Component of Statistics Canada
Catalogue no. 12-001-X Business Survey Methods Division

Article

The past is prologue

by Barbara A. Bailar

June 2000

Statistics Canada Statistique Canada

Canada

Survey Methodology, June 2000
Vol. 26, No. 1, pp. 21-30
Statistics Canada, Catalogue No. 12-001

The past is prologue

Barbara A. Bailar [1]

Abstract

Mahalanobis provided an example of how to use statistics to enlighten and inform government policy makers. His pioneering work was used by the US Bureau of the Census to learn more about measurement errors in censuses and surveys. People have many misconceptions about censuses, among them who is to be counted and where. Errors in the census do occur, among them errors in coverage. Over the years, the US Bureau of the Census has developed statistical techniques, including sampling in the census, to increase accuracy and reduce response burden. A root-mean-square-error model was developed to estimate the joint effects of variance and bias in the census. The model is used in this paper to look at the joint effects of response variance, adjustment of the bias caused by the undercount, and the use of sampling for follow-up.

Key Words: Censuses; Mahalanobis; Root-mean-square-error model; Sampling in the census.

1. Introduction

Perhaps it has always been so – that statistics, as a body of information, does not always support the actions that politicians want to take. In some countries, data from censuses are not made public, because knowledge is power. However, in our society, the power of statistics is used to inform us about needs for action, or how well we're doing as a country, or as the basis of comparison among groups. We are used to seeing and trusting statistics on an everyday basis, though most of us give little attention to how they are produced, by whom, and at what cost.

Over the last few decades, there have been many issues where statistics and politics have been in conflict. Employment and unemployment data are often used by politicians, especially in an election year. If the unemployment figures are low, the incumbents cite that figure and take the credit. If the employment figures show that many new jobs are being created, that number is cited. Either political party can use these data to make whatever political points seem salient. An attempt by the Nixon Administration to restrict access to these data led to new protections, such that the employment and unemployment data are released on the first Friday of every month by the Commissioner of the Bureau of Labor Statistics at a meeting of the Joint Economic Committee on Capitol Hill.

The definition of poverty is currently under discussion. When the poverty measure was invented by Molly Orshansky, there were not the large transfer payment systems that exist today. Because of income received or benefits paid, poverty today does not mean what poverty did 30 years ago. However, each political administration watches the poverty numbers very closely. These numbers were used by critics of the Reagan Administration to illustrate the growing burden of the poor in an administration that was alleged to be more interested in serving the rich. That Administration argued that by including medical benefits and other transfer payments, the poor were better off than before.

Probability samples of the U.S. population are now used to study sexual behavior. Much of our information on sexual behavior goes back to Kinsey. The National Opinion Research Center (NORC) at the University of Chicago has conducted two large surveys of sexual behavior in the U.S. One of these, *Sex in America*, (Michael, Gagnon, Laumann and Kolata 1994) reported on a national sample of persons aged 18-59, and was not funded by the government. The second researched the sexual behavior of adolescents and, in both cases, federal funding for these studies was questioned because powerful constituencies did not want the subject matter to be examined. The second study was finally funded by the government.

Privacy issues abound. For example, there is broad concern about the confidentiality of individual medical records and the need for researchers to access them. Privacy issues for groups are less widely recognized. Certain groups may not want to report fully in a decennial census or survey because they do not want to attract attention. Though people who are in the country illegally are supposed to be included in the census, many of them fear that government authorities looking at block statistics could use the information to raid certain blocks.

My last example here of issues in which politics and statistics are having a disagreement, is the decennial census. For decades, an undercount in the census and its differential impact on minority populations has been well-documented. The Census Bureau has studied this issue for years and now has the statistical tools and methods to represent the uncounted individuals in the census totals. Yet this "adjustment" is opposed by many politicians because of an anticipated effect on the drawing of election district boundaries. However, the uses of the census extend far beyond apportionment and redistricting. The battle before the 2000 Census has been unusually intense.

---

1. Barbara A. Bailar, National Opinion Research Center, 1155 East 60th Street, Chicago, Illinois, U.S.A.

Given these instances in which politics and statistics are confronting each other, it is useful to step back in time to review the contributions of Mahalanobis to the government of India. His methods were used successfully by the U.S. Census Bureau to learn much of what we know about errors in the census. I will review Mahalanobis' contributions, then return to a discussion of the census, the statistical tools currently used in the census, additional tools that could be used, and then conclude with a plea for Congress and the Census Bureau to follow the tradition of continuous improvement in the census through the use of statistical tools.

2. The Mahalanobis Legacy

Mahalanobis played an important role in the methodology we take for granted today. He was trained to teach physics, but became increasingly interested in statistical problems and then in building the Indian Statistical Institute. His work on the utilization of interpenetrated subsamples of the population was innovative, and gave great impetus to research on the effects of interviewers on survey and census statistics. He paid great attention to the need for pilot studies to test the implementation of survey techniques. As time went along, he enlarged his interests from sampling and surveys, in which he provided much needed information to the government, to planning and economic development. He was appointed Honorary Statistical Advisor to the Cabinet in January, 1949 and placed in charge of the Central Statistical Unit in the same year. The central role of statistics in government planning was, no doubt, due to the force of the man himself as well as his research findings. He saw the role of statistics as a system to serve the cause of planned development and envisioned a feedback arrangement between statistics and planning (Rudra 1996).

The particular contributions I wish to stress today are his major roles in sample surveys and in measuring error of all kinds – errors of observation, errors of measurement, sampling errors, copying errors, printing errors. Much of his early work on showing the variability in statistics caused by interviewers was in crop statistics (Mahalanobis 1950). He was one of the first to say, and then show, that the overall error in survey statistics was not just sampling variance but also the variance arising from the human element. One way to study such errors was by the use of interpenetrated subsamples. In the words of Mahalanobis,

> "When two (or more) samples are drawn from the same population and covered according to the same survey design, the results based on the different samples are equally valid, even though they are derived by different operational units; and divergences between the different sets of estimates supply directly some idea of the margin of uncertainty." (Mahalanobis and Lahiri 1961)

Mahalanobis demonstrated that statistics based on samples were at least comparable to, and often more accurate than statistics based on a census, in the 1940's, when sampling was still not fully accepted. He believed, as many of us now do, that samples can be better controlled than can a census. He stated (Mahalanobis and Lahiri 1961) that the magnitude of discrepancies found in a census of jute production made it appear that a census may not provide accurate estimates for small areas. The random component of the non-sampling error may add enough error that results for a large area may be no different from those obtained by a sample survey. What holds for a large area does not naturally follow for small areas.

The U.S. Census Bureau used Mahalanobis' techniques to learn more about the underlying variability of census numbers.

3. What do people think a census is

To most people, taking a census means that enumerators go out and count everyone. There are three things that people seem to think about censuses. One is that everyone is counted. A second is that an enumerator sees everyone. A third is that the census is without error. Let's look at these one by one.

Often, everyone is not supposed to be counted in a national census, and who should be counted varies from country to country, and over time within a country. For example, military personnel and their families located outside of the country could be counted or not. Civilian aliens temporarily in the country as seasonal workers could be counted or not. From these illustrations one can see that a primary necessity in census-taking is defining the scope of the census.

So, by definition, certain groups of people are not to be counted in the census. This is by design of the Census Bureau. Other people make individual or family decisions not to be counted in the census. In earlier times, some families did not report children who suffered from some diseases or retardation. Some people who have had unfortunate episodes with the legal system may decide not to be counted. These may be people who are in the country illegally, those who are hiding from law enforcement, and those who fear, for whatever reason, the consequences of being counted. In 1990, there were people who said they would not be counted because they thought the census was too intrusive.

Finally, there are people missed, not by design but by accident. Perhaps they lived in buildings that were missed, perhaps they lived on the street and were missed. Perhaps they were away during the census period. During 1998 there were many reports of how much harder it was to survey people who live in gated communities. It may be that some of these people are missed because of the overzealousness of the community guards. In some communities good maps are unavailable or not updated, so groups of people may be missed.

In any case, not everyone is counted in a census and never was.

The second myth to be refuted is that an enumerator sees everyone and knows who should be in the census or not. This never happened, even in the early censuses in the U.S., when U.S. Marshals took the census and the country was much smaller. In fact, early censuses were of households, not of individuals. This means that there were no questions asked of individuals but instead there was interest in how many people were in the household, how many were men and how many were women, how many were in different age groups, and so forth. The totals were posted in public places. Starting in 1880 the canvasser method of taking a census, where enumerators went from door to door, came into being. It is this kind of census that made some believe that an enumerator saw everyone. However, a single household member usually responded for the whole family. The enumerator did not see those who were sick, those at work, those who were away temporarily, or those who were, for some reason or another, not in the room when the enumerator visited.

Though the enumerator-type census was an improvement over one taken by the marshals, research using interpenetrated subsamples showed that census enumerators still added a considerable amount of error to the census statistics. The enumerators were influenced by their own expectations and by responses of others in their enumeration district. Also, some did not understand the instructions and reported things incorrectly. An experiment in the 1950 census showed that enumerators added considerable variability to the census statistics (Hanson and Marks 1958). Indeed, the statistics gathered from a census had the same level of variability, due to enumerators, as a 25-percent sample. This is the main reason the Census Bureau turned to the use of self enumeration in the 1960 census and progressively expanded it in later censuses. Now, if a household receives the census form by mail, fills it out, and sends it in, and no errors require resolution, no enumerator will call at the household.

The third myth is that census taking occurs without error. No one who now works on censuses or surveys believes that, but other people do. The Census Bureau encourages that belief by publishing data down to the last digit. For example, the population of the United States in 1990 was reported and published as 248,718,301 in the *Statistical Abstract*.

Even some of those who have worked closely with a census cannot see it as a statistical process that carries with it a certain amount of error. Because the error is not routinely quantified and published along with the census numbers, some cannot believe the error exists. Some persons working in the Population Division of the U.S. Census Bureau in the 1940's and 50's believed that the census was the best way to learn about any subject, and that sample surveys were inferior. Repeated demonstrations of accuracy in survey results and of bias in census data did not change their minds.

Anyone who comes into regular contact with the census now knows that there is error in the data. First, though sampling variance cannot occur for items collected on a 100-percent basis, there may still be substantial response variance introduced by effects of enumerators, respondents, and coders on census data. Second, bias affects responses to many census questions even when a person is correctly counted. Bias also affects counts when enumerators do not count everyone. The Census Bureau conducts an evaluation program as part of every census, documents the amount of error, and uses those data to attempt to improve the next census.

Large groups of people are affected by census error. The undercounting bias affects minority populations and children at a much higher rate than other populations (Edmonston and Schultze 1993). Thus, communities that are largely African-American, Hispanic, or American Indian are underrepresented in distributions of potential power and money, while those statistics that are based on children under 10 are subject to a large error.

Over the years, the Census Bureau has reported numerous studies looking at the balance between cost and accuracy. One mentioned before is the use of self-enumeration. At smaller levels of population, the effect of response variance, primarily caused by interviewers, was very high. Just as with sampling error, as the size of the area increased, and the number of enumerators who collected the data increased, the effect lessened. When the mail return rate was close to 80 percent, the response variance decreased to about one-quarter of that of a 25-percent sample (Bailar 1969).

Thus, commonly held images of the census are not always true. Also, the census is not always the same. The Census Bureau has made many changes in census taking since the first census in 1790. The number of questions, the kinds of questions, who is counted and where, who does the counting, how people are assigned to a geographic domain, how missing characteristics are handled, and the gradual increase of asking most questions of a sample have changed over the years. The next section shows how the use of statistical tools has changed the census in this century.

4. Development of statistical tools in a census

Two elements have changed the methods of the U.S. Decennial Census considerably since 1940: the use of computers; and the use of statistical techniques. At times, the two elements have complemented each other, for example in the fast processing for imputation of missing data using a "hot deck" procedure. While computers have profoundly affected the census, the remainder of this discussion will focus on the statistical methodology.

One of the major advances starting in 1940 has been the use of sampling in the census. In 1940, as documented by Waksberg and Hanson (1965), there were three major uses of sampling. One was for the collection of data deemed supplementary to the main census questions. Questions such as mother tongue, veteran status, and fertility were asked of a 5-percent sample. A second use was for certain analytic studies requiring clerical transcription and coding. To avoid a long timespan for the transcription and coding to take place, a sample of census questionnaires was selected and the transcription and coding occurred only for them. A third use was for the verification of large-scale clerical operations such as editing, coding, key-punching, and so forth. Prior to 1940, all verification was on a 100-percent basis.

To describe the next leap forward, Waksberg and Hanson said:

> "A major step forward in the use of sampling in census work took place in the 1950 Census of Population and Housing. This grew out of a profound change in attitude regarding the role of sampling. Whereas in 1940 sampling had been considered applicable only for items of supplementary and secondary interest, in 1950 the entire range of census activities was examined to determine, on a logical basis, where complete counts were necessary and where samples could provide adequate information."

The increased use of sampling for population characteristics, for sample tabulations, and for verification was successful and evaluations showed that, even with the addition of sampling error, overall error was less than if earlier techniques had been used with no sampling. This was a reinforcement of the lesson learned earlier by Mahalanobis.

During the 1950 Census, the Bureau did a great deal of research to learn the effect of response biases and response variances on census data. Waksberg and Hanson declared that it was misleading to assume that the census, without sampling, was without error. In 1950, an experiment was conducted to estimate the effect of census enumerators on census data. By using the method of interpenetrated subsamples introduced by Mahalanobis, pairs of adjacent census areas were merged and assignments to the enumerators were randomized. Since the assignments were over the same area, differences between enumerators did not reflect differences in the type of area. The main finding of the study was that a full census in which enumerators went door to door to collect the census information had response variability that made the census the equivalent of a 25-percent sample (Hanson and Marks 1958). Using that result, as well as studies of biases in various census items, Waksberg and Hanson formulated a model in which census results were subject to a relative response bias of 6 percent and a response variance equal to the sampling variance of a 25-percent household sample. They used this model to generate Table 1 which shows the magnitude of total error in census data with and without sampling.

The authors point out that for a characteristic describing 500 individuals in an area of 2,500 people, the increase in the total root mean square error arising from sampling variability is only about 25%. For larger areas and larger cells, the additional error due to sampling is even smaller.

These data were studied carefully before the decision to increase the use of sampling in the 1960 Census. In practice, sampling made even greater gains than those anticipated by the model. The authors state "Thus for a great many published statistics, the reliability was better with the use of sampling than would have been possible otherwise." (Waksberg and Hanson 1965.)

Table 1

Expected Root Mean Square Error (RMSE) of estimated cell frequencies for individual items based on a complete census and on a 25-percent sample of households

Area of 2,500 Population having RMSE based on			Area of 10,000 Population having RMSE based on			Area of 50,000 Population having RMSE based on		
Cell Frequency	Complete Census	25-percent Sample	Cell Frequency	Complete Census	25-percent Sample	Cell Frequency	Complete Census	25-percent Sample
12	7	10	50	1	20	250	34	46
50	14	19	200	30	40	1,000	85	105
125	22	31	500	52	67	2,500	180	200
500	49	62	2,000	140	160	10,000	620	650
1,250	89	102	5,000	320	330	25,000	1,520	1,530

Note 1: Computations assume a relative response bias of 6 percent and response variance equal to the sampling variance for a 25-percent sample.

Note 2: The accuracy of the results (cell frequencies) is measured by a certain kind of average of the actual errors that would occur, the root mean square error (RMSE). A useful working rule would be to assume that approximately two-thirds of all results of a census or a sample would differ from their true cell frequencies by no more than their RMSE's.

A large-scale evaluation and research program of the Decennial Census program began in the 1950's and is now an integral part of the Census. Part of the program tests new methods for possible use in the following census and part of it focuses on the evaluation of the current census. It was as part of this program that the Bureau started measuring the undercount in the census. It was also this program in which the response variance due to enumerators was measured before and after the advent of self-enumeration. (In 1960, after self-enumeration was introduced, the response variance decreased to 1/4 of the 1950 level.) Since mail-back rates have decreased substantially since 1980, that variance may have increased again, perhaps substantially.

Other studies included research on alternative ways to measure the undercount, record checks to measure the accuracy of census data, and a study of using the Post Office not only to deliver census questionnaires but to notify the Census about missed addresses and duplicate forms.

Sampling is now used extensively to control the quality of the large-scale clerical tasks associated with the census. In past censuses, verification was usually dependent, in which the verifier reviewed the coder's work and determined whether the correct codes had been assigned. The Bureau planted errors and found that dependent verification missed as many as half of the errors. This and other research caused the Bureau to develop independent verification, in which records are assigned to three coders who do not see each other's work. A "majority rule" is used to determine the best code, and statistics about such errors are used to improve the process and to identify substandard performance.

Imputation was also a necessary tool developed for use in the census. To keep within time and budget parameters, the Bureau developed a "hot-deck" imputation system, based on the assumption that people who live in proximity are likely to resemble each other for many characteristics such as educational attainment and income. Another kind of imputation was also used in 1970, 1980, and 1990 to deal with a small, residual set of addresses left on the mailing list with no information about whether or not they were occupied. No one answered the door, nor did neighbors know if anyone lived there. Thus, based on a model that assumed a high correlation between the characteristics of neighboring households, the Bureau imputed occupancy or vacancy status, and to those imputed as occupied, a number of people were imputed. In 1980, only 762,000 persons were imputed, about .003 of the total census count, but they were not spread evenly over all the States. As a result of the imputation, Indiana lost a Congressional seat to Florida. However, it should be acknowledged that doing nothing about the unclassified units would have been equivalent to imputing them all as vacant. There was information available that showed that over half of these units could be expected to be occupied so the data based on imputation were more accurate than data based on counts alone with no imputation.

5. Additional Uses of Statistical Tools

Statistical tools can be used to correct the census for the undercount. The Waksberg-Hanson root mean-square error model estimates the amount of error in the census assuming a relative response bias in the overall census of 2 percent. (The 1990 estimate was 1.6 percent.) Also assume a response variance in both the adjusted and unadjusted census equal to one-fourth the sampling variance of a 25-percent sample. That estimate may now be too low since decreasing mail-back rates have driven enumerator variances higher. However, to be on the conservative side, we shall use the 1960 and 1970 measurements.

The model is the simple mean-square error model used frequently by the Census Bureau.

$$\mathrm{MSE}(T) = \mathrm{Var}(T) + B_T^2$$

Assume T is a cell size or a size of interest in the census in an area where N is the population size. $T = NP$ where P is the proportion of the population having a certain characteristic. B is the bias in the census count. So, for example, in an area of 2,500 people, one might be interested in knowing the number of children under 10 years of age. $N = 2{,}500$ and $T = NP$.

Now the variance of an estimated proportion, p is:

$$V(p) = \frac{N-n}{N-1} \cdot \frac{1}{n} \cdot PQ$$

If we have a 25 percent sample, this reduces to

$$V(p) = \frac{3}{4} \cdot \frac{1}{n} \cdot PQ = \frac{3PQ}{N}$$

$$\begin{aligned} V(T) = V(Np) = N^2 V(p) &= 3NPQ \\ &= 3TQ \end{aligned}$$

Relative bias = (0.02) so Bias = $0.02\,T$

Now we are dealing with a census, so there is no sampling variance, but the response variance is equal to 1/4 of what the sampling variance would be. So

$$\mathrm{MSE}(T) = (0.02T)^2 + (0.25)(3)TQ$$

and

$$\mathrm{RMSE}(T) = \sqrt{(0.02T)^2 + (0.25)(3)TQ}$$

This formula has been used as the basis of the calculations in Table 2. For an unadjusted census, $\mathrm{RMSE}(T)$ would have both the bias and variance components. For an adjusted census, the relative bias is zero, so only the response variance term remains. However, this analysis presumes that the adjustment factors themselves are free from any kind of variance and bias, and that the same adjustment factors can be uniformly applied within the demographic groups.

For example, Table 2 shows that for a total of 500 in an area of 2,500, the RMSE for an unadjusted census is 20 while the RMSE for an adjusted census is 17. For the unadjusted census, the contribution from the bias term is small, $[(0.02)(500)]^2 = 100$. The contribution from the response variance is $(0.25)(3)(500)(0.8) = 300$. So $\text{RMSE} = \sqrt{400} = 20$. For the adjusted census, the bias term, 100 is removed, so the $\text{RMSE} = \sqrt{300} = 17$. However, if one considers that the estimated bias term has both variance and bias, there may be little difference between the adjusted and unadjusted results for a small area. As the total, T, gets larger, the bias term is more dominant, and the adjustment removes more error.

Table 2 shows that for a small area of 2,500 persons there is no gain for small totals, but a gain of 43 percent in accuracy for a large total of 1,500 persons. In a somewhat larger area of 10,000 persons, there is little reduction in error until a total of 1,000 is of interest, where there is a gain in accuracy of 21 percent and for a large total of 5,000, there is a gain of 61 percent. Thus, if we were talking about the number of men or women in an area of 10,000, a total that might be expected to be around half the population, there would be a large gain in the accuracy of the total from using adjusted census figures. For an area of 50,000 the bias term dominates the mean square error, even at smaller totals such as 1,000. Here the gain is 21 percent, which grows to 81 percent for a very large total of 25,000.

This illustration shows is that adjusting the census does not add to the error of the census, even for small areas and small cells, if one assumes that the bias term is measured without error. For smaller area sizes and smaller cells, the response variance dominates the mean square error, but the total error is never less than the response variance. When the census is adjusted, the bias term goes to zero, and the gains in accuracy are dramatic.

One virtue of this model is that it was developed by the Census Bureau long before the current debate on adjustment grew heated. It was used to disabuse people of the idea that the census cells have no error. It was used successfully to show critics that having most of the census questions answered by only a sample would not hurt the data unduly. Such a tried and true census model now shows the real value of adjustment.

Table 2 used the relative response bias of 2 percent based on the 1990 Census overall estimate of the undercount of 1.6 percent. However, since the undercount hits minority populations harder, let's look at a comparison of an adjusted and unadjusted census in which the relative bias is 4 percent. (The 1990 estimates of the undercount were 4.4 percent for African-Americans, 4.5 percent for American Indians, 5.0 percent for Hispanics, and 2.3 percent for Asians.)

Table 3 shows the RMSE for minority communities for the sizes 2,500, 10,000 and 50,000. Though the RMSE's for the adjusted census stay the same, since the bias has been removed, the unadjusted RMSE's are considerably larger. The gains in accuracy from an adjustment are much larger in minority communities, as one would expect. For example, as shown above, the error in the number of males in a non-minority community of 10,000 would be about 109 unadjusted and 43 adjusted. In a minority community, the errors are 205 and 43 respectively. In a larger area of 50,000 the improvement is dramatic even for a small cell of 1,000.

Now, suppose we repeal the 1976 law that specifies that there shall be no sampling for the apportionment numbers. Think about a census in which, after a certain date, the housing units not returning census forms are sampled.

Table 2
Expected Root Mean Square Error (RMSE) of estimated cell frequencies for population estimates based on a census with no adjustment for undercount and with adjustment

Area of 2,500 Population having RMSE based on			Area of 10,000 Population having RMSE based on			Area of 50,000 Population having RMSE based on		
Cell Frequency	Unadjusted Census	Adjusted Census	Cell Frequency	Unadjusted Census	Adjusted Census	Cell Frequency	Unadjusted Census	Adjusted Census
15	3	3	50	6	6	250	15	14
50	6	6	100	9	9	500	22	19
100	9	8	200	13	12	1,000	34	27
500	20	17	500	21	19	2,500	65	42
750	25	20	1,000	33	26	5,000	116	58
1,000	29	21	2,000	53	35	10,000	214	77
1,500	37	21	5,000	109	43	25,000	509	97

Note: Computations assume a relative response bias of 2 percent in the unadjusted census and 0 percent in the adjusted census. There is a response variance in both the adjusted and unadjusted census equal to 1/4 the sampling variance of a 25 percent sample.

Table 3

Expected Root Mean Square Error (RMSE) of estimated cell frequencies for population estimates in African-American, American Indian, and Hispanic communities based on a census with no adjustment for undercount and with adjustment

Area of 2,500 Population having RMSE based on			Area of 10,000 Population having RMSE based on			Area of 50,000 Population having RMSE based on		
Cell Frequency	Unadjusted Census	Adjusted Census	Cell Frequency	Unadjusted Census	Adjusted Census	Cell Frequency	Unadjusted Census	Adjusted Census
15	3	3	50	6	6	250	17	14
50	6	6	100	9	9	500	28	19
100	9	8	200	15	12	1,000	48	27
500	26	17	500	21	19	2,500	109	42
750	31	20	1,000	48	26	5,000	208	58
1,000	45	21	2,000	87	35	10,000	407	77
1,500	64	21	5,000	205	43	25,000	1,004	97

Note: Computations assume a relative response bias of 4 percent in the adjusted census and 0 percent in the adjusted census. The response variance in both the adjusted and unadjusted census equal to 1/4 the sampling variance of a 25 percent sample.

In this model, there are two components of variance, the response variance and the sampling variance. The sampling variance is based only on the nonresponse universe.

Let R be the nonresponse rate, and M the population of nonresponse households. Then $M = RN$. The total for which we are trying to estimate the sampling variance is $S = PM$. The relationship between S, the sampled part of the total, and T, the total, is through R. $S = PM = P(RN) = RT$.

So the sampling variance $= 3MPQ = 3PQRN$, assuming a 25-percent sample of the nonrespondents. This sampling rate could easily be changed for a larger rate, but for purposes of illustration, it suffices.

In Table 4, there are three contributors to the RMSE. Two of them are the terms we saw in the earlier description when sampling of the non-mail returns was not a consideration. Now we have a third term, expressing the sampling variance arising from the sample of non-mail returns. In an adjustment, only the bias term goes to zero, while the two variance terms remain. Each of the variance terms gets smaller as the cell size gets larger, but they do not vanish.

Table 4 shows the RMSE's for a census with no sampling of non-mail return households, with and without adjustment, for a 25 percent sample of non-mail return households when only half of the population mails them back and when 70 percent mail them back for the three sizes of area we have looked at before: 2,500 population, 10,000 population, and 50,000. The no sampling case is what we will have in the 2000 Census because the use of sampling for follow-up is prohibited. Look first at Section A for a population of 2,500. Where there is no sampling of non-mail return households, we see the numbers from Table 2. When half of the population mails back the census form, and the remaining half is sampled, the variance component keeps the adjusted and unadjusted RMSE's very close together. At maximum, there is a 20 percent reduction in error. There is somewhat more gain when the mailback rate is .70 and only 30 percent of the remaining population is sampled. The maximum gain in this case is 28 percent.

Small areas, such as those of 2,500 may be greatly affected by sampling, especially at a 25-percent rate if the mailback rate is low. Whether a decrease in accuracy is acceptable depends on the uses for the data. Since providing small area data is an important objective of the census, it may be that there would need to be a much larger sampling rate, if not complete follow-up for small areas. The Census Bureau has done this before with some characteristics, such as income, so that there would be less variability in the income data for areas of 2,500 or fewer persons. Following that same principle, it could be specified that there would be no sample follow-up in places of 2,500 persons or fewer, and variable follow-up rates depending on place size. Another strategy would be to use the information abundantly available about coverage error and to specify larger samples in places that have characteristics highly correlated with the undercount.

For areas of 10,000 population, we see a definite improvement from the adjustment for the bias, but the adjusted numbers with sampling are still considerably larger than the adjusted figures without sampling. However, if there is no adjustment, the sampling adds to the RMSE, but the unadjusted numbers are not much different. There is a 15 percent increase in the RMSE when only half the population returns the census form and an increase of 9 percent when 70 percent return it.

Finally, when we look at an area of 50,000 we see that the bias dominates the RMSE for all but the smallest cell sizes. When the total we are trying to estimate is 5,000 or larger, sampling adds to the RMSE, but an adjustment, with sampling, is still superior to unadjusted numbers with no sampling.

Table 5 is similar, but geared to a predominantly minority population. As in Table 3, the relative bias is 4 percent, reflecting an average undercount rate for minorities. In this table, the RMSE's for unadjusted totals are much more similar, even for smaller areas, because of the larger effect of the bias term on the RMSE.

The results for areas of 50,000, which exist in most large cities, show the devastating effects of not adjusting for the large minority undercount. The sampling variance for the larger totals has practically no effect on the RMSE, but the improvement from adjustment for all cases, sampling or no sampling is 83 percent or higher. The added error because of sampling is negligible.

Unfortunately, in many minority communities, low mail-return rates and undercounting occur together. Such communities have a 50 percent mail return rate or lower. It may be that the sample size will need to be increased in these areas.

Table 4

Expected Root Mean Square Error (RMSE) of estimated cell frequencies for population estimates based on an unadjusted census, an adjusted census, and on a 25 percent sample of non-mail return households

A. Area of 2,500 population having RMSE based on

	No sampling of non-mail return HH's		25% sample and 0.50 mailback rate		25% sample, and 0.70 mailback rate	
Cell Frequency	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
15	3	3	6	6	5	5
50	6	6	11	11	9	9
100	9	8	15	15	13	13
500	20	17	32	30	28	26
750	25	20	38	34	33	29
1,000	29	21	42	37	37	31
1,500	37	21	47	37	43	31

B. Area of 10,000 population having RMSE based on

	No sampling of non-mail return HH's		25% sample and 0.50 mailback rate		25% sample, and 0.70 mailback rate	
Cell Frequency	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
50	6	6	11	11	9	9
100	9	9	15	15	13	13
200	13	12	21	21	18	18
500	21	19	34	33	30	28
1,000	33	26	49	45	43	39
2,000	53	35	72	60	65	51
5,000	109	43	125	75	119	64

C. Area of 50,000 population having RMSE based on

	No sampling of non-mail return HH's		25% sample and 0.50 mailback rate		25% sample, and 0.70 mailback rate	
Cell Frequency	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
250	15	14	24	24	21	20
500	22	19	35	33	30	29
1,000	34	27	51	47	45	40
2,500	65	42	89	73	80	63
5,000	116	58	142	101	132	86
10,000	214	77	241	134	231	115
25,000	509	97	527	168	520	144

Table 5

Expected Root Mean Square Error (RMSE) of estimated cell frequencies for population estimates in African-American, American Indian, and Hispanic communities based on an unadjusted census, an adjusted census, and on a 25 percent sample of non-mail return households

A. Area of 2,500 population having RMSE based on

	No sampling of non-mail return HH's		25% sample, and 0.50 mailback rate		25% sample, and 0.70 mailback rate	
Cell Frequency	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
15	3	3	6	6	5	5
50	6	6	11	11	9	9
100	9	8	15	15	13	13
500	26	17	36	30	33	26
750	31	20	46	34	42	29
1,000	45	21	54	37	51	31
1,500	64	21	70	37	68	31

B. Area of 10,000 population having RMSE based on

	No sampling of non-mail return HH's		25% sample, and 0.50 mailback rate		25% sample, and 0.70 mailback rate	
Cell Frequency	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
50	6	6	11	11	9	9
100	9	9	15	15	13	13
200	15	12	22	21	18	18
500	21	19	38	33	34	28
1,000	48	26	60	45	56	39
2,000	87	35	100	60	95	51
5,000	205	43	214	75	210	64

C. Area of 50,000 population having RMSE based on

	No sampling of non-mail return HH's		25% sample, and 0.50 mailback rate		25% sample, and 0.70 mailback rate	
Cell Frequency	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
250	17	14	26	23	23	21
500	28	19	39	33	35	29
1,000	48	27	62	47	57	40
2,500	109	42	124	73	118	63
5,000	208	58	224	101	218	86
10,000	407	77	422	134	416	115
25,000	1,004	97	1,014	168	1,010	144

6. Conclusion

It has been a tradition for the Census Bureau in the latter half of this century to use statistical techniques, where possible, to make the Decennial Census more accurate and less costly. Using the techniques historically used by the Census Bureau, namely a mean-square error model, one can see that adjustment does improve census totals, even for small areas, when one assumes even a minimal level of response variance. One can also see the need for precaution if sampling is to be used for follow-up. It may be that there should be no sampling in places of 2,500 or fewer people, just as there is no sampling for certain population characteristics in these small places.

In looking at the current census controversy, it is good to remember the spirit of Mahalanobis. Not only did his ingenious use of interpenetrated subsamples give us the ability to estimate the response variance in census statistics, but his insistence that sampling and statistics should be used to solve practical problems has been the hallmark of the U.S. Census. Some of the most fundamental practical problems are those faced by the government and Mahalanobis allocated statistical resources for the solving of these problems. Likewise, the U.S. Census Bureau has a long and rich history of offering practical, cost-efficient solutions to thorny census problems.

References

Bailar, B.A. (1969). Evaluation and Research Program of the U.S. Censuses of Population and Housing, 1960: The Effect of Interviewers and Crew Leaders. Series ER 60 No. 7. Washington, DC: U.S. Bureau of the Census.

Edmonston, B., and Schultze, C. (1993). *Modernizing the U.S. Census*. Washington, DC: National Academy Press, 34-35.

Hanson, R.H., and Marks, E.S. (1958). The influence of the interviewer on the accuracy of survey results. *Journal of the* American Statistical Association, 53, 639-655.

Mahalanobis, P.C. (1946). Recent experiments in statistical sampling in the Indian Statistical Institute. *Journal of the Royal Statistical Society*, 325-378.

Mahalanobis, P.C. (1950). Why Statistics? *Sankhyā*, 10, 195-228.

Mahalanobis, P.C., and Lahiri, D.B. (1961). Analysis of errors in censuses and surveys with special reference to experience in India. *Bulletin of the International Statistical Institute*, 38, 2, 401-433 (reprinted in *Sankhyā*, 23, 325-358).

Michael, R.T., Gagnon, J.H., Laumann, E.O. and Kolata, G. (1994). *Sex in America*, *A Definitive Survey*. New York: Little Brown and Co.

Rudra, A. (1996). *Prasanta Chandra Mahalanobis*: *A Biography*. New York: Oxford University Press.

Waksberg, J., and Hanson, R. (1965). Sampling Applications in Censuses of Population and Housing. U.S. Bureau of the Census, Technical Paper No. 13.