

---

---

# A Debate on Adjusting the Census for Undercount

The Census Bureau's Response to Papers Submitted to the House  
Subcommittee on the Census on May 5, 1998, with a Rejoinder

---

---

August 28, 1998

# Contents

---

Letter from the Honorable Carolyn B. Maloney,  
ranking minority member of the Subcommittee on the Census,  
to James F. Holmes, Acting Director of Bureau of the Census

Memorandum for the Record from Preston J. Waite, Assistant to  
the Associate Director for Decennial Census

Rejoinder to the Census Bureau's Memorandum for the Record

---

DAN BURTON, INDIANA  
CHAIRMAN

BENJAMIN A. GILMAN, NEW YORK  
J. DENNIS HASTERT, ILLINOIS  
CONSTANCE A. MORELLA, MARYLAND  
CHRISTOPHER SHAYS, CONNECTICUT  
STEVEN SCHIFF, NEW MEXICO  
CHRISTOPHER COX, CALIFORNIA  
ILEANA ROS-LEHTINEN, FLORIDA  
JOHN M. McHUGH, NEW YORK  
STEPHEN HORN, CALIFORNIA  
JOHN L. MICA, FLORIDA  
THOMAS M. DAVIS III, VIRGINIA  
DAVID M. McINTOSH, INDIANA  
MARK E. SOUDER, INDIANA  
JOE SCARBOROUGH, FLORIDA  
JOHN SHADEGG, ARIZONA  
STEVE C. LATOURETTE, OHIO  
MARSHALL "MARK" SANFORD, SOUTH CAROLINA  
JOHN E. SUNUNU, NEW HAMPSHIRE  
PETE SESSIONS, TEXAS  
MIKE PAPPAS, NEW JERSEY  
VINCE SNOWBARGER, KANSAS  
BOB BARR, GEORGIA  
ROB PORTMAN, OHIO

ONE HUNDRED FIFTH CONGRESS

# Congress of the United States

## House of Representatives

COMMITTEE ON GOVERNMENT REFORM AND OVERSIGHT

2157 RAYBURN HOUSE OFFICE BUILDING

WASHINGTON, DC 20515-6143

MAJORITY (202) 225-5074  
MINORITY (202) 225-5051  
TTY (202) 225-6852

HENRY A. WAXMAN, CALIFORNIA  
RANKING MINORITY MEMBER

TOM LANTOS, CALIFORNIA  
BOB WISE, WEST VIRGINIA  
MAJOR R. OWENS, NEW YORK  
EDOLPHUS TOWNS, NEW YORK  
PAUL E. KANJORSKI, PENNSYLVANIA  
GARY A. CONDIT, CALIFORNIA  
CAROLYN B. MALONEY, NEW YORK  
THOMAS M. BARRETT, WISCONSIN  
ELEANOR HOLMES NORTON,  
DISTRICT OF COLUMBIA  
CHAKA FATTAH, PENNSYLVANIA  
ELIJAH E. CUMMINGS, MARYLAND  
DENNIS KUCINICH, OHIO  
ROD R. BLAGOSJEVICH, ILLINOIS  
DANNY K. DAVIS, ILLINOIS  
JOHN F. TIERNEY, MASSACHUSETTS  
JIM TURNER, TEXAS  
THOMAS H. ALLEN, MAINE  
HAROLD E. FORD, JR., TENNESSEE

BERNARD SANDERS, VERMONT  
INDEPENDENT

May 14, 1998

Mr. James Holmes  
Acting Director  
Bureau of the Census  
Department of Commerce  
Washington, DC 20233

Dear Mr. Holmes:

At a hearing of the Government Reform and Oversight Subcommittee on the Census on May 5, 1998, Mr. Kenneth Darga from the Office of the State Demographer, Michigan Department of Management and Budget, submitted two papers for the record. These papers are highly critical of the methods and results of the Post Enumeration Study (PES) conducted following the 1990 Census. I am writing to request that the Census Department officially respond to Mr. Darga's work so that too can be included in the hearing record.

Mr. Darga submitted two papers, *Straining at Gnats and Swallowing Camels: The Perils of Adjusting for Census Undercount* and *Quantifying Measurement Error and Bias in the 1990 Undercount Estimates*. In the first paper, the author contends that "although the results of [the PES] appeared plausible, at least at the broadest national aggregation, the method cannot produce reliable adjustments for undercount: It is not capable of counting many of the people who are missed by the Census, it is very sensitive even to extremely small sources of error, and it is subject to many sources of error that are very serious." The differences between the PES and the original Census data, in the author's opinion, may represent the difficulties in matching records between two surveys, rather than a true net undercount.

The author further contends that it is thus "not surprising to find that many of the detailed undercount measurements for 1990 were implausible and, in some cases demonstrably false. In an effort to correct a net national undercount of less than 2%, spurious undercounts of 10%, 20% and even 30% were identified for some segments of the population." Here the author seems to be referring to certain selected undercount adjustments for children under age 10 from the 1990 PES.

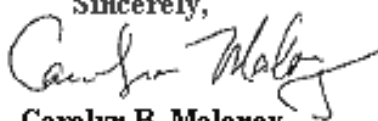
Mr. Darga concludes that using the adjustments of the PES in 1990 "would have had a devastating impact on the usefulness and accuracy of Census data at the state and local level," and that "similar problems can be expected...for Census 2000: The problems are not due to

minor flaws in methodology or implementation, but rather to the impossibility of measuring undercount through the sort of coverage survey that has been proposed.”

Mr. Darga’s second paper purports to identify and quantify several specific types of error, including survey matching error, fabrication of interviews, ambiguity or misreporting of usual residence, geocoding errors, unreliable interviews, and unresolvable cases. His analysis draws heavily on 22 unpublished reports, issued in July 1991 under the title “1990 Post-Enumeration Survey Evaluation Project” by the Census Bureau, and upon the work of Dr. Leo Breiman, an emeritus professor of statistics at the University of California, Berkeley. The quantified results of the errors identified by the author lead him to conclude that “about 70% of the net undercount adjustment that had been proposed for the 1990 Census count – 3,706,000 out of 5,275,000 persons – actually reflects identified measurement errors rather than actual undercount.”

The author infers from these data that the 1990 PES “missed a very substantial number of people who were missed by the Census, but that it also identified a large number of people as missed by the Census who actually had been counted.” His overall conclusion is that while the results of the PES and demographic analysis are very similar, the PES “cannot be relied upon to shed light on patterns of undercount for different demographic components of the population or for different geographic areas.”

As I am sure you agree, Mr. Darga’s conclusions about the reliability of the 1990 PES should not go unanswered for our hearing record, even if they represent only a small minority of scientific opinion on the issue. Please address all of the author’s criticisms, including ones I may not have mentioned. Please provide your reply by May 29, 1998. Thank you for your attention to this matter.

Sincerely,  
  
**Carolyn B. Maloney**  
**Member of Congress**

CBM/ms

cc: Rep. Dan Miller, Chairman, Subcommittee on the Census

enclosure

---

---

## Memorandum for the Record

The Census Bureau's Response to Papers Submitted to the House  
Subcommittee on the Census on May 5, 1998

---

---

Preston J. Waite, Assistant to the Associate Director  
For Decennial Census  
U.S. Bureau of the Census

**Note: Underlines have been added to indicate text which is quoted in the subsequent rejoinder. Annotations in {brackets} indicate sections of the rejoinder in which the preceding text is quoted.**

## MEMORANDUM FOR THE RECORD

From: Preston Jay Waite  
Assistant to the Associate Director for Decennial Census

Subject: Response to Darga Papers Submitted at 5/5/98 Oversight Hearing

This memorandum addresses concerns raised by Mr. Kenneth Darga in the two papers, *Straining Out Gnats and Swallowing Camels: The Perils of Adjusting for Census Undercount and Quantifying Measurement Error and Bias in the 1990 Undercount Estimates*, submitted for the record at the May 5, 1998 hearing of the Government Reform and Oversight Subcommittee on the Census. Both papers are critical of the methods and results of the 1990 Post Enumeration Survey (PES).

In the papers, Mr. Darga asserts that “although the results of the 1990 PES appeared plausible, at least at the broadest national aggregation,” it is not an acceptable method for census adjustment because it: (1) is incapable of counting people missed in the census and (2) is subject to serious errors. Further, Mr. Darga contends that the undercount rates produced from the survey are spurious for certain segments of the population and, therefore, would decrease the accuracy of local population counts. It is his belief that observed differences between the PES and the census data represent difficulties in matching census records to PES records, rather than an actual undercount in the census. In the end, Mr. Darga states that it is impossible to measure undercount “through the sort of coverage survey that has been proposed.”

The issues raised by Mr. Darga are not new to the undercount/adjustment debate. Powerful arguments about coverage measurement have been made in support of adjustment of the decennial census as well as against adjustment. {B-3} There is a growing body of literature documenting both positions in this controversy. Mr. Darga has chosen the strategy of challenging the quality of the coverage survey that provides the data used to produce the undercount rates. In the paragraphs that follow, we will attempt to address some of the statements made by Mr. Darga. The Census Bureau has a strong commitment to producing high-quality, accurate, impartial census numbers, and this commitment extends to the coverage survey data and the coverage improvement methodology.

**1. The PES is not capable of counting many of the people who are missed by the census.** The claimed inability of the PES to count people who are missed by the census appears to be based on the premise that most people who are omitted from the census are “homeless” or “people who do not want to be counted.” {A-2} The argument focuses on “drug dealers, fugitives, and illegal immigrants [who] were afraid to fill out the census form that everyone [sic] in the nation received.” This seems to imply that the American people can be divided neatly into two groups: those who are nearly impossible to count and those who are trivially easy to count. In fact, the census-taking situation is more complex. Mr. Darga’s discussion focuses heavily on correlation bias but fails to mention the inherent differences between the census enumeration process and the survey methodology. People may be missed for many reasons: (1) if their housing unit is not included on the Census Bureau’s address list; (2) if their housing unit is listed, but the post office

delivers the questionnaire to the wrong address, or the census taker goes to the wrong address; (3) if they move close to Census Day; (4) if they misunderstand the questionnaire; or (5) if the census taker fails to ask the questions correctly. The list of examples is endless. {A-2}

There are indeed people who are actively hiding from the government and who are nearly impossible to count. However, it would not be true to say that all illegal immigrants are hiding from the government and are missed by the census. The same argument pertains to “drug dealers” and “fugitives.” Thus, the very premise on which the statement is based does not hold. {A-2}

We do not believe people can be divided into two groups, one group with a near zero chance of being counted and another with a near certain chance of being counted. Rather there are many different groups with many different chances. {A-2} Clearly, most people fall into the near certain group, which is why the census is, on average, so complete, but also why the PES is capable of counting many of the people who are missed by the census.

Mr. Darga acknowledges that overall demographic analysis results are very similar to the undercount rates based on the PES data. As a general observation, we note that we expect to find differences between different approaches. Based on our knowledge of their strength and weaknesses, we find the agreement between the results produced by the demographic analysis and the Dual System Estimation (DSE) undercount rates based on the PES to be reassuring. We would go so far as to argue that the agreement between the 1990 PES and demographic analysis on the undercount rate is more than a happy coincidence; it is remarkable and strengthens our belief about the overall credibility of the results. {A-1, A-6}

## **2. The PES is subject to errors.**

Mr. Darga examines different error sources in the PES and reaches the conclusion that 70 percent of the net undercount adjustment reflects measurement error rather than actual undercount. The Census Bureau readily acknowledges that there are sampling and nonsampling errors in the PES. In fact, in 1991 and 1992, an extensive evaluation program of the PES estimates was implemented. These evaluations addressed the potential sources of nonsampling error in the PES that could bias the results, including matching error and errors in determining erroneous enumerations. The results of these studies were combined to produce an estimate of the overall bias in the net undercount rate at the U.S. level and 13 high-level geographic areas. It was the finding that at the U.S. level, when correlation bias is taken into account, about 22 percent of the revised estimate of undercount (1.6 percent) was bias and not measured undercount. This is substantially less than the figure claimed by Mr. Darga. Even if the effect of correlation bias is ignored, our estimate of bias is well below the 70 percent referenced by Mr. Darga. It also should be noted that while Mr. Darga spends considerable time discussing nonsampling errors via short illustrations, no mention is made of what we actually know about the effects of these errors. {A-16}

### **3. The undercount rates for certain segments of the population are spurious.**

Mr. Darga contends that “many of the detailed undercount measurements for 1990 were implausible and, in some cases demonstrably false.” He states that “in an effort to correct a net national undercount of less than 2 percent, spurious undercounts of 10, 20 and even 30 percent were identified for some segments of the population.” Here, the reference appears to be to certain selected undercount rates for children under age 10 from the 1990 PES. This example is misleading. The undercounts of 10, 20 and 30 percent are rare exceptions, and did not occur for major segments of the population. In fact, in subsequent analyses, we find that only two poststrata in the final set of 357 PES poststrata were over 20 percent, 10 were from 15 to 20 percent, and 17 were 10 to 15 percent. In short, 328 of the 357 poststrata were less than 10 percent—not grounds for devastating impacts. {A-8}

Furthermore, it is unclear whether Mr. Darga’s calculations for children under age 10 are based on the “smoothed” estimates that were intended for use in Census adjustment or on the “raw” estimates which were not so intended. {A-7} The sample sizes for some of these poststrata were quite small, and the resulting variance of the raw estimates was quite large. Regardless, Mr. Darga does not discuss the role of bias and variance and the ensuing consequences. Selecting the “correct” set of factors involves consideration of both bias and variance, and that is why a detailed evaluation of the original factors was done. {A-7} Given what we now know about the small sample sizes and high variances of the estimates for children under 10, Mr. Darga’s use of age data to speculate about demographic trends of Asians or black homeowners is very misleading and inappropriate (pg. 16). {A-8}

### **4. The PES cannot be used to measure undercount.**

We obviously do not agree with Mr. Darga’s statement “...similar problems can be expected... for Census 2000: The problems are not due to minor flaws in methodology or implementation, but rather to the impossibility of measuring undercount through the sort of coverage survey that has been proposed.” With the statement, Mr. Darga appears to dismiss any adjustment methodology based on data from a coverage survey. We recognize that this is a convenient argument for Mr. Darga and one that allows him categorically to dismiss the PES.

We do not share Mr. Darga’s view that measuring the undercount with a coverage survey is impossible, though we do concede it is a challenging task. Furthermore, Mr. Darga fails to acknowledge years of research and development since the 1990 census. The statement deliberately ignores the progress that has been made in our understanding of ways to improve data collection and data processing. Technological innovations to facilitate quality control and improve coverage have been adopted for implementation in the 2000 census. Throughout the decade, we have continued to enhance our knowledge about the causes of undercount and census coverage errors in general. Of course, the lessons learned and the progress made are of little relevance to Mr. Darga’s position, but that does not undo the reality of their existence. {A-9} The Census Bureau is committed to continue its quest to overcome “minor methodological flaws and implementation errors” to ensure the high quality of its data products.



Further, Mr. Darga is so determined to focus on the coverage survey itself that he fails to acknowledge the strengths of the DSE methodology that makes use of the data. The focus is strictly on data collection and the subsequent matching operation associated with the PES. No attempt is made to fully describe all the steps involved in the PES, nor to explain the DSE methodology and the statistical model of capture-recapture. Had Mr. Darga focused on the DSE methodology rather than the coverage measurement survey itself, he would probably not have stated “that based on what was provided by the Census Bureau in 1990, one could be tempted to draw the conclusion that a coverage survey can provide an incredibly accurate measure of census undercount (pg.4).” The PES does not directly provide the undercount rate. It provides the data to be used for developing adjustment factors based on the DSE methodology. {B-1} The record should also reflect that the PES data were deemed of sufficient high quality to do so on the basis of evaluation criteria that were accepted and agreed upon prior to the 1990 census. Given the facts, Mr. Darga has no valid basis to conclude that adjustments based on the PES in 1990 “would have had a devastating impact on the usefulness and accuracy of census data at the state and local level (pg. 2).” {B-2}

#### **A. Conclusion**

The arguments and the viewpoints presented by Mr. Darga are not new to the adjustment debate and have entered into many adjustment deliberations. {B-3} Mr. Darga is a state demographer. We believe his contribution may lie in pointing out the value of building consistency and demographic validity checks into the evaluation of results. Mr. Darga has demonstrated how sex ratios can be a powerful evaluation tool. The Census Bureau welcomes constructive ideas on how to improve the census and how to judge its plausibility. It is important for demographers to enter the discussion and debate on the best way to produce a census that meets both statistical and demographic standards, but the message has to be fair.

cc: James F. Holmes  
Acting Director

---

---

Rejoinder to the Census Bureau's  
"Memorandum for the Record" Regarding  
Papers on Census Undercount Adjustment

---

---

Kenneth Darga, Senior Demographer  
Office of the State Demographer  
Michigan Information Center  
Michigan Department of Management and Budget

August 28, 1998

Please address comments and suggestions regarding this analysis to:  
DargaK@state.mi.us

This rejoinder assesses the Census Bureau’s official response to two papers which I submitted as testimony to the House Subcommittee on the Census on May 5, 1998. That response was written at the request of the Honorable Carolyn B. Maloney, the ranking minority member of the Subcommittee, who asked the Bureau to respond to all of the arguments in those papers.

The first section of this rejoinder addresses the Bureau’s response to each of the major arguments in the two papers. The second section addresses several additional counter-arguments which are advanced by the Bureau.

The following abbreviations are used in referring to previous documents:

<b>Abbreviation</b>	<b>Document</b>
Camel	“Straining out Gnats and Swallowing Camels: The Perils of Adjusting for Census Undercount” (paper by Kenneth Darga submitted to the House Subcommittee on the Census on May 5, 1998)
Quantifying	“Quantifying Measurement Error and Bias in the 1990 Undercount Estimates” (paper by Kenneth Darga submitted to the House Subcommittee on the Census on May 5, 1998)
Q&A	Response by Kenneth Darga dated June 19, 1998 to 25 questions asked by Representative Carolyn Maloney
Memorandum for the Record	Response by the U.S. Bureau of the Census to a request by Representative Carolyn Maloney to address all of the criticisms contained in the two papers referenced above.

## **B. Status of Major Arguments**

The central thesis of my two papers is that a “coverage survey” or “post-enumeration survey” cannot provide a reliable basis for adjusting the census for undercount because (a) a coverage survey misses many of the same people that are missed by the census, and (b) many of the people it identifies as “missed” by the census really have not been missed at all. Thus, the adjustments for undercount derived from a coverage survey are largely based upon the pattern of errors in measuring undercount rather than upon the pattern of undercount itself. Needless to say, the errors in measuring undercount are not necessarily distributed in the same way as undercount, which causes serious errors in the adjusted population counts. In the course of establishing this thesis, the following principal points are made:

### **1. Demographic Analysis**

**As the best available measure of undercount at the national level, the Census Bureau’s “demographic analysis” method provides reasonably good information about census undercount. (Camel, pp.2-3).**

**Status:** Not directly addressed, although the Census Bureau’s Memorandum for the Record implicitly accepts the validity of the findings of demographic analysis.

This point is not central to my argument, but it is significant nonetheless because it makes my thesis much easier to prove and much more difficult to dispute. It also establishes the basis for argument A-6 below.

The net national undercount suggested by the 1990 Post-Enumeration Survey (2.1% initially, and 1.6% after revision) is quite close to the net national undercount suggested by demographic analysis (1.8%). Proponents of adjustment like to point out the closeness of these figures. As the Census Bureau’s memorandum puts it:

(W)e find the agreement between the results produced by the demographic analysis and the Dual System Estimation (DSE) undercount rates based on the PES to be reassuring. We would go so far as to argue that the agreement between the 1990 PES and demographic analysis on the undercount rate is more than a happy coincidence; it is remarkable and strengthens our belief about the overall credibility of the results. (Memorandum for the Record, page 2, paragraph 4.)

Of course, this “remarkable” agreement between the two methods is equally consistent with the thesis that the Post-Enumeration Survey (PES) misses many of the people who were really missed by the census and that it identifies other people as “missed” when they really were not. In fact, this agreement makes it much easier to prove that thesis to be true: If it can be established that the PES really does miss many of the same people, then it follows that the only plausible way the PES can get so close to the “right” level of undercount is by identifying a similar number of people as “missed” when they really have not been missed at all. Likewise, if it can be established that many of the people that the PES identifies as “missed” really were not missed, then it follows that the most plausible way the PES can get so close to the “right” level of undercount is by missing a similar number of the people who were really missed by the census. Thus, establishing either half of the thesis establishes the other half of the thesis as well.

It should be noted that my papers provide very strong arguments for both parts of the thesis, and that each part of the thesis is established independently of the other and independent of any assertions of the validity of demographic analysis. The purpose of the preceding paragraph is to show that the Census Bureau (and other readers of my papers) cannot let one part of the thesis stand without accepting the other part as well (unless, of course, they are willing to deny the findings of the Bureau’s “demographic analysis” method after they have already been cited as “remarkable” support for the credibility of their results).

## 2. Correlation Bias (Missing the Same People Missed by the Census)

**A post-enumeration survey is not capable of counting many of the people who are missed by the census. In particular:**

- (a.) A coverage survey cannot measure the undercount of homeless people, and the PES therefore does not even attempt to address this portion of the undercount. (Camel, p. 5).**
- (b.) The number of people who do not want to be counted is very substantial (Camel, p. 1), and many of them can be expected to avoid the PES as well as the census (Camel, p. 5).**

**Status:** Discussed but not refuted.

The Census Bureau's memorandum addresses this argument by distorting it and then asserting that its distortions are not valid:

- The memorandum overstates my argument by asserting that:  
The claimed inability of the PES to count people who are missed by the census appears to be based on the premise that *most* of the people who are omitted from the census are “homeless” or “people who do not want to be counted.”  
(Memorandum for the Record, page 1, paragraph 4, emphasis added.)  
While that premise may well be true, my paper neither makes nor depends upon that premise. The argument actually made in my paper is that *many* of the people missed by the census fall into these categories. (Camel, p. 1, p. 5.)
- The memorandum further states that the emphasis on people who do not want to be counted:  
. . . seems to imply that the American people can be divided neatly into two groups: those who are nearly impossible to count and those who are trivially easy to count. (Memorandum for the Record, page 1, paragraph 4.)  
The memorandum goes on to describe various reasons for being missed by the census which cause people to fall between these two extremes, and then asserts:  
We do not believe people can be divided into two groups, one group with a near zero chance of being counted and another with a near certain chance of being counted. Rather there are many different groups with many different chances.  
(Memorandum for the Record, page 2, paragraph 3.)  
The problem with this argument is that there is actually nothing in my papers which suggests that the population “can be divided neatly into two groups.” This notion is entirely an invention of the Bureau's Memorandum for the Record. In fact, the second paragraph of my paper clearly states:

A major reason for the undercount—*although not by any means the only reason*—is that quite a few people do not want their identities known by the government. (Camel, p. 1; emphasis added.)

*Of course* there are other reasons why people are missed by the census. And since those reasons often cause some of the same people to be missed by the post-enumeration survey, they bolster my argument that a coverage survey misses *many* of the same people who are missed by the census.

- The memorandum states that:

There are indeed people who are actively hiding from the government and who are nearly impossible to count. However, it would not be true to say that *all* illegal immigrants are hiding from the government and are missed by the census. The same argument pertains to ‘drug dealers’ and ‘fugitives.’ Thus, the very premise on which the statement is based does not hold. (Memorandum for the Record, page 2, paragraph 2; emphasis added.)

*Of course* that would not be true—that is why my paper never says it. In fact, my paper points out that the net undercount of 5 million persons is remarkably low given the facts that the U.S. has over 1 million people who do not make any of their required payments on court ordered child support, 5 million illegal immigrants, and over 14 million arrests each year for non-traffic offenses. (Camel, p. 1.) Obviously, a substantial number of these people are counted by the census—either they respond to the census themselves, someone else responds on their behalf, or they are added to the census count through the imputation process for non-respondents and partial respondents. The argument that “the very premise on which the statement is based does not hold” is totally impertinent, since both the premise and the statement are merely inventions of the Bureau’s Memorandum for the Record.

### 3. Extreme Sensitivity to Small Classification Errors

**A very simple and very basic statistical phenomenon causes the undercount adjustments to be extremely sensitive even to very small errors in classifying people as “missed by the census” or “erroneously enumerated.” (Camel, pp. 6-8.)**

**Status:** Not addressed.

This is one of several arguments in my paper which is sufficient by itself to totally invalidate the attempt to measure undercount through a coverage survey.

### 4. Sources of Classification Error

**The attempt to measure undercount through a coverage survey is subject to many very serious sources of error. (Camel, p. 9; argument verified in Quantifying, pp. 3-13.)**

**Status:** Not addressed.

This argument complements the preceding argument: because the adjustments are so sensitive to errors in classifying people as “missed by the census” or “erroneously enumerated,” these sources of error have a devastating impact on the accuracy of the adjustments.

## **5. Classification Errors are Reflected in the Adjustments**

**Because the coverage survey misses many of the people missed by the census and identifies other people as missed by the census when they really were not missed, the differential undercounts it suggests will largely reflect *differences in the amount of error in measuring undercount* rather than differences in the amount of undercount itself. (Camel, p. 10.)**

**Status:** Not addressed.

## **6. Inconsistency with Demographic Analysis**

**The final adjustments based on the 1990 PES are quite different from the estimates based on the Census Bureau’s “demographic analysis” method even for very broad population groups at the national level. (Camel, p. 11.)**

**Status:** The Memorandum for the Record asserts the opposite, but without presenting evidence or refuting the contrary evidence in my papers.

The memorandum states:

Mr. Darga acknowledges that overall demographic analysis results are very similar to the undercount rates based on PES data. As a general observation, we note that we expect to find differences between different approaches... (W)e find the agreement between the results produced by the demographic analysis and the Dual System Estimation (DSE) undercount rates based on the PES to be reassuring. (Memorandum for the Record, p. 2, paragraph 4.)

It should be noted that I acknowledge a similarity to the findings from demographic analysis *only for the overall total population figure*. The similarity breaks down very seriously as soon as one starts to examine the results in any detail. Thus, my paper states that:

(T)he final national PES results for 1990 are actually quite different from the estimates based on demographic analysis even for very broad population groups. The apparent undercount for black males is 42% less than the rate suggested by demographic analysis, and the rate for white, Native American, and Asian/Pacific females is 50% higher. (Camel, p. 11.)

It is hard to imagine what reassurance the Census Bureau finds in these discrepancies. (This issue is also discussed in Q&A, question 6, pp. 9-10.)

## **7. Spurious Undercount Differentials**

**Some of the large undercount differentials suggested by the 1990 post-enumeration survey are definitively shown to be spurious.**

- (a.) The 18 large differential undercounts between girls and boys displayed in Figure 3 are implausible and they follow no discernible pattern. (Camel, pp. 11-12.)**
- (b.) Because of the stability of the sex ratio in this age range, these differential undercounts can be tested definitively. (Camel, p. 13.)**
- (c.) These differentials are clearly spurious. The areas in question show no sign of differential undercount between boys and girls prior to adjustment. After adjustment based on the PES, the sex ratio in these areas is dramatically different from the norm. (Camel, pp. 14-15.)**
- (d.) The problems revealed here pertain just as much to other age groups as to children and just as much to other demographic characteristics as to the sex ratio. Because these undercount differentials are clearly spurious, we cannot trust a coverage survey to tell us which segments of the population have higher undercounts than others. (Camel, p. 15.)**

**Status:** Not refuted.

Although the Census Bureau's response does not address any of the elements of this argument listed above, it does raise two related issues:

- The Bureau's memorandum claims:  
(It is unclear whether Mr. Darga's calculations for children under age 10 are based on the 'smoothed' estimates that were intended for use in Census adjustment or on the 'raw' estimates which were not so intended. (Memorandum for the Record, page 3 paragraph 2.)

Because of conflicting information which I had received, this point was indeed unclear in the preliminary drafts of my papers which I sent to the Census Bureau for review between July 1997 and April 1998. This point was not clarified until the end of April, and the clarification is reflected in the paper submitted to the Subcommittee on the Census which the Bureau was asked to address. It is clearly stated there that these are the initial adjustment factors, prior to the application of a statistical smoothing procedure. (Camel, page 12, note 14.) It is further explained that these non-smoothed factors are pertinent for the current analysis, since they reflect the amount of apparent undercount actually identified by the PES. These non-smoothed factors are also the ones



most relevant in the context of Census 2000, since the Census Bureau does not plan to use a statistical smoothing process in the adjustment of the next census. (See Camel, pp. 15-16, note 18.)

- The Bureau's memorandum claims:

The sample sizes for some of these poststrata were quite small, and the resulting variance of the raw estimates was quite large. Regardless, Mr. Darga does not discuss the role of bias and variance and the ensuing consequences. Selecting the "correct" set of factors involves consideration of both bias and variance, and that is why a detailed evaluation of the original factors was done. (Memorandum for the Record, page 3, paragraph 2.)

If the Census Bureau is aware of any consequences of bias and high variance which are favorable to their proposed methodology, I invite them to explain how they mitigate the arguments presented in my papers.

I agree that the sample sizes for most poststrata—i.e. for most designated components of the population—were quite small, and that there were very serious problems with sampling error (i.e. the variance of both the raw estimates and the final estimates was very large). However, it would be a mistake to think that large sampling errors matter any less than large non-sampling errors. To a data user, large errors are equally serious regardless of their source.

It is certainly not accurate to suggest that my papers do not discuss the role of bias and its consequences: That is what my papers are all about. My treatment of sampling error, however, is very brief but very pertinent:

There are several types of measurement error. Although the point being made here is that the large amount of error in the adjustments is consistent with the thesis that large amounts of *non-sampling* error are inevitable, it should be noted that *sampling* error is also a very serious problem for the undercount adjustments. Actually, there is more than enough error to go around: these adjustments can reflect a very large amount of sampling error as well as a very large amount of non-sampling error. For purposes of data quality, both types of error are very problematic. (Camel, p. 13, note 15.)

The role of sampling error is discussed in more detail in my letter to Representative Maloney dated 6/19/98 in response to the 25 questions which she posed subsequent to my testimony to the Subcommittee on the Census. The limitations of sampling error as an explanation for the shortcomings of the undercount adjustments are discussed in response to question 16, and the limited impact of a larger sample size on total error and sampling error is discussed in response to questions 17 and 20. (Q&A, pp. 17-18, 20-21.)

In summary, the Bureau's response does not address my arguments directly, and the points which it does raise do not weaken my arguments.

## 8. Some Implications of Faulty Adjustments

**Inaccurate adjustments would destroy the reliability of Census data at the state and local level.**

- (a.) Errors would sometimes be large. (Camel, pp. 11-15.)**
- (b.) Errors for a given segment of the population can be expected to differ from one census to another, which would invalidate comparisons of census data over time. (Camel, p. 16)**
- (c.) These errors would have a serious impact on policy decisions and on our understanding of trends in our communities. (Camel, pp. 16-17.)**
- (d.) The presence of significant unpredictable errors would make all census comparisons unreliable. When the census suggested a change in population trends, data users would not know how much of the change represented actual demographic trends and how much represented spurious differences in the undercount adjustments. (Camel, p. 17)**

**Status:** Subpoint (a) and subpoint (c) are addressed but not refuted. The other points are not addressed.

The Bureau's memorandum makes the following two arguments:

- Subpoint (a) is addressed by stating:  
This example (i.e. the spurious differential undercounts for children) is misleading. The undercounts of 10, 20, and 30 percent are rare exceptions, and did not occur for major segments of the population. In fact, in subsequent analyses, we find that only two poststrata in the final set of 357 PES poststrata were over 20 percent, 10 were from 15 to 20 percent, and 17 were 10 to 15 percent. In short, 328 of the 357 poststrata were less than 10 percent—not grounds for devastating impacts. (Memorandum for the Record, page 3, paragraph 1; parenthetical comment added for clarity.)

This statement raises several issues and questions.

- ◆ The Bureau argues that large adjustments are “rare exceptions,” but then points out that nearly 8% of the 357 final “collapsed” adjustments were over ten percentage points—certainly opportunity enough for a substantial number of serious errors to occur.
- ◆ Large adjustments were even more frequent for the initial non-smoothed factors, which are the ones most pertinent to the analysis in my paper and most pertinent to the methodology proposed for Census 2000: 218 (16%) of the 1392 non-smoothed adjustment factors exceeded 10 percentage

points, ranging from a downward adjustment of 36 percent to an upward adjustment of 51 percent.

- ◆ What does the Census Bureau consider to be a “minor” segment of the population for which large errors would be acceptable? As one looks at the specifications for the 357 collapsed poststrata (or, more pertinently, the specifications for the 18 pairs of poststrata for which undercount rates are listed on page 12 of my paper), none of them appears to be unimportant or insignificant. The census is often relied upon for data on small segments of the population. If the community or population group on which a data user must focus is one of the ones affected by large errors, it would be of little comfort to know that most of the errors elsewhere are smaller.
- ◆ How “large” would an error have to be to be serious? Although the differential undercounts which my paper demonstrates to be spurious all exceed ten percentage points, it should be noted that even much smaller errors can be serious. Differences of a few percentage points or a few tenths of a percentage point can have significant implications for policy decisions, for resource distribution, and for understanding demographic trends in our communities.
- ◆ The impact of any errors is compounded by the fact that data users would not know which areas and population groups have serious errors and which do not: The uncertainty resulting from large errors hidden throughout the data from the next census—regardless of whether they are hidden thickly or sparsely—would have a “devastating impact” by itself.
- ◆ Finally, it should be noted that the spurious undercounts identified in my paper are accurately described and they are pertinent to the conclusions which are drawn from them. They therefore would not have been “misleading” even if the comments in the Bureau’s memorandum were valid.
- With regard to the hypothetical examples used to illustrate subpoint (c.), the Bureau states:

Given what we know about the small sample sizes and high variances of the estimates for children under 10, Mr. Darga’s use of age data to speculate about demographic trends of Asians or black homeowners is very misleading and inappropriate. (Memorandum for the Record, page 3, paragraph 2.)

I have several responses to this observation:

- ◆ I will readily grant that the adjustments for children under 10 reflect very high levels of sampling error (as well as very high levels of non-sampling error). If the statement quoted above is intended to imply that the estimates for different age groups are less beset by these problems, then some evidence to that effect should be provided.
- ◆ The argument that these illustrations are misleading and inappropriate seems to be based on the observation that the sampling error for a very small group (e.g. persons *under the age of 10* within a particular segment of the population) would tend to be higher than the sampling error for a somewhat larger group (e.g. persons *of all ages* within that segment of the population). If the variability of the adjustment factors from one census to another were solely due to statistically “well-behaved” sampling error, this criticism of the illustrations would have merit. However, several additional sources of variability must be considered:
  - Methodologies used in the coverage survey can change significantly from one census to the next. For example, changes planned between 1990 and 2000 include computing adjustment factors for individual states instead of for multi-state regions, a shorter period of time in which to conduct interviews, and a different choice of weeks for interviews.
  - A given state might have a spell of hot or rainy weather during one PES but not during the next. This can significantly affect several factors which influence the adjustments, including the rate of successful interviews, the percentage of homeless people who are found in households, and the percent of interviews fabricated by enumerators. (See, for example, *Quantifying*, pp. 5-7.)
  - The undercount rates for a given area might be strongly affected by a few aberrant blocks for one census but not for another. An extreme example of the impact of aberrant blocks is the two block clusters (out of a total of 5,290) which by themselves would have accounted for about 15% of the net national undercount in 1990 due to geocoding errors if they had not been identified and corrected. (See *Quantifying*, p. 9.) Of course, block clusters *cannot* be divided neatly into two groups: those for which errors are blatantly obvious and those for which the measure of undercount is practically perfect. Block clusters can be aberrant for many reasons. There can be a severe or modest

number of geocoding errors, a new housing development, a large number of seasonal dwellings, a university with exams during the period in which census interviews are conducted, a retirement community or another special population group, a bad census enumerator, a bad PES enumerator... The list of examples is endless. As demonstrated by the example above, a few aberrant blocks can cause variations even greater than those in the illustrations in question.

- ◆ Finally, it should be noted that the illustrations in question do not really “speculate about demographic trends of Asians or black homeowners.” Rather, they are used to illustrate how variations in error levels from one census to the next would have serious implications. Regardless of whether one likes the hypothetical examples that are used to make this point, the point itself still remains.

## 9. Relevance to Census 2000

**The problems with the 1990 adjustments can be expected to recur in Census 2000 if the proposed methodology is used. They are not due to minor correctable flaws in methodology or implementation, but rather to the impossibility of measuring undercount through the proposed coverage survey. (Camel, p. 18.)**

**Status:** Discussed but not refuted.

Most of the arguments in my papers involve problems which are either inherent in the effort to measure undercount with a post-enumeration survey, or else so intractable that they cannot be corrected.

The Bureau addresses this argument by asserting

We do not share Mr. Darga’s view that measuring the undercount with a coverage survey is impossible, though we do concede it is a challenging task. Furthermore, Mr. Darga fails to acknowledge years of research and development since the 1990 census. The statement deliberately ignores the progress that has been made in our understanding of ways to improve data collection and data processing. Technological innovations to facilitate quality control and improve coverage have been adopted for implementation in the 2000 census. Throughout the decade, we have continued to enhance our knowledge about the causes of undercount and census coverage errors in general. Of course, the lessons learned and the progress made are of little relevance to Mr. Darga’s position, but that does not undo the reality of their existence. (Memorandum for the Record, page. 3, paragraph 4.)

Although I am aware of several innovations planned for the next census and the next post-enumeration survey, I am not aware of any which would enable the next

post-enumeration survey to succeed where the previous one failed. If there are such innovations, I invite the Bureau to show specifically how they negate each of the arguments in my papers. Until that is done, I must agree that “the lessons learned and the progress made are of little relevance to Mr. Darga’s position.”

## **10. Matching Error**

**The key to measuring undercount with a coverage survey is to match each person’s survey record with the corresponding census record. However, when the same records were matched by different teams of trained personnel using the same definitions and guidelines, the disagreement rate was very high relative to the size of the net undercount that the 1990 Post-Enumeration Survey was trying to measure. (Quantifying, p. 3.) Implications of this finding include:**

- (a.) The number of difficult cases for which match status is not obvious is very large, greatly exceeding the estimated level of net undercount. This demonstrates the impossibility of measuring undercount accurately through a coverage survey even apart from any other considerations. (Quantifying, p. 4.)**
- (b.) The high level of disagreement suggests that many of the judgments reached by the final team of matchers are likely to be wrong. (Quantifying, p. 4.)**
- (c.) The level of subjectivity demonstrated by the high rate of disagreement makes the adjustments vulnerable to bias through expectations and other impertinent factors. (Quantifying, p. 4.)**
- (d.) The high level of disagreement between matchers causes the results for a given set of records to be different each time the match is performed. (Quantifying, pp. 4-5.)**

**Status:** Not addressed.

## **11. Fabrication of Data**

**Fabrication of data by interviewers is another problem that is sufficient by itself to invalidate the adjustments for undercount derived from a post-enumeration survey.**

- (a.) The level of fabrication in typical Census Bureau surveys is very substantial relative to the level of net undercount that the post-enumeration survey attempts to measure. (Quantifying, p. 5.)**
- (b.) Fabrication in either the census or the PES can cause very serious errors in the undercount adjustments. (Quantifying, pp. 5-6.)**

- (c.) **Taken together, the three studies of fabrication in the 1990 PES suggest that its level of fabrication may have been close to the level found in other Census Bureau surveys. (Quantifying, p. 6.)**
- (d.) **Apparent levels of fabrication varied substantially among regions. The regions which appeared to have the highest levels of fabrication were regions with high adjustments for undercount, and they also had very hot or rainy weather during the period in which PES interviews were conducted.**

**Status:** Not addressed.

## **12. Ambiguous “Usual” Place of Residence**

**The number of people with an ambiguous “usual” place of residence poses serious problems for undercount adjustments derived from a coverage survey.**

- (a.) **The number of people with an ambiguous “usual” place of residence is very substantial relative to the level of net undercount. (Quantifying, p. 7.)**
- (b.) **The adjustments derived from the coverage survey can have a significant impact on the regional population distribution by replacing the traditional concept of “usual” address, which is defined largely by the respondent, with a set of assignment rules developed for the coverage survey. (Quantifying, pp. 7-8.)**
- (c.) **Neighborhoods vary greatly in their proportion of people with an indistinct “usual” place of residence. The adjustments for a class of cities in an entire state or region can be determined largely by whether or not the sample includes a few blocks which are outliers in this respect. (Quantifying, p. 8.)**

**Status:** Not addressed.

## **13. Geocoding Errors**

**Geocoding errors pose very serious problems for undercount adjustments derived from a coverage survey.**

- (a.) **Geocoding errors can cause errors in classifying people as missed by the census, correctly enumerated, or erroneously enumerated. These errors can cause bias in the undercount adjustments, since they do not necessarily cancel one another out. (Quantifying, pp. 8-9.)**
- (b.) **The size of the net undercount is very sensitive to the size of the search area for records with inaccurate geographic codes. This illustrates the**

**sensitivity of the undercount adjustments to minor variations in the procedure for conducting and analyzing the coverage survey. (Quantifying, p. 9.)**

- (c.) Geocoding errors in only 2 block clusters (out of a total of 5,290) caused them to contribute nearly a million people to a preliminary calculation of net undercount. This illustrates the extreme sensitivity of the undercount adjustments to small errors, their sensitivity to a few outlier blocks, and the importance of minor variations in methodology such as criteria and methods for correcting errors that are discovered. (Quantifying, p.9.)**

**Status:** Not addressed.

#### **14. Unreliable Interviews**

**Unreliable interviews pose a very serious problem for undercount adjustments derived from a coverage survey. The percent of records which changed match status due to different information in re-interviews conducted for evaluation purposes was very large relative to the level of net undercount. (Quantifying, p. 10.)**

**Status:** Not addressed.

#### **15. Unresolvable Cases**

**The number of cases in the 1990 Post-Enumeration Survey which were unresolved even after repeated interview attempts was very substantial relative to the level of net undercount. The uncertainty resulting from such a large number of unresolved cases is a fatal flaw in the undercount measurements.**

**Status:** Not addressed.

#### **16. Combined Impact of Errors**

**A very substantial proportion of the apparent net undercount identified through the 1990 coverage survey was actually caused by bias due to various errors that were identified and documented in the Census Bureau's evaluation reports. In other words, it turned out that many of the people identified as "missed" by the census actually had not been missed at all. (Quantifying, pp. 11-13.)**



**Status:** Discussed but not refuted.

The Bureau responded to this argument by stating:

It was the finding that at the U.S. level, *when correlation bias is taken into account*, about 22 percent of the revised estimate of undercount (1.6 percent) was bias and not measured undercount. This is substantially less than the figure claimed by Mr. Darga. Even if the effect of correlation bias is ignored, our estimate of bias is well below the 70 percent referenced by Mr. Darga. It also should be noted that while Mr. Darga spends considerable time discussing non-sampling errors via short illustrations, no mention is made of what we actually know about the effects of these errors. (Memorandum for the Record, page 2, paragraph 5, emphasis added.)

The disagreement on this point between my paper and the Bureau's memorandum is much smaller than it may seem at first: I make no objection to the figures which the Bureau cites, except to note that they are not as pertinent to my arguments as the figures from the same studies which are cited in my papers.

My observations with respect to the Bureau's comments are as follows:

- ◆ Attributing 22% of the national undercount adjustments to bias should provide very scant comfort to proponents of adjustment. Even if the bias problem were no worse than that, a 22% bias would be sufficient to invalidate the adjustments.
- ◆ An overall national estimate of bias—whether the 22% figure cited by the Bureau or the more pertinent figures described below—reflects some areas and some segments of the population that have higher levels of bias and others that have lower bias or even bias in the opposite direction. Admitting an overall bias of 22% therefore amounts to an admission that some parts of the country have adjustment factors which are in error by more than 22%.
- ◆ The 22% figure cited by the Bureau reflects the amount of bias which remains *after incorporating the effects of correlation bias*. This is perfectly consistent with the central thesis of my papers, i.e. that a post-enumeration survey cannot provide a reliable basis for adjusting the census for undercount because (a) a post-enumeration survey misses many of the same people that are missed by the census (“correlation bias”), and (b) many of the people it identifies as “missed” by the census really have not been missed at all. Based on the Census Bureau's “demographic analysis” findings, I have suggested that these two errors largely cancel one another out with respect to the overall national measure of undercount. The Bureau here suggests that the two errors do not come quite so close to canceling one another out: They contend that the undercount estimates still have a 22% upward bias even after incorporating the effects of correlation bias. Although the inconsistency between this figure and the findings of demographic analysis leads me to be somewhat skeptical of it, I am willing to accept it at face value for purposes

of this discussion since, as explained below, it is irrelevant to the figures in my paper which the Bureau is attempting to challenge.

- ◆ The 70% figure which the Bureau attempts to challenge is used in my paper to make the argument that a large number of people are falsely identified as “missed” by the census, and that they are *subsequently* offset by correlation bias. Obviously, the pertinent figure to use for such a purpose would *not* already subtract the effects of correlation bias.
- ◆ There are at least two reasons why the Bureau’s estimate of bias is below the 70% figure cited in my paper even when they do not subtract the effects of correlation bias. The first reason would be evident from a more careful reading of my paper:

The analysis in the Census Bureau’s P-16 report indicates that the corrections for measurement errors in the 1990 PES would have decreased the undercount estimate from 2.1% to 1.4% [i.e. about 33%] . . . A later analysis by the same author incorporated additional corrections related to a major computer processing error discovered by Census Bureau in late 1991, the re-matching of records in some suspect blocks, and the inclusion of very late Census data that had not been available when the initial PES estimates were developed. This analysis suggested that corrections for identified measurement errors would have reduced the undercount estimate from 2.1% to 0.9% [i.e. about 57%] . . . An analysis by Dr. Leo Breiman, which built upon the Census Bureau analyses cited above, incorporated additional sources of error to arrive at an adjusted undercount estimate of only 0.6% [i.e. about 70% lower than the 2.1% figure]. (Quantifying, p. 13; bracketed phrases added for clarity.)

Thus, one of the reasons why the Bureau’s figures are different from mine is that they do not take into account all of the factors that were included in Dr. Breiman’s analysis. While it is certainly true that 33% and 57% are lower than 70%, it is hardly noteworthy. Another reason for the discrepancy is not quite so obvious: my figures are intended to reflect all of the identified errors in the adjustments that had been proposed for the 1990 Census counts, but the figures cited in the Bureau’s Memorandum for the Record reflect only errors that the Bureau did not subsequently correct. The errors that the Bureau chose to correct are just as pertinent to my arguments as the errors which the Bureau identified but did not correct.

- ◆ I am puzzled by the statement that:  
While Mr. Darga spends considerable time discussing non-sampling errors via short illustrations, no mention is made of what we actually know about the effects of these errors. (Memorandum for the Record, p. 2, paragraph 5.)  
I certainly try to mention the things that *I* know about the effect of those errors. (See Camel, pp. 13-17, Quantifying, pp. 3-11, and particularly pp. 11-13. See also Q&A, questions 3 and 4, pp. 7-8.) I invite the Bureau to share its additional insights regarding the effects of these errors, and to indicate how they amplify or refute each of the arguments in my papers.

## C. Counter-Arguments Raised by the Census Bureau

The Census Bureau's Memorandum for the Record raises several counter-arguments which are not associated with any of the individual arguments raised in my paper:

### 1. Focus on Underlying Data Instead of Subsequent Calculations

**The Bureau's memorandum states:**

**Mr. Darga is so determined to focus on the coverage survey itself that he fails to acknowledge the strengths of the DSE [Dual-System Estimation] methodology that makes use of the data. The focus is strictly on data collection and the subsequent matching operation associated with the PES. No attempt is made to fully describe all the steps involved in the PES, or to explain the DSE methodology and the statistical model of capture-recapture. Had Mr. Darga focused on the DSE methodology rather than the coverage measurement survey itself, he would probably not have stated "that based on what was provided by the Census Bureau in 1990, one could be tempted to draw the conclusion that a coverage survey can provide an incredibly accurate measure of census undercount . . ." The PES does not directly provide the undercount rate. It provides the data to be used for developing adjustment factors based on the DSE methodology. (Memorandum for the Record, page 4, paragraph 1.)**

I have several observations with respect to this counter-argument:

- ◆ Addressing the last point first, it is true that, strictly speaking, the Post-Enumeration Survey produces only raw data. The subsequent analysis of that data involves matching survey records with census records and using the results of that matching process in a formula to produce the actual undercount adjustments. Like other writings on this subject, my papers sometimes use the terms "coverage survey" or "PES" to encompass the survey itself, the subsequent analysis, and the results of that analysis. See, for example, the statement in the Bureau's Memorandum for the Record:

We would go so far as to argue that the agreement between the 1990 PES and demographic analysis on the undercount rate is more than a happy coincidence. (Memorandum for the Record, page 2, paragraph 4.)

I would also add that the phrases which are presented as a quotation from my paper appear to be merely a paraphrase by a Census Bureau analyst. The sentence which actually appears in my paper is:

Thus, one is tempted to conclude that data from a coverage survey can provide an incredibly accurate measure of Census undercount. (Camel, p. 4).

- ◆ Turning to the more substantial point raised in the Bureau's argument, I readily agree that my papers focus upon the coverage survey and upon the data

developed for use in the adjustment formula rather than upon the adjustment formula itself. In my first statistics course as an undergraduate, I was taught that the Fundamental Law of Statistics is “Garbage in, garbage out.” If the Dual System Estimation methodology has the ability to produce correct and reliable adjustments by treating survey findings as accurate when they are in fact dominated by measurement errors, I invite the Bureau to explain exactly how that is accomplished. Until that is done, I will assume that the Fundamental Law of Statistics is still in effect.

- ◆ It is also true that my papers do not use the term “capture-recapture methodology.” Although I have tried to avoid technical terminology and theoretical discussions as much as possible in my papers, it is nonetheless fruitful to review the capture-recapture model and the assumptions upon which it relies.

The capture-recapture model is most widely used in wildlife biology. A common illustration is that, if you catch and mark a certain number of fish on one day, you can estimate the total number of fish in the lake by catching some fish on another day, and then assuming that the proportion of fish with marks tells you the proportion of the total fish population that was caught on the first day. For example, if fifty percent of the fish which you catch on the second day have marks, you assume that you had marked fifty percent of the fish in the lake on the first day.

As any fisherman can probably guess, this method does not always produce accurate results. According George Seber’s *Estimation of Animal Abundance and Related Parameters* (New York, 1982), this sort of capture-recapture model can produce suitable results when certain assumptions are met, such as:

- (a.) The population is closed so that  $N$  is constant. (This assumption is obviously violated in the attempt to measure undercount with a coverage survey. A substantial number of people are born, die, or move from one place to another between the Census and the survey. It is very difficult to compensate for this problem, and it is one of the serious sources of error discussed in my papers.)
- (b.) All animals have the same probability of being caught in the first sample. (In the context of measuring undercount in the U.S. Census, this assumption translates into a premise that all people within a poststratum—i.e. within a designated component of the population—have the same probability of being counted in the Census. However, as discussed under argument A-2 above, some people purposely avoid being counted by the census, and others are not counted because of

various other factors that make them difficult to count. Thus, this assumption is also violated.)

- (c.) Marking does not affect the catchability of an animal. (In the context of measuring undercount in the U.S. Census, this translates into a premise that people counted by the census are just as likely to be counted in the coverage survey as people in the same poststratum who are missed by the census. Obviously, when people are missed because they want to be missed or because they are particularly difficult to count, this assumption is not met.)
- (d.) Animals do not lose their marks in the time between the two samples, and all marks are reported on recovery in the second sample. (In the context of measuring undercount in the U.S. Census, this assumption translates into a premise that all of the people in the sample who were counted by the Census are successfully matched with their census records. One reason for the failure of the matching process is that the marking techniques used by wildlife biologists are not suitable for use by census enumerators. Thus, as discussed at length in my papers, people counted by the census can be falsely classified as having been missed.)

Thus, one way of summarizing many of the arguments in my papers is to say that the Census Bureau's methodology for developing undercount adjustments violates the fundamental assumptions upon which the underlying capture-recapture model is based.

## 2. Sufficiency of PES Data

**The Bureau's memorandum states:**

**The record should also reflect that the PES data were deemed of sufficient high quality to do so [sic] on the basis of evaluation criteria that were accepted and agreed upon prior to the 1990 census. Given the facts, Mr. Darga has no valid basis to conclude that adjustments based on the PES in 1990 "would have had a devastating impact on the usefulness and accuracy of census data at the state and local level." (Memorandum for the Record, page 4, paragraph 1.)**

This seems to say that the 1990 undercount adjustments were good enough to be used for the 1990 Census. This overlooks the fact that the Secretary of Commerce decided not to apply the 1990 adjustments to the Census, and the director of the Census Bureau decided not to apply them to the intercensal estimates. These decisions were based on sound arguments, as demonstrated by the explanations which accompanied them. (See *Federal Register*, 7/22/91, pp. 33582-33642, and *Federal Register*, 1/4/93, pp. 69-78.)

The sentence which begins “Given the facts . . .” is difficult to reply to, since no facts appear either in this paragraph or in the remainder of the Bureau’s memorandum which support the subsequent statement. If the Bureau is aware of such facts, they should be presented along with an explanation of how they negate arguments 5, 6, 7, and 8 above.

### **3. Arguments Presented Are Not New**

**The Bureau’s memorandum states:**

**The arguments and viewpoints presented by Mr. Darga are not new to the adjustment debate and have entered into many adjustment deliberations. . . Powerful arguments about coverage measurement have been made in support of adjustment of the decennial census as well as against adjustment. (Memorandum for the Record, page 4, paragraph 2; page 1, paragraph 3.)**

If the Bureau’s Memorandum for the Record is any indication, the fact that these arguments are not new should not at all suggest that they have been refuted.

I agree that the evidence about the undercount adjustments presents a paradox. On the one hand, there is very strong evidence to show that the undercount adjustments are based on a variety of very serious errors. And yet, on the other hand, there are some respects in which the undercount adjustments look like one would expect valid adjustments to look. However, a paradox should *not* be seen as an opportunity to simply choose which evidence one wishes to accept and which evidence one wishes to ignore. A paradox demands an explanation. Either the contradictory evidence must be refuted, or else a new understanding must be reached through which the apparent contradiction can be resolved. Throughout my papers, I have therefore attempted to explain how the apparent strengths of the adjustments can be explained in the context of their weaknesses.

One of the most powerful arguments in favor of the adjustments has been the “remarkable” closeness of the overall net national undercount suggested by the 1990 PES to the overall net national undercount suggested by demographic analysis. However, this apparent strength of the adjustments can be explained very well in the context of their weaknesses:

- ◆ My papers show this finding to result from missing a lot of the same people who were missed by the census, and then identifying a similar number of people as having been missed when they actually were not missed at all. (Argument A-2, and Arguments A-3, A-4, and A-10 through 16 above.)
- ◆ The closeness of the two results also reflects the influence of expectations about undercount upon the adjustment factors. (Q&A, questions 1 and 14, pp. 1-3, 16.)

Another powerful argument in favor of the adjustment factors is the (very rough) similarity between the groups which tend to have high adjustment factors and the groups which would be expected to have high undercounts. This apparent strength of the adjustments can also be explained in the context of their weaknesses:

- ◆ It reflects a substantial overlap between the groups which are hard to count and the groups for which it is difficult to match survey records with census records. (Camel, pp. 4, 9.)
- ◆ It reflects the fact that fabrication of records, which tends to cause a low match rate and a high apparent level of undercount, tends to be more frequent in neighborhoods which interviewers perceive as dangerous. (Quantifying, p. 7.)
- ◆ It reflects the role of expectations in determining whether an uncertain survey case will be classified as “matched” or “not matched” with the census. (Quantifying, p. 4, Q&A, question 14, p. 16.)

Perhaps the greatest paradox involves the enormous error rates identified in the Census Bureau’s evaluations of the 1990 Post-Enumeration Survey (see Argument A-16 above): How can such high levels of error be consistent with the high levels of skill and care with which the PES was obviously conducted? Perhaps the Census Bureau’s readiness to reject the evidence against the PES without refuting it rests upon a failure to resolve this paradox. To those familiar with the skill, credentials, and conscientiousness of the team which designed and implemented the PES, the error levels identified by the Census Bureau’s evaluations must seem unbelievable. Nevertheless, there is a resolution for this paradox as well. Given the extreme sensitivity of a coverage survey to very small mistakes (Argument A-3 above) and the many serious sources of mistakes (Argument A-4 above), it is inevitable for the adjustments to be dominated by errors in measuring undercount. It is not the documented failure of the adjustment methodology which is unbelievable, but rather the blithe assumption that skill and hard work can overcome the fatal flaws inherent in the Bureau’s methodology.

Thus, the arguments presented here not only make a strong case against the proposed methodology, but they also show how its apparent strengths are consistent with its documented weaknesses. In order to prevail in the current debate, the Census Bureau must either refute the sixteen arguments listed above, or else show in a similar manner that the unrefuted arguments can somehow be made consistent with the thesis that we can count upon the undercount adjustments to be highly accurate. I believe that this task will prove to be as impossible as deriving accurate undercount adjustments from a coverage survey.