

Coverage of Population in Census 2000:

Results from Demographic Analysis

J. Gregory Robinson, Arjun Adlakha, and Kirsten K. West

U.S. Census Bureau, Washington, DC

Paper presented at the 2002 Annual Meeting of the Population Association of America, Atlanta, Georgia, May 8-11. This paper reports the results of research and analysis undertaken by Census Bureau staff. It has undergone a more limited review than official Census Bureau publications. This report is released to inform interested parties of research and to encourage discussion. The views expressed in this paper are those of the author(s) and do not necessarily represent the views of the U.S. Census Bureau.

Coverage of Population in Census 2000: Results from Demographic Analysis

By

J. Gregory Robinson, Arjun Adlakha, and Kirsten K. West
U.S. Census Bureau, Population Division
Washington, D.C. 20233

Demographic analysis (DA) is a well-developed coverage measurement and evaluation program. DA has served as the standard for measuring coverage trends in recent censuses and coverage differences by age, sex, and race at the national level. DA represents a macro-level approach, where population benchmarks are developed by analyzing and aggregating various types of demographic data. Examples include administrative statistics on births, deaths, legal immigration, and Medicare enrollments, as well as estimates of emigration and unauthorized immigration. The difference between the DA benchmark and the census count provides an estimate of the census net undercount. The demographic approach differs fundamentally from the survey-based Accuracy and Coverage Evaluation (A.C.E.).

In this paper, we present the results of the DA evaluation for Census 2000. The first section gives an overview of the DA method and describes its limitations. The second section addresses the initial DA results that were released in March of 2001 and compared to the A.C.E. results. The third section presents the “revised” DA results released in October of 2001, discusses the differentials in coverage measured by the DA estimates, and compares patterns of coverage in 2000 to patterns measured in the 1990 census. The last section offers conclusions and steps for future research.

The DA shows a substantial reduction in net undercount in Census 2000 compared to 1990. The reductions occur among all demographic categories: broad age groups, males and females, Black and NonBlacks. According to the DA estimates, the net census undercount rate in 2000 remains disproportionately higher than the overall average for two groups: Black men and young children.

Keywords: Demographic analysis, Census coverage, Net undercount

Introduction

In the United States, methods of Demographic Analysis (DA) have a long history of measuring population coverage trends between censuses and differences in coverage by age, sex, and race (Black, Non-Black) at the national level. Demographic Analysis is an analytic approach that has been used to measure coverage of the national population in every census since 1950 (see Siegel and Zelnik, 1966; U.S. Bureau of the Census, 1974, 1988; and Robinson et al., 1993a for the demographic evaluations of the 1960-1990 censuses). DA has documented the long-term reduction in the census net undercount rate, yet DA also reveals the persistent and disproportionate undercount of certain demographic groups (such as Black men). A goal of Census 2000 was to reduce these “differential” undercounts.

Demographic Analysis estimates serve two principal purposes in census evaluation:

1) DA estimates provide an alternative benchmark to assess completeness of coverage in the current census and document changes in coverage from previous censuses. The national DA estimates have become the accepted benchmark for tracking historical trends in net census undercounts and for assessing coverage differences by age, sex, and race. As in past censuses, DA estimates provide a new independent assessment of coverage in Census 2000 to add to the historical time series.

2) The independence and internal consistency of the DA estimation process allows us to check the survey-based Accuracy and Coverage Evaluation (A.C.E.) coverage estimates; in particular, we can assess the consistency of the age-sex results. DA and A.C.E. use entirely different methodologies. Because the sources and patterns of errors in the two estimates are sufficiently different, any disagreement in the results is important to understand.

This paper presents the results of the Demographic Analysis evaluation of coverage in Census 2000. The first section gives an overview of the demographic technique and describes the components of population changes used to construct the DA estimates. The second section addresses the initial DA results that were released in March of 2001 (U.S. Bureau of the Census, 2001a) and describes their limitations. The third section presents the “revised” DA results released in October of 2001 (U.S. Bureau of the Census, 2001b) and discusses the differentials in coverage measured by the DA estimates and compares patterns of coverage in 2000 to patterns measured in the 1990 census. The last section offers conclusions and steps for future research.

Description of the Demographic Analysis method

Demographic Analysis represents a macro-level approach for measuring coverage. The approach derives estimates of net undercount by comparing aggregate sets of data or counts. In general, DA population estimates are developed for the census date by analysis of various types of demographic data independent of the census, such as administrative statistics on births, deaths, and Medicare data and estimates of immigration and emigration. The difference between the DA estimated population (P)

and the census count (C) provides an estimate of the net census undercount (u). Dividing the net undercount by the DA benchmark provides an estimate of the net undercount rate (r):

$$u = P - C \quad (1)$$

$$r = (u/P) * 100 \quad (2)$$

The particular analytic procedure used to estimate coverage nationally for the various demographic subgroups depends primarily on the nature and availability of the required demographic data. Two principal demographic techniques are used to produce the demographic analysis estimates for 2000, one for the population under age 65 and another for the population 65 and over.

(1) Ages under 65. The Demographic Analysis estimates for the population below age 65 are based on the compilation of historical estimates of the components of population change: births since 1935 (B), deaths to persons born since 1935 (D), immigrants born since 1935 (I), and emigrants born since 1935 (E). Presuming that the components are accurately measured, the population estimates (P_{0-64}) are derived by the basic demographic accounting equation applied to each birth cohort:

$$P_{0-64} = B - D + I - E \quad (3)$$

The size of the component estimates used to develop the DA population under age 65 for 2000 is shown in Table 1 (revised estimates are given):

Table 1. DA Estimates of the Components of Change for the U.S. Resident Population:
April 1, 2000

Component	Estimate
Total Population	281,759,858
Under age 65 in 2000	
+ Births since 1935 (B)	234,860,298
- Deaths to persons born since 1935 (D)	14,766,736
+ Immigration to persons born since 1935 (I)	32,563,971
- Emigration to persons born since 1935 (E)	5,485,117
Ages 65 and over in 2000	
Medicare-based population	34,587,440

Clearly, births (234.9 million) represent by far the largest component in equation 3. The immigration component (32.6 million) is second largest, followed by deaths (14.8 million) and emigrants (5.5 million).

The actual calculations are carried out for single-year birth cohorts. For example, the estimate of the population age 40 on April 1, 2000 is based on births from April 1959 to March 1960 (adjusted for under-registration), reduced by deaths to the cohort in each

year between 1959 and 2000, and incremented by estimated immigration and emigration of the cohort over the 40-year period. As will be described, the immigration component is comprised of several subcomponents.

(2) Age 65 and Over. Administrative data on aggregate Medicare enrollments are used to estimate the population age 65 and over (P_{65+}):

$$P_{65+} = M + m \quad (4)$$

where M is the aggregate Medicare enrollment and m is the estimate of underenrollment in Medicare. The DA Population 65 and over is based on 2000 Medicare enrollments. Medicare is an administrative data set from the Health Care Financing Administration. Although Medicare enrollment is generally presumed to be quite complete, adjustments are made to the basic data to account for individuals who are omitted. An allowance is made for the estimated 1.3 million not enrolled (3.9%). Underenrollment factors are based on survey estimates of Medicare coverage and data on age at enrollment in the Medicare file. The DA population aged 65 and over (34.6 million) represents 12.3 percent of the total population. For a more complete description of the Medicare data and estimation for underenrollment, see Ahmed et al, 2001.

The demographic component estimates for the population under 65 are combined with the Medicare-based estimate for the population 65 and over to produce the total revised DA population estimate of 281.8 million as of April 1, 2000.

Estimation of Components

Estimating Births. The historical data on births, B in Equation (3), come from the vital registration system and are available at the national level beginning in 1933. As shown in Table 1, births are by far the largest component of population change involved in the demographic analysis system (234.9 million). Thus, even relatively small errors in the estimates of births can have significant effects on the demographic estimates of coverage. Tests of birth registration completeness conducted for 1940, 1950, and 1964-1968 provide correction factors for those years. The estimated level of completeness was 92.5 in 1940 (81.9 percent for Black births), 97.9 in 1950 (93.7 for Black births), and 99.2 in 1964-68. Correction factors for other years are derived by interpolation and extrapolation. In particular, the estimated number of Black births depends on the quality of the correction factors. An investigation and subsequent revision of the 1940 birth registration results led to a downward adjustment to the time series of Black births (1935-1945) and lowered the estimated net undercount for those Black cohorts (Passel, 1991; Robinson et al, 1990).

Recent research examined the assumptions underlying the completeness of birth component. The principal outcome of this research was a revision in the assumptions about registration completeness of births since 1968. The previous DA estimates assumed that all births in years since 1968 (the last year of the most recent test of birth registration completeness) were registered at the same percent (99.2 percent). Under the revision, we allow for registration completeness to gradually rise to 100 percent in 1985

(the first year natality statistics were reported electronically from all the states), and remain at 100 percent through 2000. For a more complete description of the evaluation of the birth statistics and adjustments for underregistration, see McDevitt et al, 2001 and U.S. Bureau of the Census, 1991a.

The consistency in classification of births by race is important for making comparisons between the Demographic Analysis estimates and the census. DA estimates for race groups are affected by the differences between the classification of births used in the registration system and the classifications used in the census. Race at birth is assigned on the basis of the race of the parents, and different algorithms can lead to different race assignments for births to mixed-race couples. While not affecting DA totals, this uncertainty affects DA race estimates principally for the cohorts born after 1980. Until 1989, the National Center for Health Statistics (NCHS) assigned rules for the classification of births by race, which favored race categories other than White. In 1989 NCHS adopted a new rule, which could be called the “mother rule,” which assigns the mother’s race to the child regardless of the father’s race. Alternatively, one can develop the “father rule” where the child is assigned the father’s race. Work by Passel (1990) and U.S. Bureau of the Census (1991a) shows that the “father rule” tended to match more closely with the response patterns in the census. Estimates of births consistent with the father rule were used to develop the population estimates in Equation (3).

Estimating Deaths. The historical data on deaths (D) are based on the vital registration system. An estimated 14.8 million deaths occurred between 1935 and 2000 to the population under age 65 in 2000 (Table 1). There is little information available on which to quantify empirically the possible extent of underregistration of deaths (U.S. Bureau of the Census 1991c). Therefore, unlike births for which the probable magnitude of underregistration can be empirically quantified, the magnitude of underregistration of deaths must be based on speculation. With the exception of infant deaths, all deaths are assumed to be completely registered. For infant deaths before 1960, it was assumed that deaths were underregistered at one-half the rate of underregistration of births; no adjustment for infant deaths was made for years since 1960. In addition to actual deaths, life table survival rates are used in the estimation of some immigration components (especially unauthorized immigration and legal emigration).

Estimating Immigration. The immigration component (I) in the demographic accounting equation has two major parts: measured migration and residual foreign-born migration. Measured migration consists of immigrants admitted for permanent residence, net migration from Puerto Rico, net arrival of civilians who had been living abroad, nonimmigrants living in the United States temporarily (including foreign students), and the movement of military personnel abroad.

Of these international migration components, legally admitted permanent residents and residual foreign-born migration are the largest. As shown in Table 2, the DA estimate for 2000 includes 20.3 million persons under age 65 who immigrated between 1935 and 2000 and 10.0 million other foreign-born persons (which includes unauthorized immigrants).

Table 2. DA Estimates of the Components of Immigration for the U.S. Resident Population under 65 Years of Age: April 1, 2000

Component	Estimate
Legally Admitted Permanent Residents	20,332,038
Other Measured Migration	2,249,001
Migrants from Puerto Rico	905,698
Temporary Migrants	776,002
Civilian Citizen Migration	891,940
Armed Forces Overseas	-324,639
Residual Foreign-Born Migration (includes unauthorized migrants)	9,982,932

International migration is an important component in the DA estimates. However, because administrative records for various components of international migration are incomplete or missing, assumptions about these components are particularly sensitive. Furthermore, estimates of certain components of international migration such as emigration and residual foreign-born migration, are based on census data and usually are not updated until sample data from the decennial census are processed. By using preliminary sample data from Census 2000, we were able to update these components of international migration.

Legally admitted permanent residents

Legally admitted migrants represent the largest of the legal migration components (20.3 million in Table 2). Data on legal permanent residents are based on administrative records from the Immigration and Naturalization Service (INS). The Immigration and Nationality Act defines legal immigration as the process by which a non-citizen of the United States is granted legal permanent residence. Legal immigrants, as categorized by the Census Bureau, include new arrivals to the United States and people adjusting their migrant status to legal permanent resident (including Special Agricultural Workers (SAWs) and pre-1982 entrants (LAWs)), asylees, and refugees.

The INS data are believed to be quite complete, are relatively timely (lags of one year), and require less estimation in comparison to other immigration components. However, given the wide variety of migrant situations, it is becoming more problematic to properly classify migrants into status categories (new arrivals, adjustees, refugees, asylees). Also, the race of immigrants has to be estimated, based on the race of recent immigrants by country of origin as reported in the most recent census. For a more complete description of the legal immigration data, see Perry et al., 2001.

Migrants from Puerto Rico

The migration of citizens to and from Puerto Rico (0.90 million) is a component which must be estimated and is thus subject to uncertainty. The net migration estimates for decades before 1990 are based on the cohort component analysis of successive census data of the resident population of Puerto Rico and the Puerto Rican-born population living in the U.S. Net migration from Puerto Rico for 1990 to 2000 is projected from estimates developed for the 1980's. We are exploring other data sets to improve the estimation of net migration from Puerto Rico. For a more complete description of the method and data to estimate migration from Puerto Rico, see Christenson, 2001.

Temporary Migrants

The Immigration and Nationality Act defines temporary migrants (also referred to as nonimmigrants) as aliens admitted to the United States for a specified purpose and temporary period, but not for permanent residence. Temporary migrants include those who would be considered residents of the United States for purposes of the decennial census, including foreign students and temporary workers, but excluding tourists and business workers. The number of temporary migrants used in the DA construction (0.78 million) was estimated using an algorithm to identify people with characteristics like those of temporary residents (e.g., persons who are foreign-born, not citizens, going to college, and have lived in the U.S. for less than 3 years). For a more complete description of the method and data to estimate temporary migrants, see Cassidy and Pearson, 2001.

Civilian Citizen Migration

The temporary movement abroad of citizens of the United States (0.89 million) is another component which relies wholly on estimation and is thus subject to considerable uncertainty. The temporary movement of Federally affiliated U.S. citizens is estimated by observing trends in the numbers from two sources, (1) the number of military dependents is published by the Department of Defense, and (2) the number of civilian Federal employees overseas is supplied by the Office of Personnel Management (OPM) and is inflated to include dependents.

Armed Forces Overseas

The Armed Forces overseas (AFO) component (-0.32 million) measures the net effect of three subcomponents: (1) the number of Armed Forces stationed overseas, based on data provided by the Department of Defense, (2) deaths to the AFO, and (3) net recruits to the Armed Forces from the overseas civilian population. We believe the AFO component is relatively accurate.

Residual Foreign-Born Migration

The component of residual foreign-born migration is based completely on estimation, and is subject to considerable uncertainty as to its true size. The number of residual foreign-born migrants in 2000 is estimated from analyses of data on the foreign-born population

obtained from the 2000 census. A cohort component, or “residual” methodology is utilized. An estimate of legally admitted permanent residents is developed based on administrative data on legal migrants and estimates of deaths and emigrants. The difference (the residual) from the Census 2000 count of the foreign-born (9.98 million in Table 2) is assumed to include unauthorized immigrants counted in the census. The estimation equation is as follows:

$$R = FB - [L - (M + E)] - T \quad (5)$$

Where FB = Census 2000 foreign-born population
 L = Legal immigrants
 M = Mortality to legal immigrants
 E = Emigration of legal immigrants
 T = Temporary (legal) migrants
 R = Residual foreign born

This residual foreign-born population is not an estimate of the number of unauthorized migrants. This estimate also includes persons who are here legally but are not yet included in the official estimates of immigrants and refugees. It also includes persons in a "quasi-legal" status who are awaiting action on their immigration requests. Since the estimate was derived from a residual methodology, any limitations in the methods or in the measurement of other migration components are reflected in the residual number. Here the residual may be quite different from the actual number of unauthorized migrants.

The estimate pertains only to persons counted in the census; other assumptions must be made about the undercount of residual foreign-born population; an assumed net undercount rate of 15 percent was used. Revisions to the estimates of residual foreign-born migrants may be made when final detailed sample data from the 2000 census are analyzed. For a more complete description of the methodology, see Costanzo et al, 2001.

Estimating Emigration

The component of emigration (-5.5 million in Table 1) is based completely on estimation, and is subject to uncertainty as to its true size. The emigration component (E) represents emigration of legal residents only. The estimation of foreign-born emigration for decades up to 1990 is based on using cohort component techniques applied to successive census data on the foreign-born. The volume of emigration for the 1990s is based on simple extrapolations of emigrant levels during the 1980s; the estimates are subject to greater error because no standard technique using current empirical data has yet been successfully developed for them. Native-born emigration is estimated using analytic techniques applied to census data from overseas sources and supplemented by State Department data. For a more complete description of the methods and data to estimate emigration, see Mulder et al, 2001 and Gibbs et al, 2001.

In summary, the DA method relies on current administrative data for most of the demographic components. Adjustments are made to account for certain deficiencies or incompleteness of the data sets. Emigration and residual foreign-born migration are two

relatively large components of population change based on estimation due to the lack of administrative data or large-scale survey-based sources for keeping the estimates “current”. For this decade, we are exploring the use of data on nativity and place of birth from the American Community Survey to update the estimates.

Limitations

Several limitations on the use of DA must be acknowledged:

Limited geographic and demographic detail of the DA estimates

The major DA estimates are available only at the national level and only for two broad race categories: Black and All Other Races Combined (the latter is referred to as “Non-Black”). We don’t have independent data to accurately measure internal migration within the United States (needed to do “subnational” DA estimates); historical time series of components of population change are not available for Hispanics or Asians to construct DA estimates comparable to the estimates for Blacks.

Inconsistencies in race classifications

The race categories in the DA estimates largely reflect the race assigned in the particular administrative record at the time of the event (birth, death, or enrollment in Medicare). The DA estimates of net undercount are biased to the extent that people who are classified as a particular race in DA (e.g., Black) reported a different race in the census.

The effect of the new “mark one or more” instruction for the Census 2000 question on race complicates the traditional comparison of DA estimates by race with census race tabulations. In fact, the Census 2000 tabulations do not include a category “Black” that is comparable to 1990 or earlier census tabulations. Tabulations for the Black population for 2000 contain tabulations of the number of people who reported Black only and tabulations of the number who reported Black whether or not they reported other races as well.

To deal with the reporting of more than one race, we present alternative DA estimates of census undercount using two models: (1) Model 1 compares the 2000 DA estimates for Blacks with Census 2000 tabulations for people who reported Black only, and (2) Model 2 compares the 2000 DA estimates for Blacks with Census 2000 tabulations for people who reported Black whether or not they reported any other race. At the youngest ages, the differences between the two models are the greatest. The tables and figures show the average of the two model estimates for comparison with the historical DA estimates and 2000 A.C.E. results. These averages are not necessarily the best point estimates; research on the detailed Census 2000 race and ethnicity data may provide a basis for determining at which point along the Model 1 to Model 2 range of census race tabulations the DA estimate might best be compared.

A final inconsistency affects race comparisons of the DA and the published census counts or A.C.E. estimates. In 2000, 15.3 million people (mainly Hispanics) who reported their

race as “Other Race-Not Specified” in the census were redistributed (for DA estimation) to the categories White; Black; American Indian, Eskimo, or Aleut; and Asian or Pacific Islander so that the census counts were consistent with the race categories of the historical demographic estimates.

The inconsistencies in the race data place even more importance on the use of sex ratios for making inferences about coverage by racial categories in Census 2000. Specifically, to the extent that the inconsistencies in reporting and the numbers marking more than one race are about the same for men and women, the inconsistencies will tend to cancel out in the calculation of sex ratios. We found this assumption held true: in Census 2000, the sex ratios for people who reported Black only are nearly identical to the sex ratios for people who reported Black whether or not they reported other races.

Uncertainty in DA estimates

The DA estimates contain uncertainty because the aggregate administrative data used to construct them are corrected for various types of errors and some components must be estimated. Most of these components are well measured (especially for recent decades) but several components of immigration have considerable uncertainty. Among the latter components are residual foreign-born migration, legal emigration, and the number of temporary legal migrants. We plan to develop sets of “confidence” measures for the 2000 DA estimates, extending the methods first developed by Das Gupta (1991) for assessing the accuracy of the 1990 DA estimates.

DA provides measures of net coverage, not separate components of coverage error

Finally, it should be noted that the principal DA estimates for race, sex, and age groups measure net undercount in the census. They don’t tell us about the separate effects of net coverage error (omissions, erroneous inclusions or duplicates) or content error.

DA RESULTS: March 2001

In this section we will examine the difference between the census and the initial DA estimates for 2000 released in March 2001. The first set of DA estimates that were developed and compared to the Census 2000 counts are referred to as “Base” DA. We will show how the decision was reached to create an alternative set of estimates, referred to as “Alternative” DA (which increased the estimate of unauthorized migration). In the next section we present the revised DA estimates released in October 2001.

Difference between Census and DA Estimates

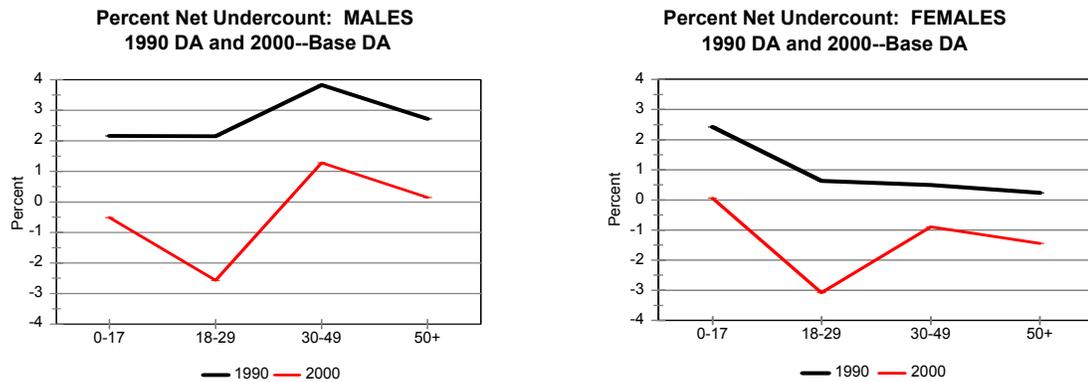
The difference between the Census 2000 count and the initial (Base) DA estimates is shown in Table 3. The census total of 281.4 million is 1.8 million higher than the Base DA estimate of 279.6 million. Relative to DA, the difference implies a net census overcount of 0.65 percent, which is expressed as a net undercount of -0.65 percent. This net coverage is dramatically different from that in the 1990 or any previous census. In 1990, the net undercount estimated by DA was 4.7 million or 1.85 percent.

Table 3. Comparison of Census and DA estimates, 1990 and 2000

	Count or Estimate (in millions)	
	2000 Base DA	1990 DA
DA (a)	279.6	253.4
Census (b)	281.4	248.7
Difference (a-b)	-1.8	4.7
Percent difference (a-b)/(a)	-0.65	1.85

The initial results by age and sex for 1990 and 2000 are illustrated in Figure 1. The 2000 estimates of net census undercount are lower than the 1990 estimates, but the differences are not uniform across age groups. Overcounts in the DA estimates for 2000 (base DA) are especially large for males and females in the age group 18 to 29. These are peak ages for immigration.

Figure 1. Estimates of Percent Net Undercount by Age and Sex, 1990 and 2000



The results for the total population in 2000 were unexpected. When examining the detailed DA estimates by age and sex, we began to realize that underestimation of immigration, particularly undocumented immigration, could be a reason for these unexpected results. We therefore conducted a systematic analysis using the error of closure.

Census-level error of closure

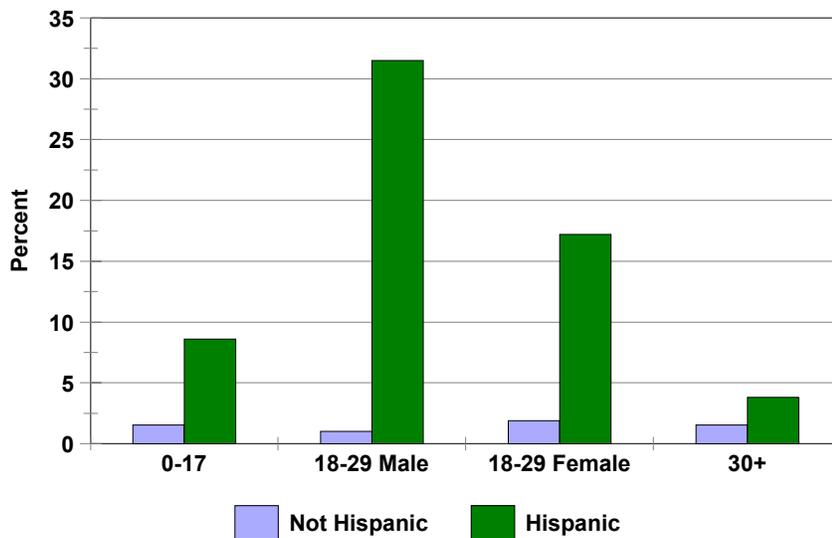
The error of closure reflects the difference of the Census 2000 count and the independent estimate based on carrying forward the 1990 census count with the components of change for 1990 to 2000. The error of closure measures the net effect of change in coverage in 1990 and 2000, and errors in the measured components of change. Although very crude, these errors of closure, when supplemented by the DA results, provide a way to assess the overall consistency of the census, A.C.E., and the demographic estimates. Major

deviations could signal a serious error in one of the systems, which should be investigated (see Robinson et al, 1993b, for an example of the error of closure analysis of the 1990 census results). In particular, we can check if the pattern of deviations by demographic characteristics such as age and sex provides clues to the possible understatement of immigration as suggested by the unusually large overcounts shown by the initial DA estimates at ages 18-29.

Figure 2 presents error of closure measures for the total population, Hispanics, and Non-Hispanics. The overall census results are 2.8 percent above the demographic benchmark for males and 2.2 percent above for females, implying higher net coverage in Census 2000 relative to 1990 and/or an understatement of the components of change (causing the 2000 DA estimate to be too “low”). The error of closure is relatively small and uniform across age-sex groups for the Non-Hispanic population. This may suggest that the 1990-2000 coverage change was relatively uniform by age and sex for Non-Hispanics, or that the error in component estimation did not disproportionately affect specific Non-Hispanic groups. The Black population falls mostly in this category.

On the other hand, the error of closure is around 10 percent for the total Hispanic population, and very different by age and sex. The error exceeds 30 percent for Hispanic men aged 18-29 (31.5 percent). Since coverage for this group undoubtedly did not improve by 31 percent relative to 1990, some of the error is likely caused by the understatement of growth (possibly through unauthorized immigration). Hispanics are largely classified in the Non-Black category for DA, so this large relative bias for Hispanic men aged 18-29 are a factor leading to the large DA overcounts for Non-Black men aged 18-29.

Figure 2: Error of Closure of Census 2000 and Demographic Component-Based Estimates



Other comparisons

We looked at other indicators, such as the estimated change in the Hispanic and the foreign-born population to judge how much the demographic components may have understated growth between 1990 and 2000. This understatement could come not only from the undocumented component, but also from an overstatement of emigration. For the purposes of simplicity, we assume that undocumented immigration between 1990 and 2000 would capture the understated growth in the initial DA estimates.

Table 4 shows the percent of the population that is foreign born, measured or implied by the 1990 census, the 2000 Current Population Survey (CPS) and DA. In 1990, 8.0 percent of the total population was foreign born, 35.8 percent of the Hispanic population was foreign born, and 5.3 percent of the Non-Hispanic population was foreign born. The March 2000 (CPS) indicates that the foreign born make up 10.9 percent of the total population, 39.4 percent of Hispanics, and 6.8 percent of Non-Hispanics. Comparing these data to the percents implied by the initial (base) DA provides further evidence of an understatement in the immigration component.

Table 4. Percent of the Population that is foreign born according to census counts, Current Population Survey estimates and Demographic estimates

	Total	Hispanic	Non-Hispanic
1990 census	7.95	35.81	5.26
Revised March CPS 2000 ¹	10.86	39.42	6.75
Implied by Base DA	10.26	36.52	6.63
Implied by Alternative DA	11.13	40.05	6.92

¹The published March 2000 Current Population Survey (CPS) data were controlled to age, race, sex and Hispanic origin estimates based on the 1990 census. The CPS data were recalculated using aggregate Hispanic origin data for Census 2000. Subsequently, the data were further refined by recalculating, using more detailed age, race, sex, and Hispanic origin numbers and the formal weighting procedures of the CPS. These results, which are included in this paper, are slightly different, but lead to the same conclusion.

Setting the alternative DA assumptions

Given the uncertainty in the estimates of the unauthorized immigration component, as well as the estimates for legal emigration and temporary migration, we examined the results of assuming various mixtures of increasing unauthorized immigration and temporary migration and decreasing emigration during the 1990's. Given the large proportion of Hispanics among unauthorized immigrants and the implied deficit of Hispanics in the base DA estimates, we chose to increase the unauthorized immigration component.

How much should we increase the estimate of net undocumented immigration? A doubling of the net increase in unauthorized immigrants during the 1990 to the 2000 decade, from 2.77 million to 5.53 million, would result in an estimated percent foreign born population of 11.1 percent, which is slightly above the CPS estimates; and an estimated Hispanic population of 12.7 percent in 2000, which is slightly above Census 2000. This leads us to believe that any further increase in the assumed net flow of unauthorized immigrants during the decade would likely be too high. In other words, to provide a reasonable upper estimate of unauthorized immigration, we chose to assume a doubling of the net flow of unauthorized from 2.8 million to 5.5 million.

The alternative DA estimate is shown in the second column of Table 5. The DA estimate of 282.3 million was 0.9 million above the Census 2000 count, implying a small net undercount of 0.32 percent. This is still much lower than the 1990 DA estimate of 1.85 percent, indicating a major reduction in net undercount (although not as large as suggested by the Base DA estimates).

Also shown in Table 5 are the estimates of net undercount for Census 2000 based on the Accuracy and Coverage Evaluation (A.C.E.). The DA estimate of 284.7 million was 3.3 million above the Census 2000 count, measuring a net undercount of 1.15 percent. Either of the two DA sets-the Base DA or Alternative DA-were below the March A.C.E. estimate of 284.7 million, and imply a much greater reduction in net undercount from 1990. These inconsistencies between the DA and the A.C.E. estimates of the population is the major reason the Census Bureau issued the recommendation of the Executive Steering Committee for A.C.E. Policy (ESCAP) on March 1, 2001 that the Census 2000 Redistricting Data not be adjusted for net census undercount based on data from the A.C.E.

Table 5. Comparison of Census 2000, Initial (base) DA and Alternative DA estimates, and Accuracy and Coverage Evaluation (A.C.E.) Estimates

	Count or Estimate (in millions)		
	Base DA	Alt DA	A.C.E.
Estimate (a)	279.6	282.3	284.7
Census 2000 (b)	281.4	281.4	281.4
Difference (a-b)	-1.8	0.9	3.3
Percent Difference (a-b)/a	-0.65	0.32	1.15

DA RESULTS: October 2001

Analysis of DA Components, March to October 2001

We knew that the Alternative DA estimates were interim numbers that would be replaced by a more complete analysis of data that became available later. Between March and October of 2001, we conducted an extensive review of the components of population change used to construct the DA estimates. This review led to the revision of

components (the revised components were presented in the section on DA methodology). The research was focused on two fronts:

1) Analysis of the administrative records used in the DA estimates (births, deaths, legal migration, Medicare).

First, we systematically reviewed the historical data on births and deaths. In particular, this review led to new assumptions about birth registration completeness since 1968, where we allowed registration completeness to gradually rise to 100 percent by 1985 (the first year natality statistics were reported electronically from all the States), rather than “freezing” the factor at the 99.2 percent of the 1968 test. This change lowered estimated births from 1968 to 2000 by 0.7 million.

Second, we reviewed the Medicare data for the population 65 and over, which led to some small changes in the estimation of underenrollment; raising the overall Medicare-based estimate by 0.1 million. Third, we systematically reviewed the legal immigration data from INS. This led to revisions of specific migration categories, lowering overall measured migration by 0.9 million. For example, our initial estimates included about 200,000 refugees who did not in fact receive permanent residence by April 1, 2000—these persons were removed from the “legal” category (and shifted to a “quasi-legal” group that were picked up in the residual foreign-born migrant estimates).

2) Re-calibration of the DA international migration components with current data.

The initial estimates of residual foreign-born migrants were based mainly on extrapolations of trends from the 1980’s. We needed to use current data to revise these estimates. To do this, we exploited early Census 2000 and Census 2000 Supplementary Survey data on the foreign born. These data were not available in March 2001. First, we revised the component that includes unauthorized migration using the residual method and Census 2000 foreign-born data (which raised the estimate by 1.4 million). We revised the estimate of temporary migrants using Census 2000 Supplementary Survey data (which raised the estimate by 0.1 million). Also, we revised the estimate of emigration (which raised the estimate by 0.2 million).

Revised DA Estimates

The DA estimates for 2000 that resulted from the analysis of various administrative records and the re-calibration of the international migration components are referred to as the “Revised” DA. Although the various analyses led to changes in the estimated components of births and international migration, the total DA population and demographic composition of the revised DA estimates are not significantly different from the Alternative DA estimates of March.

Compared to the Census 2000 count of 281.4 million, the Revised DA estimate of 281.8 million implies a net census undercount of 0.12 percent (see Table 6). The net census undercount in 2000 remains dramatically different from that in the 1990 under the revised DA set. In 1990, the revised net undercount was 4.2 million, or 1.65 percent.

Table 6. DA Estimate and Net Census Undercount Total Population

Category	1990 Census	2000 Census	
	Revised DA (October 2001)	Base DA (March 2001)	Revised (October 2001)
DA	252.9	279.6	281.8
Diff from Census	4.2	-1.8	0.3
% Difference	1.65	-0.65	0.12

The revised DA estimates demonstrate that the substantial reduction in net census undercount from 1990 to 2000 was shared by almost all demographic groups. The net census undercount of males and females each fell by about 1.5 percentage points (to an estimated net census undercount of 0.86 percent for males and estimated net census overcount of 0.60 percent for females in 2000). The estimated net undercount rate dropped more for Blacks (estimated net census undercount of 2.78 percent in 2000) than Nonblacks (estimated net census overcount of 0.29 percent in 2000), reducing the “differential” undercount of Blacks relative to Nonblacks from 4.4 percentage points in 1990 to 3.1 points in 2000.

Table 7. Demographic Analysis Estimates of Percent Net Census Undercount: 1990 and 2000

(a minus sign denotes a net overcount)

Category	Revised 1990 DA	Revised 2000 DA
Total	1.65	0.12
Male	2.39	0.86
Female	0.93	-0.60
Black	5.52	2.78
NonBlack	1.08	-0.29
Black Male, ages 20-64	11.31	8.44
Children, ages 0-4	3.72	3.84

Note: DA estimates represent revised estimates released in October 2001.

The revised DA estimates for 2000 show a broad decline in net census undercount rates for almost all race-sex-age categories. According to the revised DA estimates for 2000, the net census undercount rate remains disproportionately higher than the overall average for only two groups—Black men and young children. The estimated net census undercount of 8.4 percent for Black men aged 20-64 and 3.8 percent for children under age 5 are exceptions to the achievement of very high levels of net coverage in Census 2000.

Figure 3 displays more detailed estimates of percent net undercount for race, sex, and age groups in Census 2000. In terms of level of percent undercount, the high rates for Black men age 25-64 is the most notable pattern, where the estimated net undercount ranges from 6 to 11 percent. For Black males and females, the undercount rates for ages 0-4 are relatively high (though not as high as the rates for Black adult men). Also in contrast to the high net undercount estimates for Black adult men are the relative low undercount rates of Black adult women. For Non-Black males and females, the estimates of net undercount for 2000 are quite low across all ages—with the exception of young children (especially under age 5). In fact, the net census undercount estimates for Non-Black females fall below the zero undercount line for all 5-year age groups over age 10.

Figure 3. Percent Net Census Undercount by Race, Sex, and Age: 2000 Revised DA

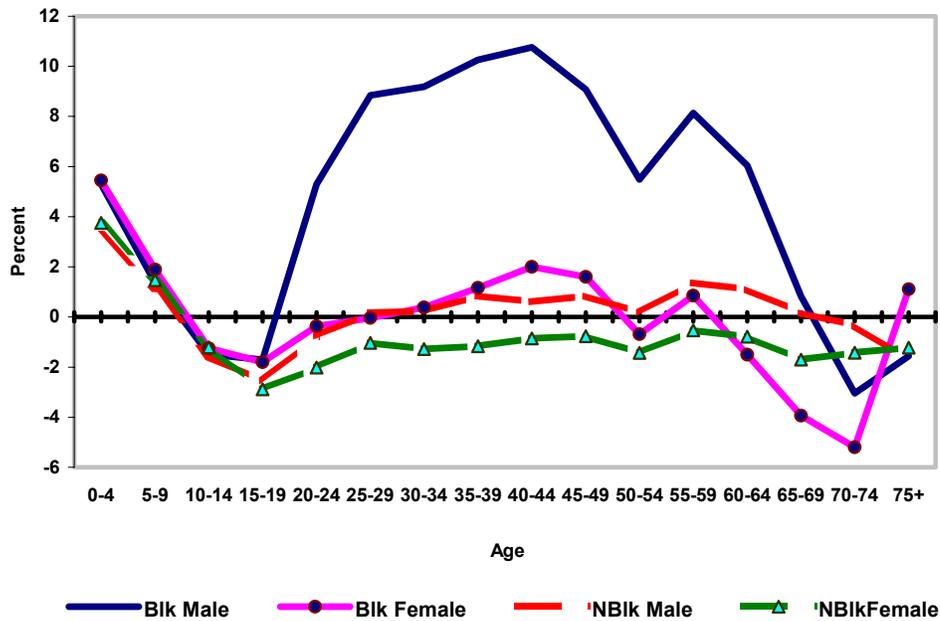
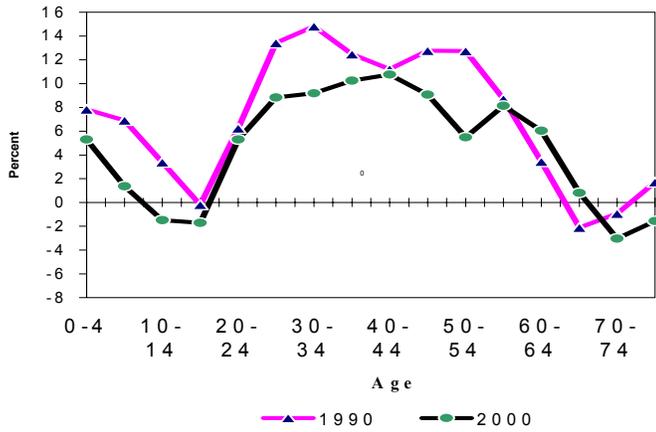


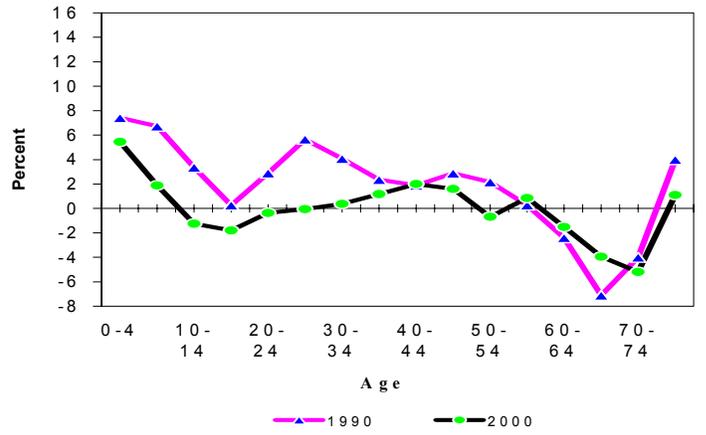
Figure 4 compares DA estimates of percent net census undercount for 5-year age groups in 1990 and 2000 (see also Appendix Tables 1 and 2). Compared to 1990, the revised DA estimates for 2000 show a broad decline in net census undercount rates for almost all race-sex-age categories. The net undercount rates for adult Black men aged 20-64 remain high in 2000 compared to other groups, though the rates are somewhat reduced from the even higher levels in 1990. The undercount rates for Black male and Black female children fell from 1990 to 2000. For Nonblacks, the net undercount rates were reduced from 1990 with one notable exception—children under age 5. For Nonblack females, ages 0-4 and 5-9 are the only age groups for which net census undercounts are observed; net census overcounts are measured for all age groups 10 and over. For Nonblack males, ages 0-4 is the only age group for which net undercounts of over 1.5 percent are measured in 2000 (in 1990, net undercount rates of 1.5 percent or more were estimated for ten 5-year age groups).

Figure 4. Revised Demographic Analysis Estimates of Percent Net Census Undercount: 1990 and 2000

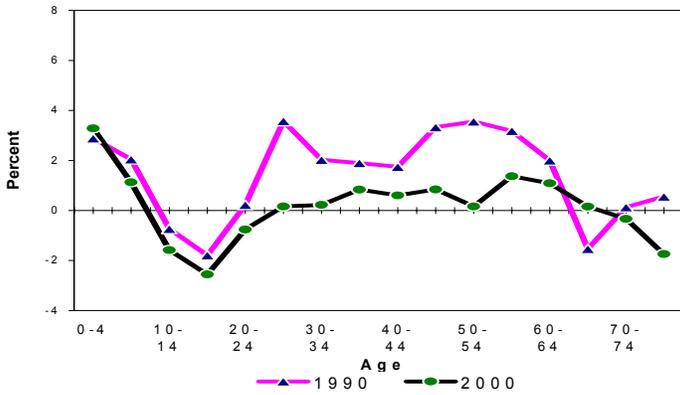
Black Male



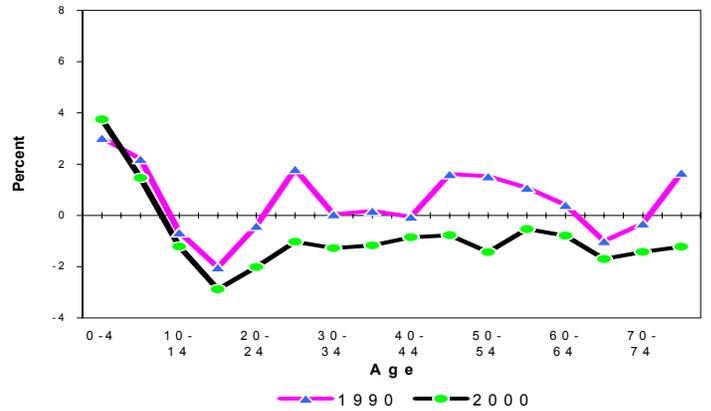
Black Female



Nonblack Male



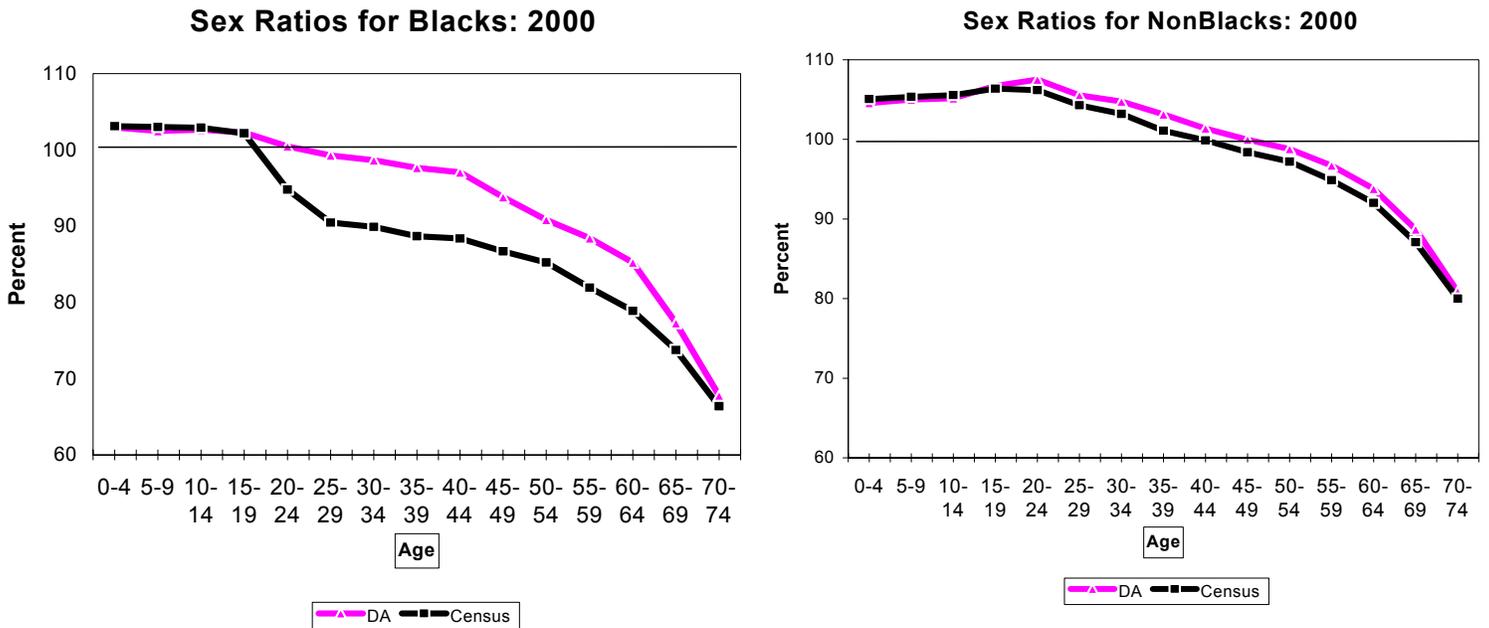
Nonblack Female



Sex ratios

The last evaluation tool of demographic analysis to be considered is the comparison of sex ratios (males per 100 females) of the demographic estimates and the census. The sex ratio from demographic analysis is generally considered to be one of the method's most robust measures. As illustrated in Figure 5, the sex ratios of the census counts for Blacks fall well below the "expected" demographic ratios between ages 20 and 65. These sex ratio patterns imply that the net undercount rates of Black men in the 2000 census are high *relative* to the undercount rates of Black women. The "gap" for Non-Blacks is much smaller in comparison, denoting smaller differences in the undercount rates of Non-Black men and women. These observations are consistent with the direct estimates of percent net undercount for race-sex-age groups shown in Table 3.

Figure 5: Comparison of Sex Ratios based on the Census and DA: 2000



Summary

In summary, the DA measures a substantial reduction in net undercount in Census 2000 compared to 1990. The reduction is found for all three sets--the initial (base) DA, the alternative DA, and the revised DA sets. The reduction occurs among all demographic categories: all broad age groups, males and females, Blacks and Non-Blacks. For only two of the demographic groups measured by DA do differential undercounts continue to persist in 2000—adult Black men and young children.

We will continue research to assess the quality of the DA estimates and identify ways to improve the estimates for future evaluations. We continue to evaluate the components of population change, especially the components of international migration. We are looking at ways to update the more difficult-to-measure migration components, such as using results from the American Community Survey.

We plan to expand the demographic scope of DA to include coverage evaluation benchmarks for Hispanics and other groups, and will continue the development of subnational demographic benchmarks.

Finally, we plan to develop a set of “confidence” measures for the 2000 DA estimates, expanding upon the earlier efforts to develop statistically-based assessments of accuracy of the DA results.

References

Ahmed, Bashiruddin, Douglas K. Sater, David L. Word, Myoung-Ouk Kim, Brenda J. Skillern. October 2001. "The Use of Medicare Data in the Demographic Analysis (DA) and in the Population Estimates Programs." Population Division Working Paper Series xx, U. S. Census Bureau, Washington, D.C. 20233 (Pending Release)

Cassidy, Rachel and Lucinda Pearson. 2001. "Evaluating Components of International Migration: Temporary (Legal) Migrants". Population Division Working Paper Series No. 60, U.S. Census Bureau, Washington, D.C. 20233

Christenson, Matthew. December 2001. "Evaluating Components of International Migration: Migration Between Puerto Rico and the United States", Population Division Working Paper Series No. 64, U.S. Census Bureau, Washington, D.C. 20233

Costanzo, Joseph, Cynthia J. Davis, Caribert Irazi, Daniel M. Goodkind, and Roberto R. Ramirez. 2001. "Evaluating Components of International Migration: The Residual Foreign Born". Population Division Working Paper Series No. 61, U. S. Bureau of the Census, Washington, D.C. 20233

Das Gupta, Prithwis. 1991. DA Evaluation Project D11: "Models for Assessing Errors in Undercount Rates Based on Demographic Analysis." Preliminary Research and Evaluation Memorandum No. 84, U.S. Bureau of the Census.

Gibbs, Jim C., Gregory S. Harper, Marc J. Rubin, and Hyon B. Shin. 2001. "Evaluating Components of International Migration: Native Emigrants," Population Division Working Paper Series No. 63, U.S. Census Bureau. Washington, D.C. 20233 (Pending Release)

McDevitt, Tom, Martin O'Connell, Colleen Joyce, 2001. "Evaluating the Components of Births and Deaths used in Demographic Analysis." Population Division Working Paper Series xx, U. S. Census Bureau, Washington, D.C. 20233 (Pending Release)

Mulder, Tammany, Betsy Guzman, and Angela M. Brittingham. 2001. "Evaluating Components of International Migration: Foreign-Born Emigrants". Population Division Working Paper Series No. 62. U.S. Census Bureau, Washington, D.C. 20233 (Pending Release)

Passel, Jeffrey S. 1991. "Age-Period-Cohort Analysis of Census Undercount Rates for Race-Sex Groups, 1940-1980." Proceedings of the Social Statistics Section of the American Statistical Association.

Passel, Jeffrey. S. 1990. "Demographic Analysis: A Report on Its Utility for Adjusting the 1990 Census," Washington, D.C.: The Urban Institute.

Perry, Marc J., Barbara J. Van der Vate, Lea Auman, and Kathleen Morris. 2001. "Evaluating Components of International Migration: Legal Migrants". Population Division Working Paper Series No. 59. U. S. Census Bureau, Washington, D.C. 20233

Robinson, J. Gregory, B. Ahmed, P. Das Gupta, and K. A. Woodrow. 1993a. "Estimation of Population Coverage in the 1990 United States Census Based on Demographic Analysis." *Journal of the American Statistical Association*, Vol. 88, No. 423, pp. 1061-1071.

Robinson, J. Gregory, B. Ahmed, and E. Fernandez. 1993b. "Demographic Analysis as an Expanded Program for Early Coverage Evaluation of the 2000 Census." *Proceedings of the 1993 Annual Research Conference*. U.S. Bureau of the Census. Arlington, Virginia, pp. 166-200.

Robinson, J. Gregory, P. Das Gupta, and B. Ahmed. 1990. "Case Study of the Investigation of Errors in Estimates of Coverage Based on Demographic Analysis: Black Adults Aged 35 to 54 in 1980." In *Proceedings of the Social Statistics Section of the American Statistical Association*.

Siegel, Jacob S. and Melvin Zelnik. 1966. "An Evaluation of Coverage in the 1960 Census of Population by Techniques of Demographic Analysis and by Composite Methods." In *Proceedings of the Social Statistics Section of the American Statistical Association*.

U.S. Bureau of the Census. 2001a. "Accuracy and Coverage Evaluation: Demographic Analysis Results," by J. Gregory Robinson. *DSSD Census 2000 Procedures and Operations Memorandum Series B-4*, March 2, 2001

U.S. Bureau of the Census. 2001b. "ESCAP II: Demographic Analysis Results," by J. Gregory Robinson. *Executive Steering Committee on Accuracy and Coverage Evaluation Policy II*, Report Number 1, Oct. 13, 2001.

U.S. Bureau of the Census. 1991a. DA Evaluation Project D1: "Error in the Birth Registration Completeness Estimates" *Preliminary Research and Evaluation Memorandum No.74 (PERM)* by J. Gregory Robinson.

U.S. Bureau of the Census. 1991b. DA Evaluation Project D9: "Inconsistencies in Race Classifications of the Demographic Estimates and the Census" *PERM No. 82*, by J. Gregory Robinson and Susan Lapham.

U.S. Bureau of the Census. 1991c. DA Evaluation Project D7: "Uncertainty Measure for Other Components" *PERM No.80*, by J. Gregory Robinson, Karen A. Woodrow, and Bashir Ahmed.

U.S. Bureau of the Census. 1974. "Estimates of Coverage of Population by Sex, Race, and Age: Demographic Analysis," by Jacob S. Siegel, Evaluation and Research Reports, PHC(E)-4.

U.S. Bureau of the Census. 1988. "The Coverage of Population in the 1980 Census," by Robert Fay, Jeffrey S. Passel, and J. Gregory Robinson. Evaluation and Research Reports, PHC80-E4.

Appendix Table 1.

Demographic Analysis Estimates of Percent
Net Undercount by Age and Sex: 1990 and 2000

Age groups	Male		Female	
	1990	2000	1990	2000
All ages	2.4	0.9	0.9	-0.6
0-4	3.7	3.6	3.8	4.0
5-9	2.8	1.2	2.9	1.5
10-14	-0.1	-1.6	0.0	-1.2
15-19	-1.5	-2.4	-1.7	-2.7
20-24	1.1	0.1	0.1	-1.8
25-29	4.9	1.4	2.4	-0.9
30-34	3.7	1.4	0.6	-1.0
35-39	3.2	2.1	0.5	-0.8
40-44	2.8	1.9	0.2	-0.5
45-49	4.3	1.8	1.8	-0.5
50-54	4.5	0.7	1.6	-1.3
55-59	3.7	2.0	1.0	-0.4
60-64	2.1	1.6	0.1	-0.9
65-69	-1.6	0.2	-1.5	-1.9
70-74	0.1	-0.5	-0.6	-1.8
75+	0.6	-1.7	1.9	-1.0

Appendix Table 2.

Demographic Analysis Estimates of Percent Net Undercount
by Age, Sex, and Race: 1990 and 2000

Age groups	Black				Non-Black			
	Male		Female		Male		Female	
	1990	2000	1990	2000	1990	2000	1990	2000
All ages	8.1	5.1	3.1	0.5	1.6	0.2	0.6	-0.8
0-4	7.8	5.3	7.4	5.4	2.9	3.3	3.0	3.8
5-9	6.9	1.4	6.7	1.9	2.0	1.1	2.2	1.5
10-14	3.4	-1.5	3.4	-1.2	-0.7	-1.6	-0.7	-1.2
15-19	-0.2	-1.7	0.3	-1.8	-1.8	-2.6	-2.0	-2.9
20-24	6.2	5.3	2.9	-0.4	0.2	-0.8	-0.4	-2.0
25-29	13.4	8.8	5.7	0.0	3.6	0.2	1.8	-1.0
30-34	14.8	9.2	4.1	0.4	2.0	0.2	0.0	-1.3
35-39	12.5	10.3	2.4	1.2	1.9	0.8	0.2	-1.2
40-44	11.2	10.8	1.9	2.0	1.7	0.6	-0.1	-0.9
45-49	12.8	9.1	2.9	1.6	3.3	0.8	1.6	-0.8
50-54	12.7	5.5	2.2	-0.7	3.6	0.2	1.5	-1.4
55-59	8.7	8.1	0.3	0.9	3.2	1.4	1.1	-0.5
60-64	3.4	6.0	-2.4	-1.5	2.0	1.1	0.4	-0.8
65-69	-2.1	0.8	-7.1	-3.9	-1.5	0.2	-1.0	-1.7
70-74	-0.9	-3.0	-4.0	-5.2	0.1	-0.3	-0.3	-1.4
75+	1.7	-1.6	4.0	1.1	0.6	-1.7	1.7	-1.2