Dr. David W. Murray
Census Monitoring Board
Congressionally Appointed Member
Hearing on Year 2000 Decennial Census
Committee on Commerce, Science and Transportation
United States Senate
March 28, 2001

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

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

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

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

In the resultant morality play, adjustment advocates usually came off as earnest advocates for the poor, who could be aided by a simple application of statistical justice. Those who favored an enumerated count, on the other hand, were

often cast as stubbornly refusing to use a readily available technical means to solve a social problem – "correcting" the undercount by statistics. Lost in the fracas were genuine arguments about the feasibility and advisability of supplanting the standard enumeration with these technical means – a position ultimately validated not only by the Supreme Court decision of January, 1999, but as well on February 28, 2001 by the decision of the Census Bureau itself. The enumerated count prevailed for good technical, not political, reasons.

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

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

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

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

system and make good estimates about the real population of the lake.

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

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

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

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

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

But let's put all those concerns aside for a moment, and presume that all went well in the measurement process. Have we really taken a sample with the A.C.E. that can then be used to correct the count? The answer is still, um, not exactly. There are several remaining steps to go. There is first the problem of selecting the 314,000 A.C.E. households. Every pollster knows that a proper sample, which will be used to extrapolate opinions "upward," as it were, onto the entire population (not, as has been noted, what the A.C.E. actually does), must guard against being skewed or biased in its selection of participants. The best guard is to have a systematically random probability process for the selection. But this can not be done exactly for the A.C.E. design, since there are constraints on the sample that result in trade-offs. For instance, there must be some households allocated to every one of the 50 states, and further, we must ensure that the households are "distributed" in such a way that they represent demographic groups of interest.

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

But the most important step was to divide up the 314,000 households into what were termed "post-strata." That is, the population in the sample was stratified and assigned into multiple cells that, according to the A.C.E.'s sociology, represented appropriate "types" of Americans. The appropriateness of a type was related to the probability that an individual in one of those cells would be missed in the enumerated census. The post-strata (think of the cells in an Excel spreadsheet) represented the intersection of variables like race/ethnicity, sex, age, and tenure (homeownership or not), the whole apparatus further divided by four regions of the country and by type of community (a range of larger-to-smaller metropolitan areas continuing to rural). As it turns out, only the largest post-stratified type, non-Hispanic Whites, were subdivided into the full set of post-strata distinctions. For most other types, the cells had to be conflated because of small

absolute numbers, meaning that Asian Americans, for instance, were placed into two national cells (owner

and non-owner) without regional breakdowns.

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

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

Further, all of these purely quantitative concerns must be coupled with the apparent arbitrariness and uncertainty about the sociological assumptions underlying the choice of American "types." Were the assumptions actually legitimate models of the probability of being enumerated? We simply don't know. Overall, we must realize that the census represents the intersection of sophisticated quantification (assumptions about numbers) with real human beings (assumptions about which are, unfortunately, anything but a sophisticated science). The A.C.E. design represents the place where two sets of very complicated models of the world derived from two very different disciplines interact, with any errors (in theory or in implementation) compounding each other. The results are then magnified by becoming the basis for adjusting the data on 281 million other people, the A.C.E. being considered the last word in accuracy, and hence, the benchmark standard for calibrating the entire US data collection system. Suffice it to say, the stakes are high for such a probability mechanism of unproven reliability.

And all that has been discussed above transpires before the results are released to the public, and ultimately encounters the requirements of the legal and constitutional system, fundamental provisions of which contradict the A.C.E. activity on the face of it. Finally, in a development beyond the scope of this discussion, we must remember the indeterminacy added to this census by the first-ever multiple-race selection, cross-cutting the

whole system with 126 possible choices of racial/ethnic self-identification (which choices have themselves been acknowledged to be completely arbitrary governmental categories with no basis in scientific fact; moreover, the choices are unstable even in single individuals at different times).

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

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

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

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

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

unduplication of records or imputations).

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

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

At any rate, we are now at a point in our analysis where we can adjust the census statistically to "correct" the count. If a post-stratum has a positive number, such as 1.08, that means that we found more people in the A.C.E. survey in that stratum than the enumeration had recorded. Rather than 100 people, let us say, the A.C.E. is telling us there likely are108 people. Hence, every time we find a record back in the enumeration national census for some other one in that post-stratum, we do not record just a one; we instead write down 1.08 for each one found. That means for every group of 100 people we find anywhere in the country who fit this profile, we "add" 8 more people of that type (where, exactly, do we put them? More in a minute...). Now we are rolling at last. For every 10,000 found, we write down 10,800. For every ten million, well, let us see here, the model tells us we have got 800,000 more people just like that, which we have to place somewhere on the map, even though we've never actually met nor counted them directly.

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

Even more remarkable, we have just engaged in a process that is not really "sampling" at all, but

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

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

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

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

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

In actuality, we begin to see many difficulties with the operation of the A.C.E. conceptually, over and above those concerns linked to problems of measurement error and implementation issues. One of the central conceptual difficulties is that the statistical estimation, incorporating, as it must, a certain probability margin of error that is ineradicable (it being inherent in the operation of probability), only begins to "even out" its errors at certain levels of aggregation. That is, for the gross level of the total population (the aggregate count of the total number of the population, roughly 281.4 million persons), the probability errors (the inevitable pluses and minuses wavering from the actual target) do "average out." For instance, for every

100,000 measurement "pluses" that are too high there will also occur about 100,000 corresponding measurement "minuses" that cancel each other out. But a lower levels of aggregation (state, county, district) the possibility starts to magnify that they do not all "average out," and we may well be left with residual error – a less accurate count than we began with in the enumeration.

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

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

Problems abound. By virtue of the adjustment design, we have generated estimated people (virtual people) who have never been contacted nor identified, yet must be placed in some concrete location in an actual census block. The principle for assigning them a "local habitation and a name" is arbitrary and based on unproven assumptions. Yet their presence can have consequence in the apportioning of political power and funding. Moreover, the Congressional-appointed members of the Board further demonstrated in our report of September, 1999 that the effect of adjustment is to fail to position the undercounted correctly and proportionally in the communities where they were actually missed. By applying a "blanket" adjustment to every sector across the country, the adjustment gives the illusion of a remedy, because the actual undercounted are not uniformly distributed across the country. Those communities that "lose" in the undercount do not receive a commensurate adjustment.

Further, it could be argued that the adjustment design, an effort to statistically "model" the population and then reformulate it, could have the effect of introducing more political features into the census than are found in the actual enumeration. Let it be noted that the "political" aspect of the census adjustment does not have to necessarily imply the active intervention of partisan concerns. As with budgetary or income tax battles, any process that is "assumption-dependent" is thereby open for political debate. Whoever sets the assumptions, or establishes the criteria for which factors are considered important (and in which order), can largely constrain the possible outcomes of the strictly quantitative process. All census activities are, of

course, assumption-dependent, in this sense – witness the dispute between Utah and North Carolina over the allocation of the last House seat based on population. The issue hinged on whether or not overseas missionaries were assumed to be equivalent to overseas military in terms of their state assignment. It follows that every aspect of the census has political implications, in that it constructs political definitions and quantifies what are properly political entities – human beings in groups. Nevertheless, even given these caveats, census enumeration is relatively more assumption-independent than is the alternative -- modeling the population for statistical adjustment purposes, where changed assumptions have the power to radically alter the entire nature of our national self-portrait.

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

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

We face a genuine dilemma, which the ESCAP report issued at the end of February well captured. The actual enumerated count from Census 2000 placed us in a new and perplexing landscape, one unanticipated by the designers of the A.C.E. (and many other parties as well). The count of 281.4 million surpassed the best reckoning of the population provided by the Demographic Analysis (DA) by nearly 2 million people. This is an anomalous outcome, since the DA has traditionally stood above the enumerated count and told us the magnitude of the undercount in the enumeration. But now we have broken through the DA measurement. By the reckoning of the DA, we actually have an OVERCOUNT in the enumerated census. Perhaps, on the other hand, the enumerated census is right on the mark, having itself nearly eliminated the

heretofore undercount. Various scenarios to explain this have already been proposed by the ESCAP report and others, but none has achieved a reconciliation without introducing yet other anomalies.

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

with either of the two other measurements.

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

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

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

Thank You.