

Census Coverage

Presentation to INFO 747

April 26, 2007

Decennial Census Undercount

1. What is it...
2. How big?
3. How do we know?
4. What do participant observation studies tell us about who is missed and why?
5. Why does the UC matter?

What is it?

Estimate population by some alternative means; produce estimate of Census net undercount = persons missed - erroneous inclusions (e.g., duplicate records).

P_i is alternative est. of a population i

C_i is Census count of population i

Census Net Undercount = $(P_i - C_i)$

Net Undercount % = $(P_i - C_i) / P_i$

Differential Net Undercount Biases Rate Comparisons

U_i and U_j are net undercount proportions for populations i and j .

E_i and E_j are event counts for i and j (e.g., deaths in 2000)

$$RR = (E_i/P_i) / (E_j/P_j)$$

$$\text{Estimated } RR = E_i/C_i / E_j/C_j$$

$$[E_i/P_i(1-U_i)] / [E_j/P_j(1-U_j)] \rightarrow$$

$$\text{Estimated } RR = RR \times (1-U_j)/(1-U_i)$$

How do We Know?

Three Coverage Assessment Methods

Demographic Analysis

Participant observation studies

e.g., 30 studies in 1990 “Alternative Estimations”

Dual System Estimation

“Post-Enumeration Survey” (1990)

A.C.E. (Accuracy & Coverage Evaluation 2000)

capture/recapture methods +

Demographic Analysis

Basic DA Method

- Cohorts Under age 65 at 4/1/00

$$P = B - D + I - E$$

- Cohorts 65+ at 4/1/00

Medicare enrollment + under-enrollment adjustment

- Admin. Data Sources: Birth and Death registrations, Medicare Records, INS, ...

Basic DA Method

- Immigration: legally admitted permanent residents (INS) + other authorized immigrants + “residual” immigrants including unauthorized (estimated from a preliminary Census sample)
- Emigration
state department and foreign sources

From Robinson et al. 2002

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

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

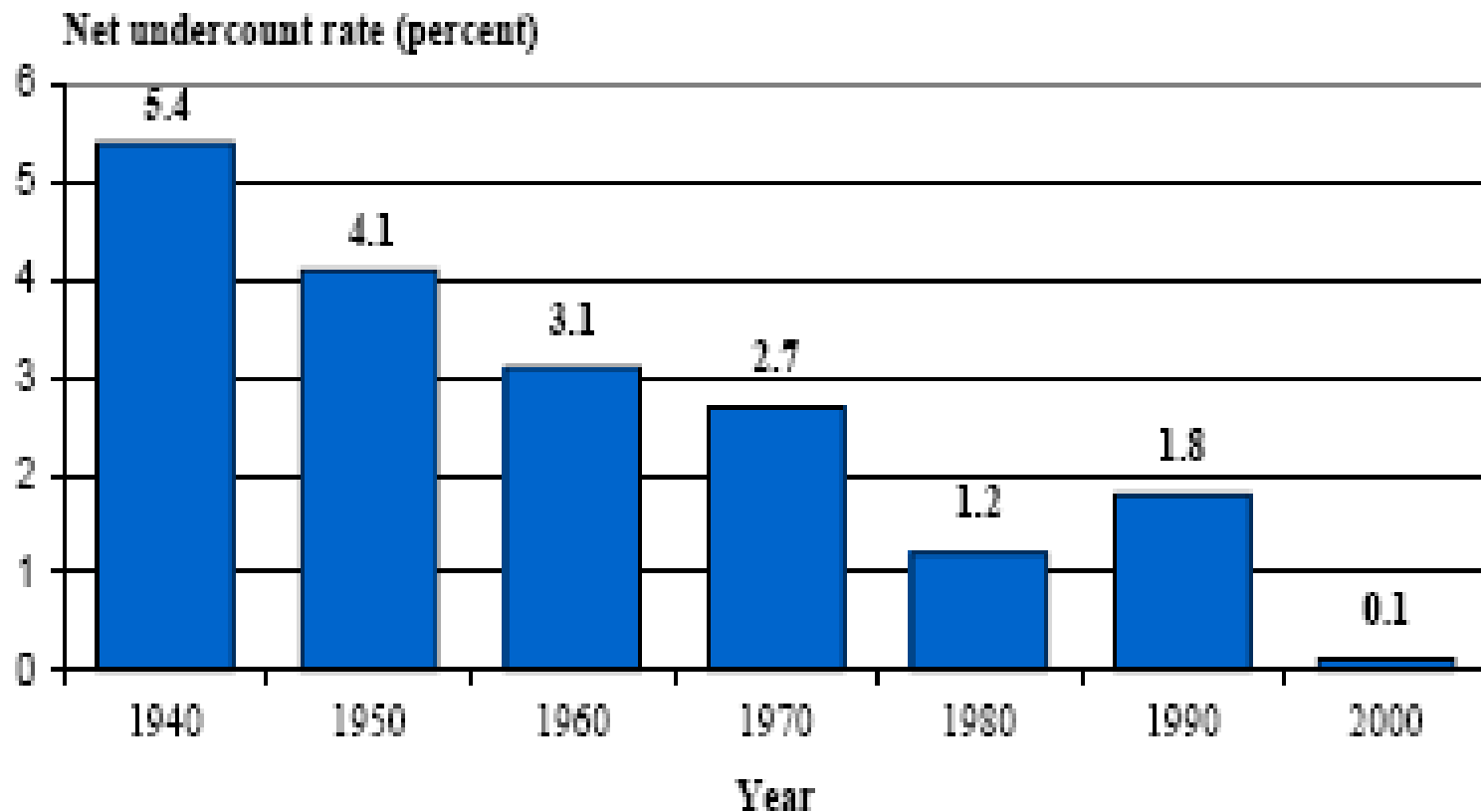
Where FB = Census 2000 foreign-born population
L = Legal immigrants
M = Mortality to legal immigrants
E = Emigration of legal immigrants
T = Temporary (legal) migrants
R = Residual foreign born

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-

Trends and demographic patterns of Census Undercount from DA

Coverage improved 1940 - 2000 except 1980-90

Figure 2: Net Population Undercount Rates, 1940-2000



Differential net undercount of Blacks has not improved much

Figure 3: Difference in the Net Undercount Rate Between Blacks and All Other Population Groups, 1940-2000

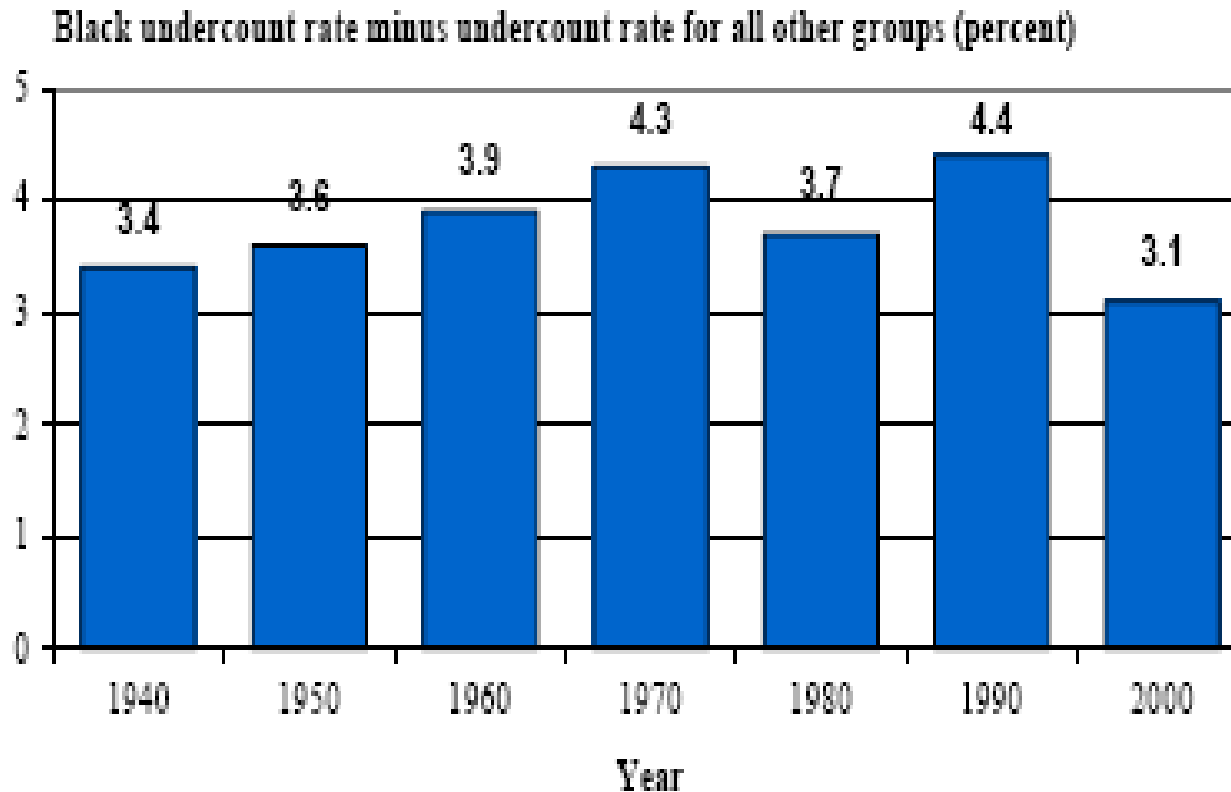


TABLE 2.1 Net Population Undercount in the Census by Demographic Analysis, 1940-1990

Undercount Rates and Numbers	1940	1950	1960	1970	1980	1990
Total						
Population (millions)	131.7	150.7	179.3	203.3	226.6	248.7
Undercount rate (%)	5.4	4.1	3.1	2.7	1.2	1.8
Number undercounted (millions)	7.0	6.3	5.6	5.5	2.8	4.7
Nonblacks						
Population	118.8	135.7	160.5	180.7	199.9	218.2
Undercount rate	5.0	3.8	2.7	2.2	0.8	1.3
Number undercounted	5.9	5.2	4.3	4.0	1.6	2.9
Blacks						
Population	12.9	15.0	18.9	22.6	26.7	30.5
Undercount rate	8.4	7.5	6.6	6.5	4.5	5.7
Number undercounted	1.1	1.1	1.3	1.5	1.2	1.8
Difference: black-nonblack net undercount rate	3.4	3.6	3.9	4.3	3.7	4.4

Note: Alaska and Hawaii became states in 1959. For 1950 and earlier, the population data and undercount estimates are for the 48 coterminous states. For 1960 and after, the data include Alaska and Hawaii.

Source: Population data: Bureau of the Census (1993b, Table 1). Undercount rates: Robinson et al. (1993).

Prominent Race X Age Patterns

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

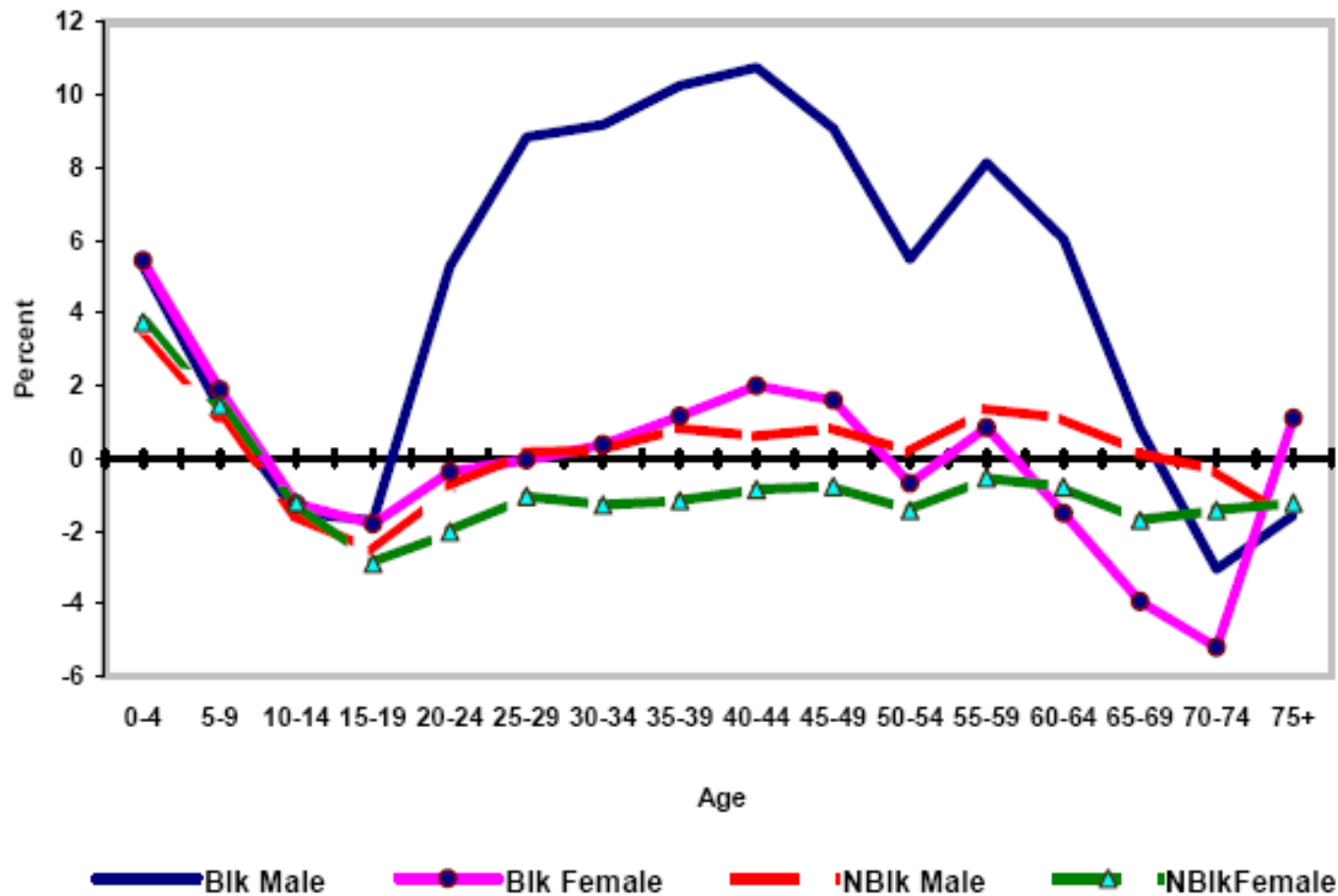
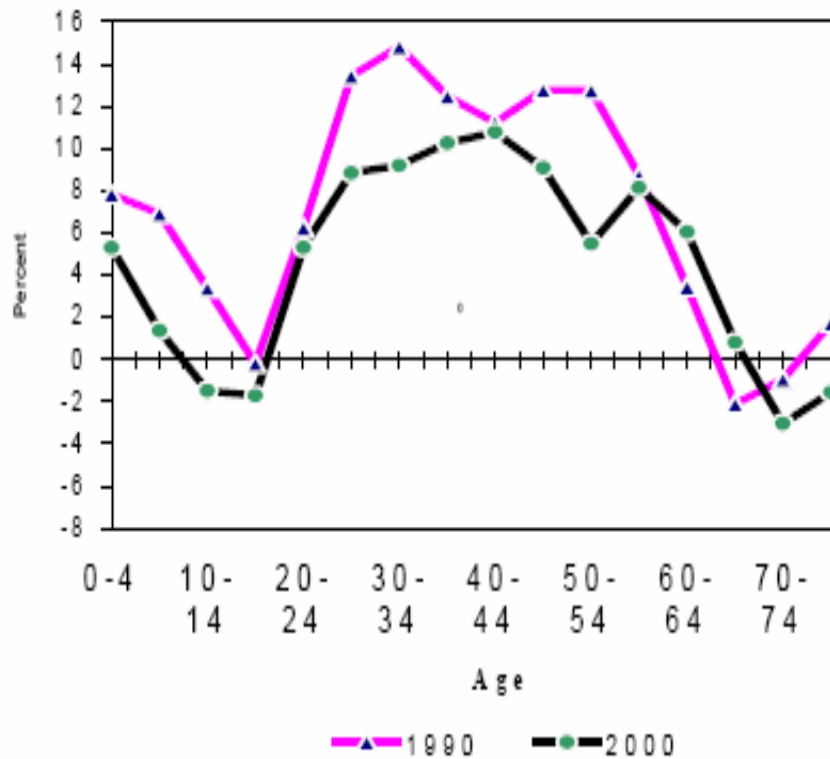
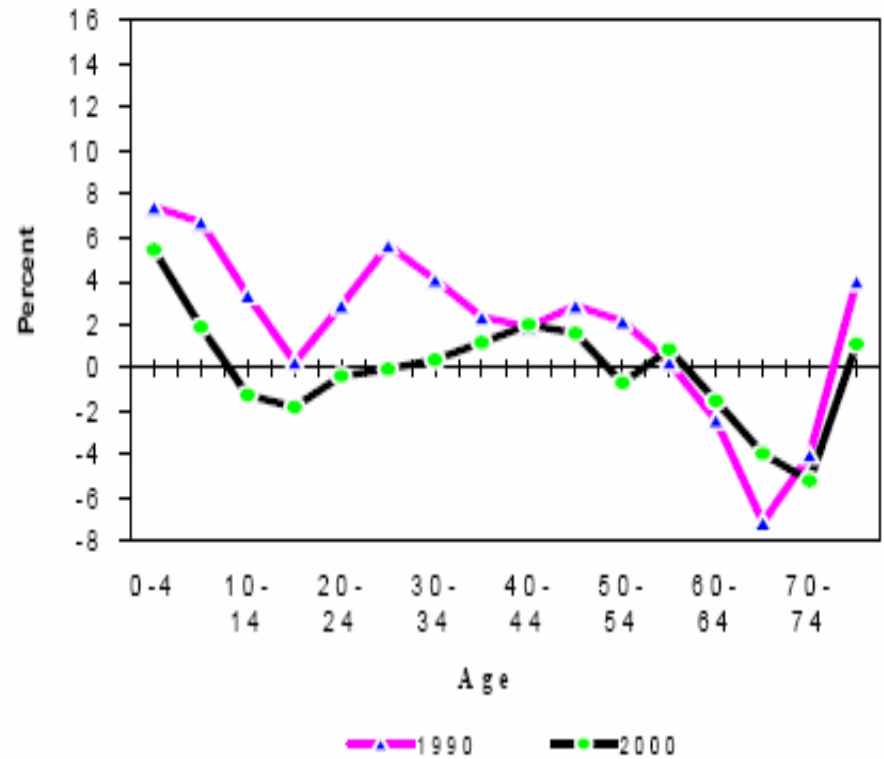


Figure 4. Revised Demographic Analysis Estimates of Percent Net Census Undercount: 1990 and 2000

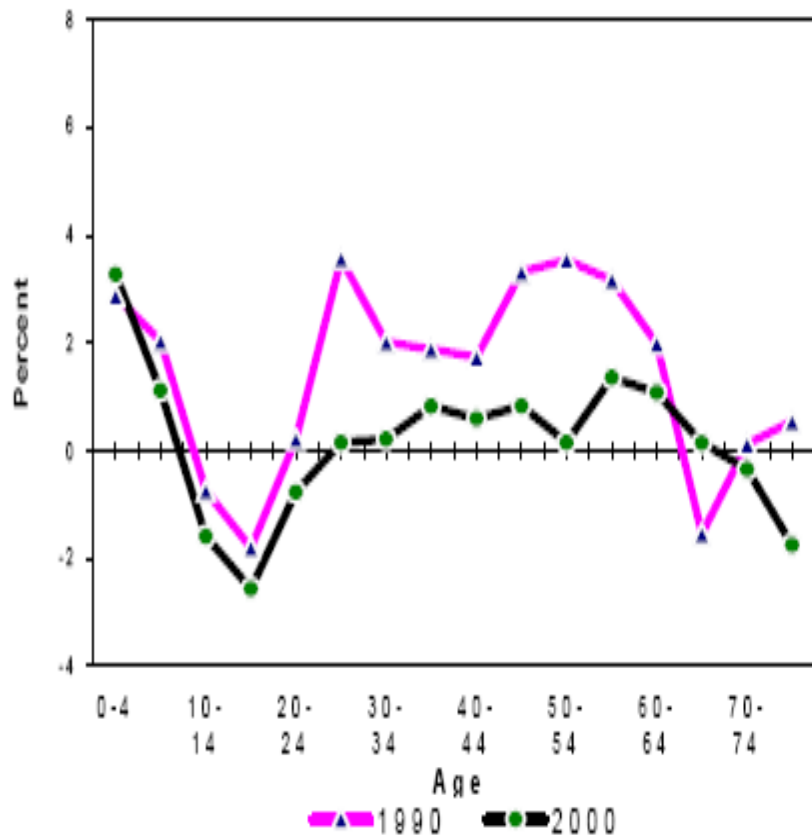
Black Male



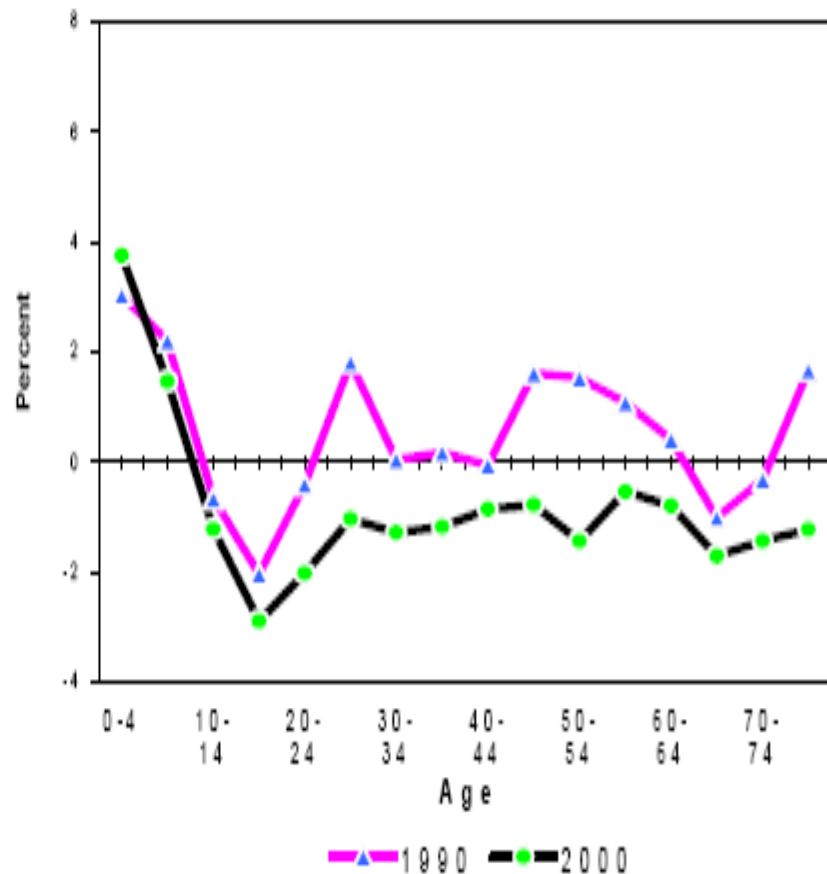
Black Female



Nonblack Male



Nonblack Female



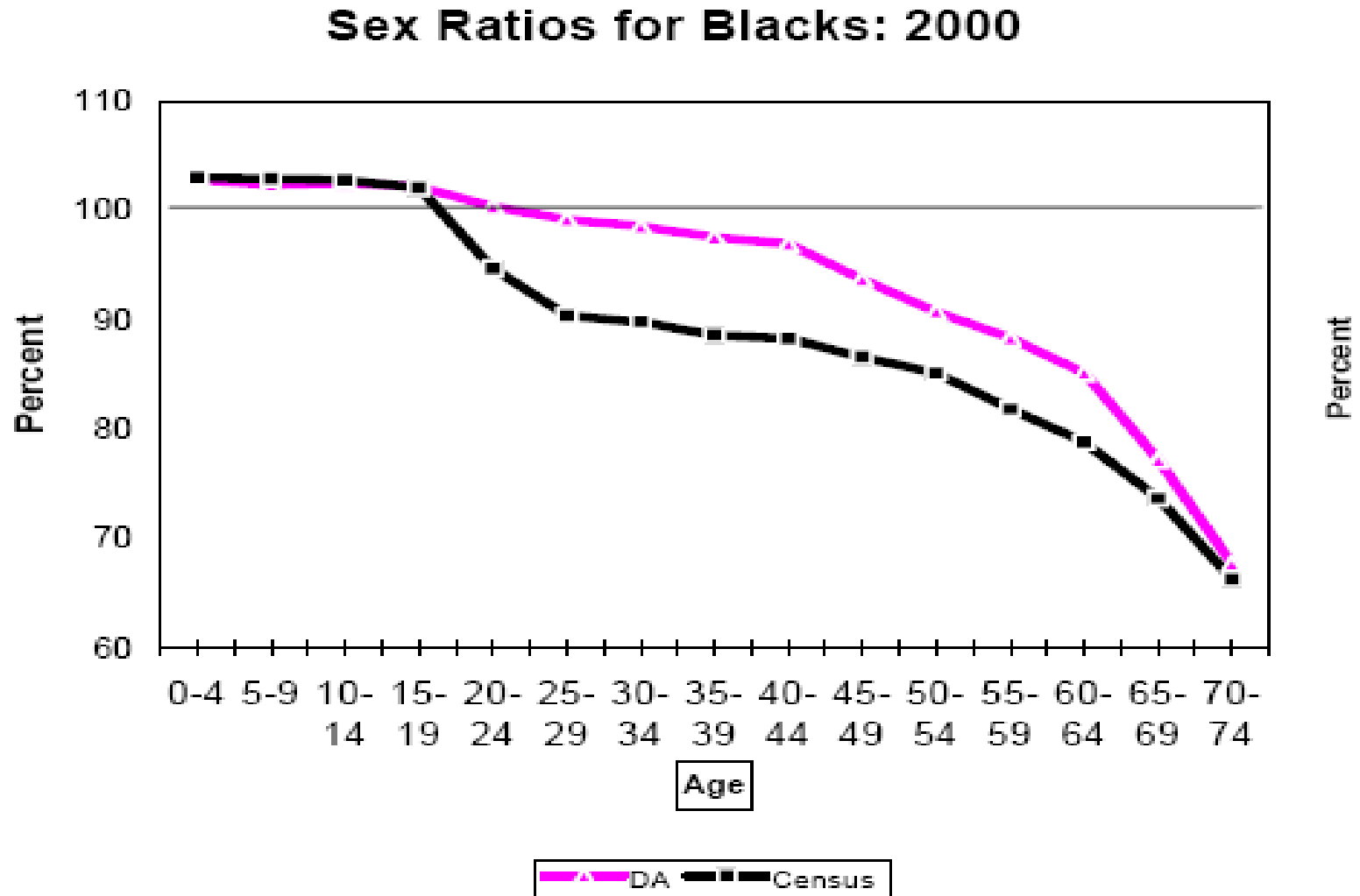
**Table 7. Demographic Analysis Estimates of Percent Net Census Undercount:
1990 and 2000**

(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.

Why Undercount Matters: Biased estimates of sex ratios



Some Sources of Error in DA

- Incomplete birth registration, esp. <1965
- Incomplete death registration, esp. infants
- Inconsistency in race classification:
vital records vs. Census
e.g., births to parents of different races.
Since 1989: race of mother
Pre-1989: race of father, unless mother black
and father white (ex. Hawaiian)
DA: race of father;
- Unauthorized immigration

Limitations of DA

Race: mostly black vs. non-black

Only national: lack internal migration ests.

Multiple race reporting in 2000 Census
thought to complicate comparisons

Provides estimates of net coverage only.
Do two wrongs make a right?

Participant Observation Studies of Low Income, Immigrant, Minority Communities

B. Lee Sung: Chinatown

Six six-story tenement buildings with about 100 housing units (apts.).

Estimate: 14% to 18% net undercount.

Census missed 10 housing units out of 98 (often “non-obvious partitions”)

Most people who were missed lived in households that were enumerated.

Sung: Causes, most to least important

1. Inaccessibility: physical & time
2. Resistance: fear, distrust, concealment
government, authority, crime
immigration, IRS, landlord, etc.
3. Language
4. “Wrong” enumerators
5. Cultural differences: age, name,
6. Immigrants (old and new)
7. “Irregularities”: subdivided apts., mobility,
multiple residences

A. Hamid, Central Harlem

Housing units 002 to 011 were in such a brownstone. I am sure that there are records to show that once upon a time, this was a modest brownstone housing a snug nuclear family of husband, wife and young children. This view of the building remains unchanged in the 1990 Census. One questionnaire was mailed out to the address, and an enumerator, walking past it (he could not have gone in!) listed it as one housing unit, containing three individuals. The reality, of course, is quite different. In 1984, a minister of a nearby Baptist church bought the property "to respond to the problem of homelessness." He divided the brownstone into ten housing units, and rented each apartment at \$350-\$375 a month. Since few repairs have been made to the heavy wear-and-tear on public spaces like stairways and shared bathrooms, the aspect inside the building is grimly Victorian. Twenty one individuals live in the building today, including two children under 10. Figure 1 illustrates the situation and the census outcome.

Figure 1

Actual situation	Census outcome
Converted Brownstone # 1 Housing units A 17 002 to 011	
21 people in 10 housing units	Family of 3 in 1 housing unit (single family house)
	Totals
<hr/> Census Day residents: 21	<hr/> Censused population: 3

Five More Converted Brownstones

Actual Situation

Census outcome

Converted Brownstone # 2: HUs A17 002 to 011 = B17 563]

22 people living in
11 housing units

1 person living in
1 housing unit

Converted Brownstone # 3: HUs A17 028 to 035 = B17 571]

4 people living in
4 housing units plus
4 other vacant apartments

1 person living in
1 housing unit

Converted Brownstone # 4: HUs A17 045 to 051 = B17 639 et al]

19 people living in
7 housing units

5 people living in
5 apartments plus
3 other vacant apartments

Converted Brownstone # 5 : HUs A17 082 to 087 = B17 886

6 people living in
6 housing units plus
1 vacant apartment

1 person living in
1 housing unit

Converted Brownstone # 6 : HUs A17 096 to 102 = B17 017 et al

19 people living in
7 housing units

4 people living in
3 housing units plus
6 vacant apartments

Totals

Census Day Residents: 70

Censused population: 12

Some Causes (summary)

- unwilling to cooperate (welfare; crime; landlord; immigration; fear; inconvenience)
- unable (e.g., no "usual place of residence"; complex family/HH defs)
- low social visibility (e.g. homeless)
- missed units (e.g. illegally subdivided)

S. Mahler, Salvadorans on Long Island

Census missed about 47% (though 31 / 118 “misgeocoded”)

80% “whole household” misses

20% persons missed from enumerated households.

Of 118 people missed, 50 were Salvadorans.

About 50% of those missed were undocumented immigrants.

Causes

expected: fear of deportation;

actual: fear that disclosing borders will lead to eviction
(landlord, illegal subdivide, etc....)

- Mobility
- Irregular housing
- Irregular households
- Civil war in home country keep Salvadorans from getting to know too much about one another in US.
- Language & literacy: perhaps 50% illiterate in Spanish
- Newness of the group/lack of political power

Strengths of P.O. Studies

- Insights on behavioral and other causes of undercount
- Suggestions for improving coverage
- Can target populations known to be difficult to enumerate
- Learn something about low-income and immigrant populations, and interaction with government.

Weaknesses of P.O.

Missed in AE site, but counted elsewhere?

Generalize?

Enumerator motives?

Dual System Estimation (PES / ACE)

Some ACE II Highlights

- Used Dual System Estimation
- ACE II focus on erroneous inclusions (vs. past DSEs focus on omissions)
- Used DA sex ratios to improve estimates of correlation bias.

Basic PES Method:

Sample again, with independence

Capture-recapture...to estimate total pop.

“modified for Census context”

Example: “fish in a lake”

- 800 in first (tag); 600 in second, of which 500 are “tagged” (i.e., 500 in both)

- If independent: $p(\text{both}) = p(\text{first}) * p(\text{second})$

Replace population probs w/sample proportions

$500/\text{pop} = (600/\text{pop}) * (800/\text{pop}) \dots$ solve for “pop”
(= 960).

More...

Some Challenges

1. How well are matches determined?

- sufficient info to determine match?

Drop “Insuff. Info.” cases from Census

2. "Wiley trout"...

Are p_1 & p_2 independent? Corr. bias corrections.

And pop of unknown size has zero prob. of capture.

• ACE II: used DA for some corr. bias corrections.

3. Adjust census count down for incorrect

© 2007 John M. Abowd, Sanders Korenman,
all rights reserved

enumerations

Return to simple example

DS Estimated **pop** = $(800 \times 600) / 500 = 960$

ACE

E-sample: eliminate erroneous enumerations
(800 too high) → DS **pop** high → UC too high

P-sample match rate errors, e.g.,
(500 too low) → **pop** high → UC too high

ACE II reduced UC estimate from about 1.2% to net overcount of 0.5

Table 13: Census Count, Demographic Analysis (DA) Estimate and A.C.E. Revision II Estimate for the U.S. Resident Population: April 1, 2000 (a minus sign indicates a net overcount)

	Count or Estimate
1. Census Count	281,421,906
2. DA Estimate	281,759,858
3. A.C.E. Revision II Estimate	280,090,250
Net Census Undercount (Amount)	
4. DA Estimate (=2-1)	337,952
5. A.C.E. Revision II Estimate (=3-1)	-1,331,656
Net Census Undercount (Percent)	
6. DA Estimate (=4/2*100)	0.12
7. A.C.E. Revision II Estimate (=5/3*100)	-0.48

Source: U.S. Census Bureau

Note: 1) A.C.E. Revision II estimate includes an adjustment for correlation bias, based on the DA sex ratios for adult males.
 2) DA estimate reflects revised estimate published in U.S. Bureau of the Census, 2001, ESCAP II, Report No. 1, October 13.

Some Findings from ACE II

Census 2000 net overcount

Diff. net undercount by race reduced

Overcount of NH Whites = 1.13 %

Undercount of NH Blacks = 1.85 %

ACE II provides net est. only.

additional assumptions needed for gross
coverage errors...however

Some Findings ACE II

Census 2000 count includes at least 5.8 million duplicates.

Census 2000 includes other erroneous enumerations (non-duplicates) such as fictitious persons, non-residents.

Do two wrongs make a right?

If same “relevant characteristics” and geography of omissions, then erroneous inclusions improve Census accuracy.

Lessons from ACE II for 2010?

From: NAS Counting under Adversity

- Groundbreaking work on duplication by name and birthdate matching
- Residence rule (“usual residence”) may be problematic, especially for
 - College students
 - Children in joint custody
 - Others with more than one residence.
- Proxy data are error prone

Have these problems grown over time?

Why Undercount Matters

Some Census Uses

Congressional apportionment (art. 1, sec.2)

Allocation of Federal / state funds

Research: biased estimates

– Direct

- Census counts as denominators in rate estimation
- Estimates of proportions & means from Census

– Indirect

e.g., Census counts and estimates used as population controls for social surveys; weighting & validation

References

- Committee on National Statistics, NAS. 2004. *The 2000 Census: Counting Under Adversity*. National Academy Press. Chapters 5 and 6.
- Sung, Betty Lee. 1991 "[Behavioral Causes of Census Undercount: New York City's Chinatown](#)". Report #6. U.S. Bureau of the Census, Center for Survey Methods Research. August.
- Hamid, Ansley. 2001. "[Ethnographic follow-up of a Predominantly African American Population Sample in A Sample Area in Central Harlem, New York: Behavioral Causes of the Undercount of the 1990 Census.](#)" Report #11. U.S. Bureau of the Census, Center for Survey Methods Research. January.
- Mahler, Sarah. 1993. "[Alternative Enumeration of Undocumented Salvadorans on Long Island](#)". Report #26. U.S. Bureau of the Census, Center for Survey Methods Research. January.
- Edmonston, B. 2002. The undercount in the 2000 Census.
- Robinson, J. G., A. Adlakha and K. West. "[Coverage of Population in Census 2000: Results from Demographic Analysis](#)". Paper presented to the 2002 Meetings of the Population Association of America. Atlanta. May 8, 2002.
- U.S. Bureau of the Census. [Technical Assessment of A.C.E. Revision II](#). March 12, 2003.
- U.S. Bureau of the Census. [Decision on Intercensal Population Estimates](#). March 12, 2003.