. process cannot be shown to be more accurate for local communities than the Census enumeration data that they were to replace, and may even encourage a false hope that the undercount problem has been "solved" while simultaneously reducing the incentive to adopt Census enumeration strategies and practices that could prove more effective in ensuring that undercounted communities are identified, accurately enumerated, and accordingly, are provided with the political power and public funding that are their just expectation.

RESPONSE TO WRITTEN QUESTIONS SUBMITTED BY HON. JOHN MCCAIN TO ARTURO VARGAS

Question 1. You have emphasized that over one million Latinos were missed by the Census Bureau. What suggestions would you have to improve the Census Bureau's enumeration process for counting Latinos?

Answer. The first place to begin in determining how to improve the enumeration of Latinos for the next census would be to build upon those elements of Census 2000 which were particularly effective in this regard. From our extensive work in conducting community outreach and public education for the census, we believe the following efforts contributed to the relative success of Census 2000:

- The early waiving of the U.S. citizenship requirement for the hiring of enumerators helped ensure that adequate numbers of bilingual individuals were available to be hired by the Bureau in areas with large non-English speaking populations. However, there continued to be a shortage of bilingual enumerators in areas with small and emerging non-English speaking populations.
- The Census in the Schools helped to convey the importance of the census to families by using schoolchildren as key messengers.
- The paid advertising campaign in English and non-English languages helped raise awareness in the census considerably.
- The partnership program was particularly successful in involving as many community leaders and organizations as possible in being messengers of the importance of the census.

One of the major obstacles to ensuring a full census count of Latinos, however, is a structural element of the enumeration process. As long as the Census Bureau continues to rely primarily on the mail-out/mail-back method, Latinos and other

populations with higher rates of poverty than the general population, non-English speaking populations, and groups of individuals who fear contact with the government, will be particularly vulnerable to being missed in the census. It is our understanding that the census enumeration method may be vastly different in 2010. Any changes to the enumeration process should be specifically designed to remove these inherent barriers to the counting of low-income populations, children, immigrants, and others at historical risk of being undercounted. We understand that a fundamental aspect of "recreating" the census is the introduction of the American Community Survey (ACS), which will be able to provide annual data on the U.S. population and eliminate the need for the long form in the decennial census. We believe that the ACS has particular promise in collecting data that will be critical in designing a more effective decennial census and targeting resources for outreach and enumeration in 2010.

Based on our experience with Census 2000 and information we have received from other Latino organizations and leaders involved in census promotion and outreach, we make the following recommendations:

- Promoting the census must become an ongoing initiative, and not an activity that is left to being a decennial task.
- Congress should promote annual appropriations for the Census Bureau to carry out sustained partnerships between the Bureau and the community, including Latino organizations. The Census Bureau and the Department of Commerce have expressed a commitment to maintaining these partnerships throughout the decade, but it will depend on the willingness of Congress to support these partnerships through appropriations.
- The Congress should support and monitor the American Community Survey, an annual sample survey of the American population that is designed to collect the data now asked in the long form. It is the Census Bureau's plan to eliminate the need for the long form in the decennial census through implementation of the ACS. The ACS can also serve as an ongoing tool to reinforce the confidentiality of the census throughout the decade.
- The American Community Survey must be fully funded by Congress every year.
- Members of Congress should take a leadership role in promoting the census through the public schools by encouraging the school districts in their states and congressional districts to incorporate the census as a permanent element of the school curriculum. Census in the Schools was a particularly successful element of the 2000 Census outreach plan and should happen every year.
- The Congress should support and promote supplemental outreach funding by cities and states in the decennial census. In 2000, jurisdictions such as California and Houston allocated their own resources to census outreach and promotion with considerable success in increasing their mail-back response rates.
- The Congress should advocate for funding for the Census Information Centers. Several Latino organizations have been designated as Census Information Centers, but are unable to raise the funds necessary to analyze, publish and distribute the data they receive. Congress should appropriate resources to make these Centers viable.
- The Congress should carefully evaluate all aspects of the Census 2000 and work with the Bureau and the community to improve those areas that fell short. Some key concerns in Census 2000 include the following issues:

    - As in 1970, 1980 and 1990, it again was difficult for the public to receive the Spanish-language form in a timely and easy manner, which contributed to confusion and frustration during the enumeration. We strongly recommend that the Census Bureau determine how to make the non-English language census forms more easily accessible to the public. For Census 2010, data from the ACS on levels of English-language proficiency and non-English language use would be especially valuable in identifying areas where non-English language forms could be made more readily available, including by mailing the non-English language forms in targeted communities with particular non-English languages predominate.
    - While it appears that adequate numbers of bilingual enumerators were employed in areas with large Latino populations, such as Los Angeles and south Texas, there was a severe shortage of bilingual census workers in areas with emerging Latino communities, such as western Pennsylvania and western New York. For Census 2010, data from the ACS on levels of English-language proficiency, non-English language use, and nativity would be especially valuable in identifying all the areas where bilingual enumerators will be required.
    - While the paid advertising campaign appeared to be successful overall, there were limitations on the media outlets that were selected to broadcast the

spots. Future advertising should be distributed to those media outlets most watched/listened/read by Latinos.

• The Congress should encourage and prod the Census Bureau to hire more Latinos in key policy making position at the national office where Latinos presently are significantly under-represented. Latinos also are under-represented in senior management positions in the regional offices as well. Having Latinos in key positions throughout the Census Bureau is key to sustaining a successful decade-long outreach effort in the Latino community.

Question 2. You also highlight the importance of an accurate Census count to the Latino community. The Census Bureau says that there is a large discrepancy between their Demographic Analysis, official Census data, and A.C.E. results. What factors should Secretary Evans consider as he reviews criteria for releasing adjusted data?

Answer. We are particularly concerned about the discrepancy suggested by the Demographic Analysis; we believe the Census Bureau should fully research possible errors in this analysis. First, analyses of the 2000 Census data suggest that thousands of persons may have been counted in 2000 who were in the country on April 1, 1990 and missed by the 1990 Census. Second, it also has been suggested by the analyses from Census 2000 and the A.C.E. that the assumptions about the levels of immigration during the 1990s used the Demographic Analysis were significantly miscalculated. These possible errors would have a particular impact on the Latino population and we strongly recommend that the Secretary consider these possibility of greater error in the 1990 Census than what has been assumed, as well as misassumptions about immigration levels during the 1990s, which may have flawed the Demographic Analysis.

In addition, I would like to bring to the attention of the Senate Committee the recommendations on this point forwarded to Secretary Evans by the Secretary of Commerce's Advisory Committee on the Decennial Census. These recommendations were forwarded by the Acting Chair of the Committee to Secretary Evans in a letter dated April 6, 2001. The recommendations read as follows: (1) The Committee recommends that the Department of Commerce and the Bureau of the Census complete the work of evaluation of the Census 2000 and Accuracy and Coverage Evaluation (A.C.E.) data, report the data as completely as possible including by race, ethnicity, and geography, and complete the report expeditiously. (2) If the Bureau of the Census recommends adjusting the data, the Committee recommends that the Secretary publish the adjusted data at all levels of geography and incorporate the adjusted data into all data products. (3) If the Bureau of the Census does not find the adjusted data to be more accurate than Census 2000 data, then the Committee recommends that the adjusted data should be available for research purposes. (4) If the Bureau of the Census makes the recommendation to adjust the data, and the Department of Commerce decides to not release the adjusted data for official purposes, then the decision and the adjusted data should be made public.

RESPONSE TO WRITTEN QUESTIONS SUBMITTED BY HON. JOHN MCCAIN TO EUGENE P. ERICKSEN

Question 1. The Census Bureau's ESCAP Panel recommended against releasing adjusted Census numbers, because of the large discrepancy between the Demographic Analysis, the Accuracy and Coverage Evaluation (A.C.E.) sample, and the actual Census data. What may have caused this discrepancy?

Answer. The national estimates of population provided by these three methods are: Demographic analysis: 279.6 million; Census count: 281.4 million; A.C.E. estimate: 284.7 million.

The demographic analysis estimate is probably too low. There are indications that the volume of international immigration, especially of the undocumented type, was greater than the Census Bureau had estimated. I believe that the valid construction of such estimates has become more difficult than it was in 1980 or 1990, and we must consider the possibility that valid demographic estimates may not be possible for 2000.

The census count is probably too low, but perhaps not by very much. As indicated in the ESCAP report (see pp. 4–6), both demographic analysis and the A.C.E. indicate a Black undercount rate between 2 and 3 percent. Estimates for the non-Black population provided by the two methods differ, primarily because they differ greatly for Hispanics. Indications from the A.C.E. suggest that net undercounts for Hispanics and similar to those for Blacks; undercount rates for non-Hispanic Whites and Asians are lower, but for Native Americans they are higher.

The A.C.E. estimate may be too high. My reading of the ESCAP report, and various backup reports, suggests that the possibility that the Census Bureau underestimated the rate of erroneous enumeration. The bureau features two factors leading to this possibility. One is "balancing error," resulting from an inconsistency in the geographic search areas used in the P- and E-samples in the A.C.E. The other is due to high rates of duplication in the "late census adds."

When deciding whether individual people were counted correctly in the A.C.E., the Census Bureau searches the block(s) surrounding each sample member's address. For example, if a person were missed at her own address, but counted next door, the bureau would not count such a person as an omission. Problems in locating surrounding addresses in the correct blocks could have created an imbalance in the searches. The Census Bureau is now investigating this possibility.

The Census Bureau used a sophisticated computer program, and other methods to estimate that 6.5 million people had been counted twice at the same address. The bureau removed them from the count. Upon later reflection, the bureau decided that about 2.2 million of them may not have been counted twice, and they returned them to the count as "late census adds." It is likely that a substantial number of these people were in fact duplicates although evidence for this is inconclusive.

I have been told that the bureau may have underestimated erroneous enumerations by as many as 2 million people (see ESCAP report, p. 25). Should this be the case, the net undercount would be reduced to 1.3 million people. It is likely that racial differential undercounts would remain, and that some groups would have net overcounts.

Question 2. In your testimony, you argue that comparison of adjusted 1990 and 2000 Census data is required to accurately measure population growth. Why cannot sets of 1990 and 2000 unadjusted Census data be used to measure population growth.

Answer. My comments referred to estimates of population growth, for local areas, where the rates of change are more variable than they are for the nation as a whole. As I testified, the reductions in the net undercount between 1990 and 2000 are much greater for minority than for the non-Hispanic White population.

Let us consider two cities of 100,000 true population. One, City A, is 90 percent White and 10 percent Black and the other, City B, is 30 percent White and 70 percent Black. Let us further assume that the local undercounts in each place mirror the national undercounts for each of the 1990 and 2000 Censuses. This is to say that the White undercount is 0.7 percent in both years, but that the Black undercount was 4.6 percent in 1990 and 2.2 percent in 2000.

The undercount in City A was 1.09 percent in 1990 and 0.85 percent in 2000, a small improvement of 0.24 percent. For City B the comparable numbers are 3.43 percent in 1990 and 1.75 percent in 2000, a larger improvement of 1.68 percent. The City A population count would appear to have grown from 98,100 in 1990 to 99,050 in 2000. The City B count would appear to have grown even more from 96,570 to 98,250.

For cities where the shares of rental housing are greater, the year-to-year discrepancies would be greater than this. As a consequence, a city's or county's population could appear to grow even though its population had not changed, due to the fact that the Census Bureau counted better in 2000.

Many of the population counts obtained in 2000 have surprised local officials, and they are inconsistent with the Census Bureau's local population estimates for 1998 and 1999.

I have taken these estimates and projected a 2000 estimate by projecting the 1998–99 change. For example if the Bureau's estimate was 660,000 in 1998 and 650,000 in 1999, the expected 2000 estimate was 640,000. I then compared the expectation to the actual count. I did this for all counties with at least 500,000 people in the 1990 Census.

As shown in the attached table, most of the expected values were lower than the counts. This was especially true in those counties with large minority populations. I obtained the following results:

Average Percentage:

Percent Black Or Hispanic, 2000: 40 percent or more/Difference, Estimate And Count: 3.84 percent.

Percent Black Or Hispanic, 2000: 20 to 39.9 percent/Difference, Estimate And Count: 3.09 percent.

Percent Black Or Hispanic, 2000: 0 to 19.9 percent/Difference, Estimate And Count: 1.96 percent.

The difference is defined as (count–estimate)/count. A positive number means that the count was greater than the estimate. Of the 97 counties in my studies, 86 had

positive discrepancies. The above results show that the discrepancies were on average large in counties with greater shares of Black or Hispanic populations.

Errors in Population Estimates for 2000 by Counties and Percent Minority

State	County	Percent Hispanic or NH Black	1990 Population	2000 Population	1998 Estimate	1999 Estimate	2000 Estimate	Error
Texas	El Paso	80.98%	591,610	679,622	694,603	701,908	709,213	-4.35%
New York	Bronx	79.62	1,203,789	1,332,650	1,191,319	1,194,099	1,196,879	10.19
Florida	Dade	76.28	1,937,094	2,253,362	2,150,877	2,175,634	2,200,391	2.35
Maryland	Prince Georges	69.29	729,268	801,515	776,907	781,781	786,655	1.85
District of Columbia	67.31	606,900	572,059	521,426	519,000	516,574	9.70	
Maryland	Baltimore City	65.74	736,014	651,154	645,664	632,681	619,698	4.83
Georgia	De Kalb	61.71	545,837	665,865	592,870	596,853	600,836	9.77
Texas	Bexar	61.24	1,185,394	1,392,931	1,354,837	1,372,867	1,390,897	0.15
New Jersey	Essex	55.72	778,206	793,633	748,322	747,355	746,388	5.95
New York	Kings	54.21	2,300,664	2,465,326	2,266,242	2,268,297	2,270,352	7.91
California	Los Angeles	54.03	8,863,164	9,519,338	9,223,807	9,329,989	9,436,171	0.87
New Jersey	Hudson	51.92	553,099	608,975	553,030	552,819	552,608	9.26
Texas	Harris	51.15	2,818,199	3,400,578	3,202,021	3,250,404	3,298,787	2.99
Pennsylvania	Philadelphia	51.07	1,585,577	1,517,550	1,434,968	1,417,601	1,400,234	7.73
Tennessee	Shelby	50.98	669,018	897,472	867,804	873,000	878,196	2.15
Georgia	Fulton	50.13	648,951	816,006	737,222	744,827	752,432	7.79
Texas	Dallas	49.95	1,852,810	2,218,899	2,045,309	2,062,100	2,078,891	6.31
California	Fresno	49.03	667,490	799,407	755,051	763,069	771,087	3.54
California	San Bernardino	47.94	1,418,380	1,709,434	1,635,967	1,669,934	1,703,901	0.32
Illinois	Cook	45.79	5,105,067	5,376,741	5,192,396	5,192,326	5,192,256	3.43
Michigan	Wayne	45.69	2,111,687	2,061,162	2,116,540	2,106,495	2,096,450	-1.71
California	Kern	44.11	543,477	661,645	631,615	642,495	653,375	1.25
New York	Queens	43.93	1,951,598	2,229,379	1,993,172	2,000,642	2,008,112	9.93
New York	New York	42.45	1,487,536	1,537,195	1,546,508	1,551,844	1,557,180	-1.30
California	Riverside	42.19	1,170,413	1,545,387	1,480,708	1,530,653	1,580,598	-2.28
Alabama	Jefferson	40.77	651,525	662,047	660,039	657,422	654,805	1.09
							Average	3.84%
Texas	Travis	37.21	576,407	812,280	709,182	727,022	744,862	8.30
Florida	Broward	36.78	1,255,488	1,623,018	1,507,770	1,535,468	1,563,166	3.69
Massachusetts	Suffolk	36.36	511,433	689,807	641,333	641,695	642,057	6.92
Florida	Orange	36.31	677,491	896,344	804,489	817,206	829,923	7.41
California	Ventura	35.21	669,018	753,197	732,143	745,063	757,983	-0.64
North Carolina	Mecklenburg	34.12	649,623	695,454	630,813	648,400	665,987	4.24
California	Alameda	33.60	1,279,182	1,443,741	1,397,050	1,415,582	1,434,114	0.67
Wisconsin	Milwaukee	33.07	959,275	940,164	911,536	906,248	900,960	4.17

Florida	Hillsborough	32.43	834,054	998,948	925,413	940,484	955,555	4.34
Texas	Tarrant	32.36	1,170,103	1,446,219	1,354,040	1,382,442	1,410,844	2.45
California	Orange	32.26	2,410,558	2,846,289	2,723,782	2,760,948	2,798,114	1.69
Arizona	Pima	32.19	666,880	843,746	790,333	803,618	816,903	3.18
California	San Diego	32.18	2,498,016	2,813,833	2,766,123	2,820,844	2,875,565	-2.19
Florida	Duval	31.64	672,971	778,879	734,664	738,483	742,302	4.70
Nevada	Clark	30.79	741,459	1,375,765	1,161,259	1,217,155	1,273,051	7.47
Ohio	Cuyahoga	30.59	1,412,140	1,393,978	1,380,428	1,371,717	1,363,006	2.22
Tennessee	Davidson	30.36	1,497,577	569,891	533,258	530,050	526,842	7.55
New York	Westchester	29.17	874,866	923,459	900,861	905,572	910,283	1.43
Missouri	Jackson	28.48	633,232	654,880	655,055	654,484	653,913	0.15
Arizona	Maricopa	28.38	2,122,101	3,072,149	2,783,779	2,861,395	2,939,011	4.33
Indiana	Marion	27.89	797,159	860,454	812,662	810,946	809,230	5.95
New Jersey	Camden	26.96	502,824	508,932	504,268	503,093	501,918	1.38
California	Contra Costa	26.84	803,732	948,816	917,970	933,141	948,312	0.05
California	Santa Clara	26.62	1,497,577	1,682,585	1,641,848	1,647,419	1,652,990	1.76
Maryland	Montgomery	26.33	757,027	873,341	839,158	852,174	865,190	0.93
Florida	Palm Beach	25.91	863,518	1,131,184	1,032,872	1,049,420	1,065,968	5.77
California	Sacramento	25.66	1,041,219	1,223,499	1,166,699	1,184,586	1,202,473	1.72
California	San Mateo	25.24	649,623	707,161	701,080	702,102	703,124	0.57
Ohio	Hamilton	24.45	866,228	845,303	847,202	840,443	833,684	1.37
Oklahoma	Oklahoma	23.55	599,611	660,448	632,865	636,539	640,213	3.06
Connecticut	Hartford	22.65	851,783	857,183	827,706	829,671	831,636	2.98
New Jersey	Middlesex	22.18	671,780	750,162	712,638	717,949	723,260	3.59
Maryland	Baltimore County	21.77	692,134	754,292	721,556	723,914	726,272	3.71
California	San Francisco	21.67	723,959	776,733	745,756	746,777	747,798	3.73
Connecticut	Fairfield	21.48	827,645	882,567	837,476	841,334	845,192	4.23
Illinois	Lake	21.15	516,418	644,356	608,348	617,975	627,602	2.60
Ohio	Montgomery	21.03	573,809	559,062	570,141	565,866	561,591	-0.45
Connecticut	New Haven	20.92	804,219	824,008	792,879	793,208	793,537	3.70
Kentucky	Jefferson	20.53	664,937	693,604	671,595	672,900	674,205	2.80
Missouri	Saint Louis	20.38	993,529	1,016,315	997,347	996,181	995,015	2.10
Ohio	Franklin	20.01	961,437	1,068,978	1,021,578	1,027,821	1,034,064	3.27
							Average	3.09%
New York	Nassau	19.72	1,287,348	1,334,544	1,300,995	1,305,057	1,309,119	1.91
Virginia	Fairfax	19.41	818,584	969,749	927,895	945,717	963,539	0.64
Rhode Island	Providence	19.19	596,270	621,602	573,701	574,108	574,515	7.58
New York	Monroe	18.66	713,968	735,343	714,936	712,419	709,902	3.46
New York	Suffolk	17.10	1,321,864	1,419,369	1,370,549	1,383,847	1,397,145	1.57
Oklahoma	Tulsa	16.81	503,341	563,299	543,417	548,296	553,175	1.80

Errors in Population Estimates for 2000 by Counties and Percent Minority—Continued

State	County	Percent Hispanic or NH Black	1990 Population	2000 Population	1998 Estimate	1999 Estimate	2000 Estimate	Error
New York	Erie	16.03	968,532	950,265	933,702	925,957	918,212	3.37
Pennsylvania	Delaware	15.87	547,651	550,864	542,592	541,502	540,412	1.90
Michigan	Kent	15.70	500,631	574,335	544,781	550,388	555,995	3.19
New Jersey	Bergen	15.33	825,380	884,118	854,428	857,052	859,676	2.76
Ohio	Summit	14.00	514,990	542,899	537,160	537,856	538,552	0.80
New Jersey	Monmouth	13.96	553,124	615,301	603,214	611,444	619,674	-0.71
Florida	Pinellas	13.45	851,659	921,482	877,273	878,499	879,725	4.53
Pennsylvania	Allegheny	13.20	1,336,449	1,281,666	1,267,963	1,256,806	1,245,649	2.81
Oregon	Multnomah	13.05	583,887	660,486	630,573	633,224	635,875	3.73
Massachusetts	Essex	13.03	670,080	723,419	700,370	704,407	708,444	2.07
Minnesota	Hennepin	12.91	1,032,431	1,116,200	1,058,943	1,064,419	1,069,895	4.15
Utah	Salt Lake	12.83	725,956	898,387	845,913	850,243	854,573	4.88
Michigan	Oakland	12.45	1,083,592	1,194,156	1,175,057	1,179,978	1,184,899	0.78
Washington	Pierce	12.27	586,203	700,820	675,962	688,807	701,652	-0.12
Illinois	DuPage	11.98	781,666	904,161	880,996	892,547	904,098	0.01
Washington	King	10.77	1,507,319	1,737,034	1,654,329	1,664,846	1,675,363	3.55
Pennsylvania	Montgomery	9.41	678,111	750,097	719,569	724,087	728,605	2.87
Massachusetts	Worcester	9.20	709,705	750,963	730,769	738,629	746,489	0.60
Hawaii	Honolulu	8.94	871,768	876,156	871,768	864,571	857,374	2.14
Massachusetts	Middlesex	7.73	1,398,468	1,465,396	1,422,465	1,426,606	1,430,747	2.36
Pennsylvania	Bucks	5.52	541,174	597,635	587,863	594,047	600,231	-0.43
Massachusetts	Bristol	5.43	506,325	534,678	516,975	520,258	523,541	2.08
Massachusetts	Norfolk	4.94	616,087	650,308	642,089	643,580	645,071	0.81
Michigan	Macomb	4.26	717,400	788,149	786,866	792,082	797,298	-1.16
.....	.....	.....	.....	.....	.....	.....	Average	1.96%

Source: ***www.census.gov/People estimates/County Population Estimates/ County Population Estimates for July 1, 1999 and the Population Change from July 1, 1998 to July 1, 1999/American FactFinder/Race Hispanic or Latino (2000)***

RESPONSE TO WRITTEN QUESTIONS SUBMITTED BY HON. BARBARA BOXER TO DONALD L. EVANS

Question 1. Mr. Secretary, I am requesting two things. First, for the Census Bureau to immediately release all the data collected and the adjusted numbers it believes are inaccurate so that independent experts can examine the numbers and assumptions.

Answer. In keeping with our goal of openness and transparency in decisions and processes, the Census Bureau signed a Memorandum of Understanding (MOU) with the U.S. House of Representatives Committee on Government Reform Subcommittee on the Census, the Census Monitoring Board, and the National Academy of Sciences' Committee on National Statistics, granting access to data to facilitate independent review of Census 2000. Presently, a second MOU is under consideration by these oversight groups. This MOU provides the opportunity for outside researchers to study the A.C.E. Census 2000 and demographic analysis data while appropriately controlling access to sensitive data and protecting the ongoing deliberative process of Census Bureau staff and other officials of the Department of Commerce. The data now being made available under this MOU have been used by the Census Bureau for assessing the accuracy of the Accuracy and Coverage Evaluation (A.C.E) survey results. We believe that providing these materials will further the oversight responsibilities of the Congress and the other above-mentioned entities as well as satisfy the requests by many Members of Congress and other interested parties that the A.C.E. data be made available. Although the Senate had not previously participated in this review process, we have contacted the committees of jurisdiction to extend the same opportunity.

Question 2. Second, for the Census Bureau to move with immediate haste to resolve the issues it believes have distorted the adjusted numbers and provide Congress with adjusted numbers well before the Fall deadline for determining distribution of federal funds. Can you assure myself and the Committee that these requests will be met?

Answer. The Census Bureau is currently engaged in a comprehensive research and evaluation effort to address all outstanding issues and to resolve questions raised by the Executive Steering Committee for Accuracy and Coverage Evaluation Policy (ESCAP) report issued on March 1, 2001. Our commitment to provide the Congress and the public with the most accurate and statistically sound data remains our primary objective. As soon as possible, the Census Bureau plans to release its findings and make a recommendation to the Secretary of Commerce concerning the accuracy of the adjusted data. We will provide these findings to the Congress as swiftly as possible.

[From USA TODAY, March 28, 2001]

REPORT: CITIES, MINORITIES LOSE IN CENSUS UNDERCOUNT

(By Haya El Nasser and Paul Overberg)

California could claim half a million more people, Texas another 350,000 and New York 300,000 if the 2000 Census were adjusted to make up for the people who were missed, according to a Democratic-backed, state-by-state undercount estimate out today.

The undercount numbers for 50 states, the District of Columbia and five big cities are the first to indicate where an estimated 3.3 million missed by the Census were living.

The numbers, released by Democrats on the Census Monitoring Board, are likely to put more pressure on the Census Bureau to release adjusted population counts for all areas.

The Census Bureau is still reviewing the numbers.

The estimates were done by Eugene Ericksen, a statistics professor at Temple University and a consultant to Democrats on the monitoring board. He used the methods outlined by the Census Bureau in its own study of whether Census numbers should be adjusted.

His analysis shows the largest percentages of people missed were in states with a high number of minorities and big cities. The net undercounts were smallest in some Midwestern and Plains states.

Undercounts have broad implications for states and cities because Census numbers are used to redraw political districts and distribute $185 billion in Federal funds annually.

Most Republicans are opposed to the estimates because they say it creates "virtual" people. Democrats say it's the only way to correct the disproportionate undercount of minorities and the poor.

"At this point, we don't have the confidence in (our adjusted) numbers to release them," says John Thompson, the Census Bureau's associate director for the decennial Census.

Acting Census Director William Barron said he could not comment on Ericksen's estimates because he had not seen his methodology. Commerce Secretary Donald Evans, who oversees the Census Bureau, is to testify about the numbers today before a Senate panel.

The estimates claim about 188,500 people were missed in New York City (2.3 percent undercount); more than 62,000 in Chicago (2.1 percent); 44,000 in Houston (2.2 percent); 21,500 in Philadelphia (1.4 percent); and 9,400 in Atlanta (2.2 percent).

Census 2000 was the most accurate count ever. Undercount estimates released today show that 35 states—including the two biggest, California and Texas—had a smaller undercount in 2000 than in 1990.

In 2000, officials said the Census missed at least 6.4 million and counted at least 3.1 million twice, a net undercount of 3.3 million. In 1990, it missed 8.4 million and counted 4.4 million twice, a net undercount of 4 million.

The Census Bureau recommended against using adjusted numbers because it couldn't explain huge gaps between the Census, adjusted numbers and estimates based on birth, death and immigration records.

Undercount examined

Democratic members of the Census Monitoring Board, a bipartisan oversight group, are releasing today the first state estimates of people missed in Census 2000. Estimates were calculated by Eugene Ericksen, a statistician at Temple University in Philadelphia. He used the same method the Census Bureau adopted in its study of the undercount but has not released. He used published and unpublished Census data, but he did not have access to all the data the bureau is using in its study.

State	Net Undercount	Number Missed	Official total	Adjusted total
D.C.	2.4%	14,067	572,059	586,126
Nevada	2.0%	40,157	1,998,257	2,038,414
Alaska	1.7%	11,102	626,932	638,034
Idaho	1.7%	22,244	1,293,953	1,316,197
Texas	1.7%	356,296	20,851,820	21,208,116
Wyoming	1.7%	8,386	493,782	502,168
Hawaii	1.6%	19,325	1,211,537	1,230,862
Montana	1.6%	14,390	902,195	916,585
New Mexico	1.6%	29,766	1,819,046	1,848,812
California	1.5%	529,782	33,871,648	34,401,430
New York	1.5%	290,938	18,976,457	19,267,395
Vermont	1.5%	9,146	608,827	617,973
Maryland	1.4%	75,204	5,296,486	5,371,690
Arkansas	1.3%	35,212	2,673,400	2,708,612
Delaware	1.3%	10,241	783,600	793,841
Florida	1.3%	208,867	15,982,378	16,191,245
Maine	1.3%	16,923	1,274,923	1,291,846
Oregon	1.3%	46,118	3,421,399	3,467,517
Tennessee	1.3%	74,351	5,689,283	5,763,634
Virginia	1.3%	89,601	7,078,515	7,168,116
Arizona	1.2%	60,739	5,130,632	5,191,371
Colorado	1.2%	53,564	4,301,261	4,354,825
Georgia	1.2%	101,947	8,186,453	8,288,400
Louisiana	1.2%	54,737	4,468,976	4,523,713
North Carolina	1.2%	95,292	8,049,313	8,144,605
Oklahoma	1.2%	41,557	3,450,654	3,492,211
Alabama	1.1%	49,462	4,447,100	4,496,562
Kentucky	1.1%	45,781	4,041,769	4,087,550
Mississippi	1.1%	31,348	2,844,658	2,876,006
New Hampshire	1.1%	13,240	1,235,786	1,249,026
New Jersey	1.1%	91,867	8,414,350	8,506,217
Rhode Island	1.1%	11,874	1,048,319	1,060,193
South Carolina	1.1%	43,393	4,012,012	4,055,405
Utah	1.1%	25,523	2,233,169	2,258,692
Washington	1.1%	67,968	5,894,121	5,962,089
West Virginia	1.1%	20,853	1,808,344	1,829,197
Illinois	1.0%	119,115	12,419,293	12,538,408
Connecticut	0.9%	29,195	3,405,565	3,434,760
Pennsylvania	0.9%	112,784	12,281,054	12,393,838
Massachusetts	0.8%	48,623	6,349,097	6,397,720
Indiana	0.7%	45,331	6,080,485	6,125,816
Nebraska	0.7%	12,758	1,711,263	1,724,021
Wisconsin	0.7%	37,810	5,363,675	5,401,485
Iowa	0.6%	16,480	2,926,324	2,942,804
Kansas	0.6%	15,412	2,688,418	2,703,830
Missouri	0.6%	30,944	5,595,211	5,626,155
Ohio	0.6%	68,530	11,353,140	11,421,670
Michigan	0.5%	51,950	9,938,444	9,990,394
South Dakota	0.5%	3,793	754,844	758,637
North Dakota	0.4%	2,385	642,200	644,585
Minnesota	0.3%	16,288	4,919,479	4,935,767

Source: Eugene Ericksen, Democratic appointees of Census Monitoring Board