BUREAU OF THE CENSUS
STATISTICAL RESEARCH DIVISION REPORT SERIES

SRD Research Report Number: CENSUS/SRD/RR-86/20

CAPTURE-RECAPTURE ESTIMATION IN THE PRESENCE OF A KNOWN SEX RATIO

by

Kirk M. Wolter
Chief, Statistical Research Division
U.S. Bureau of the Census
Room 3536, FOB #3
Washington, D.C. 20233

(301) 763-3807

Recommended by: Kirk M. Wolter

Report Completed: August 26, 1986

Report Issued: September 4, 1986

August 26, 1986

Capture-Recapture Estimation in the Presence of a Known Sex Ratio

by

Kirk M. Wolter*

1. Introduction

Potential failure of the independence assumption has been cited as a key problem for the dual-system or capture-recapture model for more than 30 years. In this note, a certain solution to the problem is achieved by replacing the independence assumption by another assumption thought generally to be less objectionable. The solution may have wide areas of application, but is specifically applicable to the problem of estimating coverage error in a census of a human population.

Let there be two, two-way tables corresponding to males and females, respectively:

<u>Males</u>

		List B: in	List B: out	
List A	in	p_{111}	p_{112}	p_{11+}
	out	p_{121}	p_{122}	p_{12+}
		p_{1+1}	p_{1+2}	1

*Kirk M. Wolter is Chief, Statistical Research Division, U.S. Bureau of the Census and Professorial Lecturer, The George Washington University. The author gratefully acknowledges comments by Gary Koch on an early draft of the article, and computational assistance provided by Andrea Meier.

Females

		List B in	List B out	
List A	in	p_{211}	p_{212}	p_{21+}
	out	p_{221}	p_{222}	p_{22+}
		p_{2+1}	p_{2+2}	1

The observed data are x_{1jk} and x_{2jk} for $(j,k) = (1,1), (1,2)$ and $(2,1)$. As usual, the (2,2) cells are not observed, nor are the total population sizes

$$N_1 = x_{111} + x_{112} + x_{121} + x_{122}$$

$$N_2 = x_{211} + x_{212} + x_{221} + x_{222}.$$

We shall study the problem of estimating N_1 and N_2 under the usual capture-recapture or dual-system assumptions, except we shall not require that the two lists be independent.

Instead, we study two models where the sex ratio

$$r = N_1/N_2$$

is assumed known. In the first model, the cross-product ratios

$$\theta_1 = \frac{p_{111}\, p_{122}}{p_{112}\, p_{121}}$$

$$\theta_2 = \frac{p_{211}\, p_{222}}{p_{212}\, p_{221}}$$

are assumed equal (i.e., $\theta_1 = \theta_2 = \theta$). Thus, we permit list association, but to the same degree for males and females. In the second model, we assume independence for females

$$\theta_2 = 1 ,$$

but not for males.

For either model, the estimation problem now consists of six parameters and six obervable statistics. Two of the original eight parameters are eliminated by the assumptions. We proceed to construct estimators for the remaining parameters, including the applicable cross-product ratio.

The estimation schemes are presented in Section 2, and we follow that with an example involving data from the 1980 U.S. Census in Section 3. The article closes with a short summary.

2. Estimation Schemes

First, we consider Model 1. Under multinomial sampling, the log-likelihood is

$$\begin{aligned} \mathcal{L} = {} & \log N_1! - \log x_{111}! - \log x_{112}! \\ & - \log x_{121}! - \log(N_1 - x_{1(1)})! \\ & + x_{111} \log p_{111} + x_{121} \log p_{121} \\ & + x_{112} \log p_{112} \\ & + (N_1 - x_{1(1)}) \log p_{122} \\ & + \log N_2! - \log x_{211}! - \log x_{212}! \\ & - \log x_{221}! - \log (N_2 - x_{2(1)})! \\ & + x_{211} \log p_{211} + x_{221} \log p_{221} \end{aligned} \qquad (1)$$

$$+ x_{212} \log p_{212}$$

$$+ (N_2 - x_{2(1)}) \log p_{222} \quad ,$$

where

$$x_{1(1)} = x_{111} + x_{112} + x_{121}$$

$$x_{2(1)} = x_{211} + x_{212} + x_{221} \ .$$

Given the restrictions $N_1 = N_2 r$ and $\theta_1 = \theta_2 = \theta$, the maximum likelihood estimators of the critical unobservable parameters are

$$\hat{N}_2 = \frac{Kx_{2(1)} - x_{1(1)}}{K - r} \quad , \qquad (2)$$

$$\hat{N}_1 = r\hat{N}_2 \quad , \qquad (3)$$

$$\hat{x}_{222} = \frac{rx_{2(1)} - x_{1(1)}}{K - r} \qquad (4)$$

$$= \hat{N}_2 - x_{2(1)} \quad ,$$

$$\hat{x}_{122} = K \frac{rx_{2(1)} - x_{1(1)}}{K - r} \qquad (5)$$

$$= \hat{N}_1 - x_{1(1)} \quad ,$$

$$\hat{\theta} = \frac{rx_{2(1)} - x_{1(1)}}{\frac{x_{121}\,x_{112}}{x_{111}} - r\,\frac{x_{221}\,x_{212}}{x_{211}}} \tag{6}$$

$$= \frac{x_{111}\,\hat{x}_{122}}{x_{112}\,x_{121}}$$

$$= \frac{x_{211}\,\hat{x}_{222}}{x_{212}\,x_{221}} \quad ,$$

where

$$K = \frac{x_{112}\,x_{121}\,x_{211}}{x_{111}\,x_{212}\,x_{221}} \quad .$$

Second, we consider Model 2, where the corresponding restrictions are $N_1 = N_2 r$ and $\theta_2 = 1$. The maximum likelihood estimators are given by

$$\hat{N}_2 = \frac{x_{21+}\,x_{2+1}}{x_{211}} \quad , \tag{7}$$

$$\hat{N}_1 = r\hat{N}_2 \quad , \tag{8}$$

$$\hat{x}_{222} = \frac{x_{221}\,x_{212}}{x_{211}} \tag{9}$$

$$= \hat{N}_2 - x_{2(1)} \quad ,$$

$$\hat{x}_{122} = r\,\frac{x_{21+}\,x_{2+1}}{x_{211}} - x_{1(1)} \tag{10}$$

$$= \hat{N}_1 - x_{1(1)} \quad ,$$

$$\hat{\theta}_1 = \frac{x_{111}\,\hat{x}_{122}}{x_{121}\,x_{112}} \quad . \tag{11}$$

For both Models 1 and 2, application of standard Taylor series methods provides a way of determining the approximate covariance matrix of $(\hat{N}_1, \hat{N}_2)$.

3. Example

To illustrate the new methods, we consider data from the 1980 Post Enumeration Program (PEP), undertaken by the U.S. Bureau of the Census to estimate coverage error in the decennial census. Previous discussion of these data was presented in Wolter (1986). Also see Cowan and Bettin (1982) for documentation on the PEP.

The PEP produced the following data

		April 1980 CPS: in	April 1980 CPS: out	
1980 Census	in	$\tilde{x}_{i11}$	$\tilde{x}_{i12} = x_{i1+} - \tilde{x}_{i11}$	x_{i1+}
	out	$\tilde{x}_{i21} = \tilde{x}_{i+1} - \tilde{x}_{i11}$		
		$\tilde{x}_{i+1}$		

where the 1980 Census is List A and the April 1980 Current Population Survey (CPS) is List B. The census marginals x_{i1+} are the official 1980 census counts of the noninstitutional population reduced by an estimate of the number of duplicates encountered in the census. The estimators $\tilde{x}_{i11}$ of the matched cells and $\tilde{x}_{i+1}$ of the CPS marginals were prepared in accordance with the CPS sampling design. These statistics are survey estimators of the quantities x_{i11} and x_{i+1} that correspond to a situation in which List B is a 100%

enumeration of the conceptual CPS population. See Wolter (1986) for an extended discussion of the distinction between x's and $\tilde{x}$'s.

Table 1 presents the basic data disaggregated by age, race, and sex. Columns 4 and 6 sum to x_{i1+}, and the difference between x_{i1+} and column 7 is an estimate of census duplications and such. Note that $\tilde{x}_{i12}$ assumes negative values for blacks in the older age groups. This nonstandard behavior arises because of sampling variability in the CPS and because this cell is defined as the difference $x_{i1+} - \tilde{x}_{i11}$, where the x_{i1+} are not subject to sampling variability. In the results that follow, only $\tilde{x}_{i11}$, $\tilde{x}_{i+1}$ and x_{i1+} enter the estimation formulas for $\hat{N}_i$, and we calculate these formulas as is, with no special modification for the negative values of $\tilde{x}_{i12}$. For notational convenience, we henceforth omit the "~" where no confusion will result.

Table 2 presents sex ratios by age and race corresponding to the census year 1980. We take these ratios to be known, although in practice they are subject to some unknown degree of error. These ratios are derived by demographic accounting methods utilizing data sources essentially independent of the 1980 Census, including birth, death, and immigration statistics; historical census data; and data from sample surveys. For details see Passel and Robinson (1984).

Results of applying Model 1 to these data are given in Table 3. Clearly there are problems. The estimated cross-production ratio $\hat{\theta}$ behaves more erratically than it should and is occasionally outside of its parameter space $(0, \infty)$. Also, the estimated census capture probabilities $\hat{p}_{i1+}$ are unbelievably low in many cases. For example, consider black males age 20-24 where $\hat{p}_{11+} = .143$. Census coverage of this domain is known to be low, but there is strong evidence to suggest that many more than 14.3% of these individuals are counted. Although some of the problems with Model 1 may be ameliorated by smoothing the data across age groups prior to fitting the model, these results generally suggest to us that the model is misspecified.

Model 2 results appear in Tables 4 and 5 and seem much more reasonable. The erratic behavior of $\hat{\theta}$ in Table 3 is now dampened somewhat in Table 4 and the estimated capture probabilities are now more in agreement with our knowledge of the extent of census undercounting.

Unfortunately, some of the $\hat{\theta}$ continue to lie outside of the parameter space. We address this difficulty by modifying the Model 2 estimators in the following manner:

$$\hat{\theta} = \max\left\{0, \frac{\hat{x}_{122}\, x_{111}}{x_{121}\, x_{112}}\right\} \tag{12}$$

$$\hat{N}_1 = \max\{x_{1(1)}, \hat{N}_2\, r\} \quad . \tag{13}$$

All other estimators remain unchanged. This modification guarantees that $\hat{\theta}$ will be greater than or equal to zero and that $\hat{N}_1$ will at least equal the number of males actually observed (aside from sampling variability) in either the CPS or the census. Results of this modification are given in Table 5. Observe that $\hat{\theta} = 0$ for the young and the very old.

To further illustrate the new methods, Figures 1, 2, and 3 present plots of some of the results. The various sex ratios are depicted in Figure 1. Note the great discrepancy between the expected sex ratio on the one hand and the census and PEP sex ratios on the other. This discrepancy was our original motivation for developing Models 1 and 2. Model 2 attempts to preserve the expected sex ratio. But note that there remains a small discrepancy between the expected and Model 2 sex ratios for the very young and old, arising because of the modification in equation (13) which permits $\hat{N}_1 > r\,\hat{N}_2$. Some of the behavior of the sex ratios for the elderly may be attributed to age misreporting, known to be a significant problem.

Percent net undercount is defined by

$$\hat{U}_i = 100 \left(\frac{\hat{N}_i - \text{1980 Census Noninstitutional Population}}{\hat{N}_i} \right),$$

and three sets of $\hat{U}_i$ are pictured in Figures 2 and 3, including PEP, Model 2, and demographic analysis (refers to the same demographic accounting methods used to produce r). These data pertain to the black population and we have achieved similar results for nonblacks and for the total population. For males there is a great discrepancy between the PEP and demographic analysis results, particularly in the middle age groups, but not so for females. We feel that this fact provides some justification for the specification of Model 2.

Model 2 seems to eliminate much of the discrepancy for males, while for females Model 2 and the PEP are identical. High undercount rates in the middle age groups for black men are now predicted by both Model 2 and demographic analysis. Thus, in large measure, the problem of "correlation bias", first noticed in census undercount studies in the Census Bureau's 1950 Post Enumeration Survey, has been ameliorated by replacing the independence assumption by the sex ratio assumption.

It is of some interest to note the location of the "hump" in Figure 2. For demographic analysis, the hump occurs approximately at age 40, while in 1970 the corresponding hump occurred approximately 10 years earlier, say at age 30. This shift in the pattern of the age-specific undercount rates has been a subject of considerable debate among demographers and statisticians. While there is no widely accepted reason for the shift, it may signal some kind of cohort effect: either this particular cohort is less cooperative with decennial censuses, the independent data sources used in forming the demographic estimates are especially flawed for this cohort, or there are residual effects of the Viet Nam war on this cohort of males. The 1980 Model 2 estimates, however, return the hump to the vicinity of age 30, and in this sense follow more closely the pattern of age-specific undercounting seen in 1970 than in 1980.

For young black males, say age 0-15, the PEP estimates are higher than the demographic estimates and Model 2 exaggerates the difference. The demographic estimates, however, may be particularly accurate in this age range because of the completeness of the U.S. birth registration system in recent years. Since there is little reason to believe in a strong correlation bias for these young males, we may conclude that Model 2 should be restricted within this age range to the same independence assumption used for females.

Finally, we should stress that the real utility of Model 2 is not necessarily at the national level as studied here. Aside from the fact that illegal aliens are not included in the basic data sources, demographic analysis may provide a satisfactory view of the extent of undercounting at the U.S. level. But demographic analysis does not provide satisfactory estimates of the net undercount at subnational levels, principally because of a lack of satisfactory data on internal migration. Thus, the only information available on the extent of such undercounting comes from capture-recapture studies and it is here that Model 2 may be most useful. These studies have been reasonably accurate for females but less accurate than would be desirable for males because of the so-called "correlation bias". Improvements in accuracy at subnational levels can now be achieved, however, by incorporating knowledge of sex ratios.

4. Summary

The models studied in this paper illustrate a general method for circumventing the independence assumption, and thus for incorporating an allowance for list association in the estimation of population size. The method involves two features:

(i) consideration of two or more 2-way tables simultaneously, each representing a different domain of the population;

(ii) linking the parameters of the various tables.

Enough restrictions must be placed on the parameters so that the model is identifiable. In the present examples this was accomplished by assuming (1) the sex ratio is a known biological constant, and either (2) the degree of association for both males and females is identical or (3) the independence assumption holds only for females. Another potential application of the general method occurs when the age distribution of the population is considered known.

The Model 2 assumptions are thought to be reasonable in the study of human populations. Alternative assumptions may be more realistic in the study of wildlife populations.

REFERENCES

Passel, J.S. and Robinson, J.G. (1984), "Revised Estimates of the Coverage of the Population in the 1980 Census Based on Demographic Analysis," Proceedings of the Social Statistics Section, American Statistical Association, Washington, D.C.

Wolter, K.M. (1986), "Some Coverage Error Models for Census Data," Journal of the American Statistical Association, 81, 338-346.

Table 1. Data from the 1980 Post Enumeration Program

AGE	RACE	SEX	IN BOTH $\tilde{x}_{i11}$	IN CPS ONLY $\tilde{x}_{i21}$	IN CENSUS ONLY $\tilde{x}_{i12}$	1980 CENSUS NONINSTITUTIONAL POPULATION
0-4	NONBLACK	M	6512842	384170	310025	7130377
		F	5935424	340786	548527	6774369
	BLACK	M	1011733	196512	120759	1227002
		F	976904	226459	137317	1207726
5-9	NONBLACK	M	6849371	341181	161775	7275583
		F	6400716	298563	224148	6919777
	BLACK	M	1059613	175118	108599	1253382
		F	1080406	154802	71444	1234037
10-14	NONBLACK	M	7431362	302916	202399	7941700
		F	7013102	254091	291348	7581911
	BLACK	M	1167728	198777	86910	1335321
		F	1123208	155304	123080	1325434
15-19	NONBLACK	M	8261040	519041	452150	9174395
		F	7857926	525709	624466	8884746
	BLACK	M	1188923	160998	156269	1450091
		F	1235125	149533	154093	1487858
20-24	NONBLACK	M	7478714	734723	1165140	9252032
		F	7814541	642151	846266	9206680
	BLACK	M	849668	140210	284036	1237004
		F	1002533	153627	306185	1417551
25-29	NONBLACK	M	7221914	458359	841425	8535184
		F	7642141	390646	503372	8557193
	BLACK	M	721528	113707	191769	1028997
		F	987598	139234	148400	1230185
30-34	NONBLACK	M	6838363	365037	567179	7742192
		F	7056332	276446	461994	7846505
	BLACK	M	615040	92438	147697	834499
		F	909946	123859	37476	1012900

Table 1 (Continued)--Page 2

AGE	RACE	SEX	IN BOTH $\tilde{x}_{i11}$	IN CPS ONLY $\tilde{x}_{i21}$	IN CENSUS ONLY $\tilde{x}_{i12}$	1980 CENSUS NONINSTITUTIONAL POPULATION
35-39	NONBLACK	M	5481687	283242	403762	6155104
		F	5839157	224278	220029	6291367
	BLACK	M	481085	84605	116113	643503
		F	624304	65589	125575	792325
40-44	NONBLACK	M	4624611	184771	287863	5108676
		F	4825161	194404	256262	5261917
	BLACK	M	435197	44972	87998	554897
		F	589229	61543	50955	682054
45-49	NONBLACK	M	4446400	152925	199614	4842492
		F	4789237	138489	97907	5057107
	BLACK	M	381677	49562	85285	506413
		F	531796	51153	49701	624874
50-54	NONBLACK	M	4665765	152797	216666	5082882
		F	5033803	113185	201874	5442785
	BLACK	M	420627	37974	37969	496445
		F	568647	39433	16009	621016
55-59	NONBLACK	M	4648511	111835	125948	4977214
		F	5150707	122547	171391	5533852
	BLACK	M	369676	55628	67473	459114
		F	503404	34963	29077	566422
60-64	NONBLACK	M	3848124	121230	237568	4243719
		F	4495213	114049	190423	4892527
	BLACK	M	305087	35255	42171	378690
		F	471257	36037	-19804	481068
65-69	NONBLACK	M	3236111	101812	130854	3522115
		F	4128981	103704	75483	4377901
	BLACK	M	300352	15826	8664	324635
		F	428569	23853	-19080	438865

Table 1 (Continued)--Page 3

AGE	RACE	SEX	IN BOTH $\tilde{x}_{i11}$	IN CPS ONLY $\tilde{x}_{i21}$	IN CENSUS ONLY $\tilde{x}_{i12}$	1980 CENSUS NONINSTITUTIONAL POPULATION
70-74	NONBLACK	M	2351004	69511	107575	2562421
		F	3278376	86371	114390	3522156
	BLACK	M	243692	15782	-26938	227787
		F	303404	18141	-4250	321037
75+	NONBLACK	M	2820695	69534	78405	3028753
		F	4872110	191982	102862	5177135
	BLACK	M	251914	12156	-11821	263905
		F	404358	29011	-3592	430334

Table 2. Expected Sex Ratios by Age and Race for 1980

AGE	RACE	SEX RATIO (MALES/FEMALES) r
0-4	NONBLACK	1.054
	BLACK	1.023
5-9	NONBLACK	1.053
	BLACK	1.022
10-14	NONBLACK	1.048
	BLACK	1.012
15-19	NONBLACK	1.044
	BLACK	1.000
20-24	NONBLACK	1.036
	BLACK	.976
25-29	NONBLACK	1.030
	BLACK	.971
30-34	NONBLACK	1.017
	BLACK	.965
35-39	NONBLACK	1.007
	BLACK	.960
40-44	NONBLACK	1.002
	BLACK	.957
45-49	NONBLACK	.987
	BLACK	.957
50-54	NONBLACK	.962
	BLACK	.940
55-59	NONBLACK	.931
	BLACK	.901
60-64	NONBLACK	.888
	BLACK	.846
65-69	NONBLACK	.815
	BLACK	.772
70-74	NONBLACK	.725
	BLACK	.715
75+	NONBLACK	.539
	BLACK	.568

Table 3. Results of Applying Model 1 to the 1980 Post Enumeration Program

AGE	RACE	SEX	$\hat{N}_i$	$\hat{p}_{i+1}$	$\hat{p}_{i1+}$	$\hat{\theta}$
0-4	NONBLACK	M	7223922	.955	.944	.923
		F	6853816	.916	.946	.923
	BLACK	M	1219533	.991	.929	-4.667
		F	1192115	1.009	.935	-4.667
5-9	NONBLACK	M	7521505	.956	.932	20.994
		F	7142930	.938	.927	20.994
	BLACK	M	1324314	.932	.882	-1.060
		F	1295806	.953	.889	-1.060
10-14	NONBLACK	M	7981638	.969	.956	5.450
		F	7616067	.954	.959	5.450
	BLACK	M	1666703	.820	.753	14.417
		F	1646940	.776	.757	14.417
15-19	NONBLACK	M	8910501	.985	.978	-11.325
		F	8534962	.982	.994	-11.325
	BLACK	M	1781189	.758	.755	12.995
		F	1781189	.777	.780	12.995
20-24	NONBLACK	M	10078220	.815	.858	6.112
		F	9728008	.869	.890	6.112
	BLACK	M	7943854	.125	.143	142.305
		F	8139195	.142	.161	142.305
25-29	NONBLACK	M	9058792	.848	.890	10.057
		F	8794944	.913	.926	10.057
	BLACK	M	1671463	.500	.546	21.325
		F	1721383	.655	.660	21.325
30-34	NONBLACK	M	8170305	.882	.906	13.203
		F	8033732	.913	.936	13.203
	BLACK	M	1084680	.652	.703	10.339
		F	1124021	.920	.843	10.339
35-39	NONBLACK	M	6436827	.896	.914	12.852
		F	6392082	.949	.948	12.852
	BLACK	M	947877	.597	.630	13.030
		F	987372	.699	.759	13.030

Table 3 (Continued)--Page 2

AGE	RACE	SEX	$\hat{N}_i$	$\hat{p}_{i+1}$	$\hat{p}_{i1+}$	$\hat{\theta}$
40-44	NONBLACK	M	6979159	.689	.704	163.627
		F	6965228	.721	.730	163.627
	BLACK	M	803187	.598	.651	25.845
		F	839276	.775	.763	25.845
45-49	NONBLACK	M	5071059	.907	.916	39.637
		F	5137851	.959	.951	39.637
	BLACK	M	668041	.646	.699	13.682
		F	698057	.835	.833	13.682
50-54	NONBLACK	M	5322229	.905	.917	40.448
		F	5532463	.930	.946	40.448
	BLACK	M	626067	.733	.733	37.778
		F	666028	.913	.878	37.778
55-59	NONBLACK	M	4164007	1.143	1.147	-238.372
		F	4472618	1.179	1.190	-238.372
	BLACK	M	515304	.825	.848	2.219
		F	571925	.941	.931	2.219
60-64	NONBLACK	M	4336260	.915	.942	17.281
		F	4883176	.944	.960	17.281
	BLACK	M	406191	.838	.855	4.859
		F	480132	1.057	.940	4.859
65-69	NONBLACK	M	3536619	.944	.952	16.479
		F	4339410	.975	.969	16.479
	BLACK	M	328311	.963	.941	7.598
		F	425273	1.064	.963	7.598
70-74	NONBLACK	M	2509833	.964	.980	-5.740
		F	3461838	.972	.980	-5.740
	BLACK	M	226207	1.147	.958	3.628
		F	316373	1.016	.946	3.628
75+	NONBLACK	M	4377802	.660	.662	729.084
		F	8122081	.623	.613	729.084
	BLACK	M	241305	1.094	.995	19.186
		F	424833	1.020	.943	19.186

Table 4. Results of Applying Model 2 to the 1980 Post Enumeration Program

AGE	RACE	SEX	$\hat{N}_i$	$\hat{p}_{i+1}$	$\hat{p}_{i1+}$	$\hat{\theta}$
0-4	NONBLACK	M	7226468	.954	.944	1.063
		F	6856231	.915	.946	1
	BLACK	M	1404080	.861	.807	3.201
		F	1372512	.877	.812	1
5-9	NONBLACK	M	7301378	.985	.960	-6.323
		F	6933883	.966	.955	1
	BLACK	M	1345860	.917	.868	.141
		F	1316889	.938	.875	1
10-14	NONBLACK	M	7932414	.975	.962	-.517
		F	7569097	.960	.965	1
	BLACK	M	1435633	.952	.874	-1.202
		F	1418610	.901	.879	1
15-19	NONBLACK	M	9448074	.929	.922	7.598
		F	9049879	.926	.937	1
	BLACK	M	1557407	.867	.864	2.420
		F	1557407	.889	.892	1
20-24	NONBLACK	M	9709909	.846	.890	2.895
		F	9372499	.902	.924	1
	BLACK	M	1473042	.672	.770	4.248
		F	1509264	.766	.867	1
24-29	NONBLACK	M	8818746	.871	.914	5.562
		F	8561890	.938	.951	1
	BLACK	M	1258565	.664	.726	7.662
		F	1296154	.869	.876	1
30-34	NONBLACK	M	7945690	.907	.932	5.784
		F	7812872	.939	.962	1
	BLACK	M	1038709	.681	.734	8.268
		F	1076382	.960	.880	1
35-39	NONBLACK	M	6335958	.910	.929	8.018
		F	6291915	.964	.963	1
	BLACK	M	795514	.711	.751	5.569
		F	828661	.833	.905	1
40-44	NONBLACK	M	5296724	.908	.927	17.344
		F	5286152	.950	.961	1
	BLACK	M	676646	.710	.773	11.929
		F	707049	.920	.905	1

Table 4 (continued) Page 2

AGE	RACE	SEX	$\hat{N}_i$	$\hat{p}_{i+1}$	$\hat{p}_{i1+}$	$\hat{\theta}$
45-49	NONBLACK	M	4963094	.927	.936	23.911
		F	5028464	.980	.972	1
	BLACK	M	610021	.707	.765	8.443
		F	637431	.915	.912	1
50-54	NONBLACK	M	5149972	.936	.948	16.171
		F	5353401	.961	.978	1
	BLACK	M	587687	.780	.780	26.582
		F	625199	.973	.935	1
55-59	NONBLACK	M	5072761	.938	.941	61.539
		F	5448723	.968	.977	1
	BLACK	M	513087	.829	.852	2.000
		F	569464	.945	.935	1
60-64	NONBLACK	M	4266411	.930	.958	7.949
		F	4804517	.959	.975	1
	BLACK	M	411135	.828	.845	5.873
		F	485976	1.044	.929	1
65-69	NONBLACK	M	3512702	.950	.959	10.670
		F	4310064	.982	.975	1
	BLACK	M	333720	.947	.926	19.448
		F	432280	1.047	.947	1
70-74	NONBLACK	M	2524560	.959	.974	-1.110
		F	3842151	.966	.974	1
	BLACK	M	226684	1.145	.956	3.354
		F	317041	1.014	.944	1
75+	NONBLACK	M	2787173	1.037	1.040	-93.886
		F	5171007	.979	.962	1
	BLACK	M	243967	1.082	.984	14.519
		F	429519	1.009	.933	1

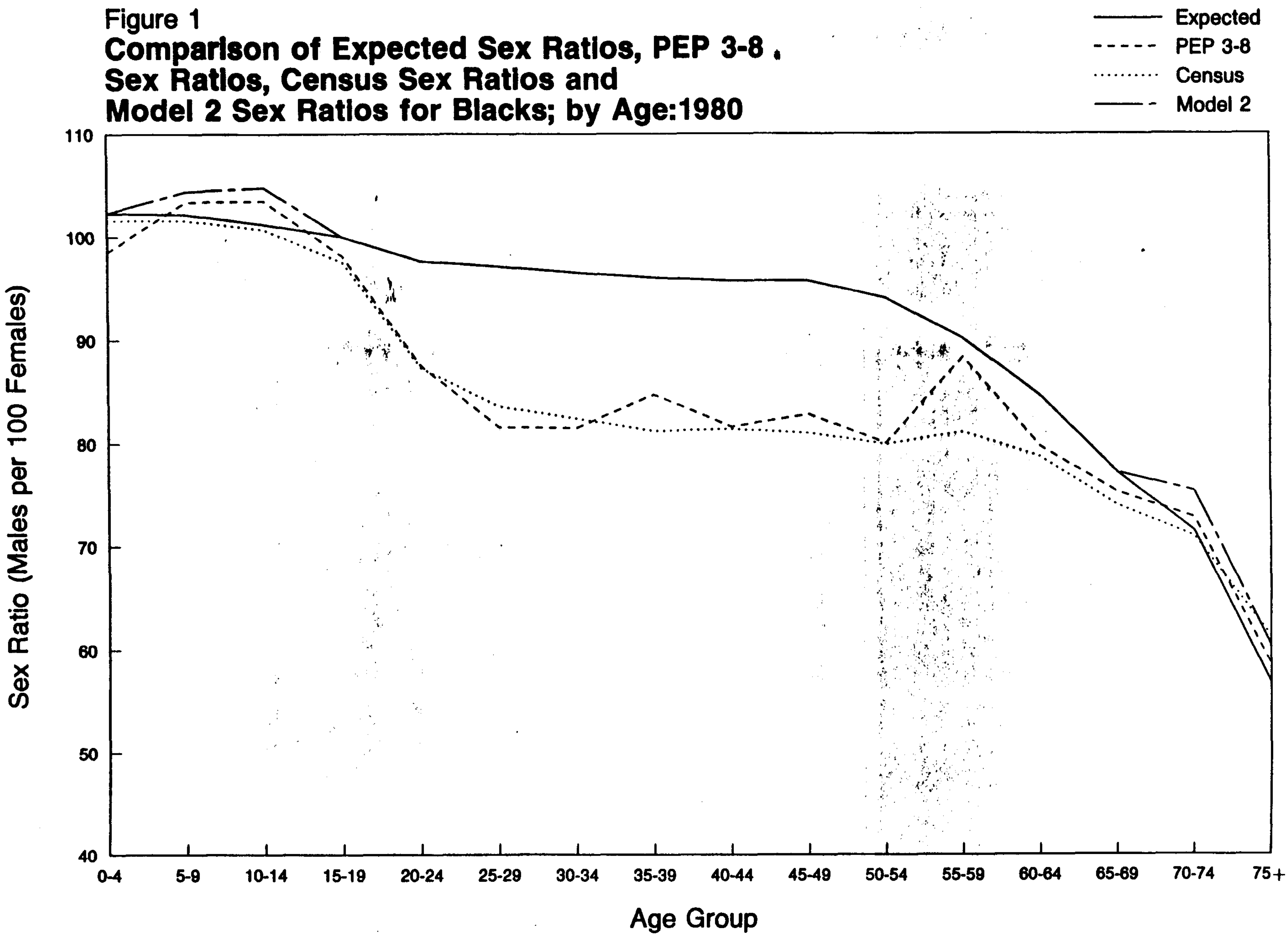
Table 5. Results of Applying Model 2 to the 1980 Post Enumeration Program, Modified Estimation Procedure

AGE	RACE	SEX	$\hat{N}_i$	$\hat{p}_{i+1}$	$\hat{p}_{i1+}$	$\hat{\theta}$
0-4	NONBLACK	M	7296051	.945	.935	1.063
		F	6856231	.915	.946	1
	BLACK	M	1404080	.861	.807	3.201
		F	1372512	.877	.812	1
5-9	NONBLACK	M	7438466	.967	.943	0
		F	6933883	.966	.955	1
	BLACK	M	1375206	.898	.849	.141
		F	1316889	.938	.875	1
10-14	NONBLACK	M	8027117	.964	.951	0
		F	7569097	.960	.965	1
	BLACK	M	1486796	.919	.844	0
		F	1418610	.901	.879	1
15-19	NONBLACK	M	9448074	.929	.922	7.598
		F	9049879	.926	.937	1
	BLACK	M	1557407	.867	.864	2.420
		F	1557407	.889	.892	1
20-24	NONBLACK	M	9709909	.846	.890	2.895
		F	9372499	.902	.924	1
	BLACK	M	1473042	.672	.770	4.248
		F	1509264	.766	.867	1
25-29	NONBLACK	M	8818746	.871	.914	5.562
		F	8561890	.938	.951	1
	BLACK	M	1258565	.664	.726	7.662
		F	1296154	.869	.876	1
30-34	NONBLACK	M	7945690	.907	.932	5.784
		F	7812872	.939	.962	1
	BLACK	M	1038709	.681	.734	8.268
		F	1076382	.960	.880	1
35-39	NONBLACK	M	6335958	.910	.929	8.018
		F	6291915	.964	.963	1
	BLACK	M	795514	.711	.751	5.569
		F	828661	.833	.905	1

Table 5 (continued) Page 2

AGE	RACE	SEX	$\hat{N}_i$	$\hat{p}_{i+1}$	$\hat{p}_{i1+}$	$\hat{\theta}$
40-44	NONBLACK	M	5296724	.908	.927	17.344
		F	5286152	.950	.961	1
	BLACK	M	676646	.710	.773	11.929
		F	707049	.920	.905	1
45-49	NONBLACK	M	4963094	.927	.936	23.911
		F	5028464	.980	.972	1
	BLACK	M	610021	.707	.765	8.443
		F	637431	.915	.912	1
50-54	NONBLACK	M	5149972	.936	.948	16.171
		F	5353401	.961	.978	1
	BLACK	M	587687	.780	.780	26.582
		F	625199	.973	.935	1
55-59	NONBLACK	M	5072761	.938	.941	61.539
		F	5448723	.968	.977	1
	BLACK	M	513087	.829	.852	2.000
		F	569464	.945	.935	1
60-64	NONBLACK	M	4266411	.930	.958	7.949
		F	4804517	.959	.975	1
	BLACK	M	411135	.828	.845	5.873
		F	485976	1.044	.929	1
65-69	NONBLACK	M	3514908	.950	.958	10.670
		F	4310064	.982	.975	1
	BLACK	M	333720	.947	.926	19.448
		F	432280	1.047	.947	1
70-74	NONBLACK	M	2562551	.945	.959	0
		F	3482151	.966	.974	1
	BLACK	M	238999	1.086	.907	0
		F	317041	1.014	.944	1
75+	NONBLACK	M	3009894	.960	.963	0
		F	5171007	.979	.962	1
	BLACK	M	259456	1.018	.925	0
		F	429519	1.009	.933	1

Figure 1
Comparison of Expected Sex Ratios, PEP 3-8 Sex Ratios, Census Sex Ratios and Model 2 Sex Ratios for Blacks; by Age:1980



CENGRAF [UPF3 GRANID: BIN816 PROJID: SRD ACCTNO: 4205441 DATE: 071486 TIME: 162707 MACHINE: A REEL: GXSPOL FILE: CBOSOUT84566 FRAME: 1

Figure 2

Percent Undercount for Black Males: 1980 Demographic Analysis, 1980 PEP 3-8 and Model 2

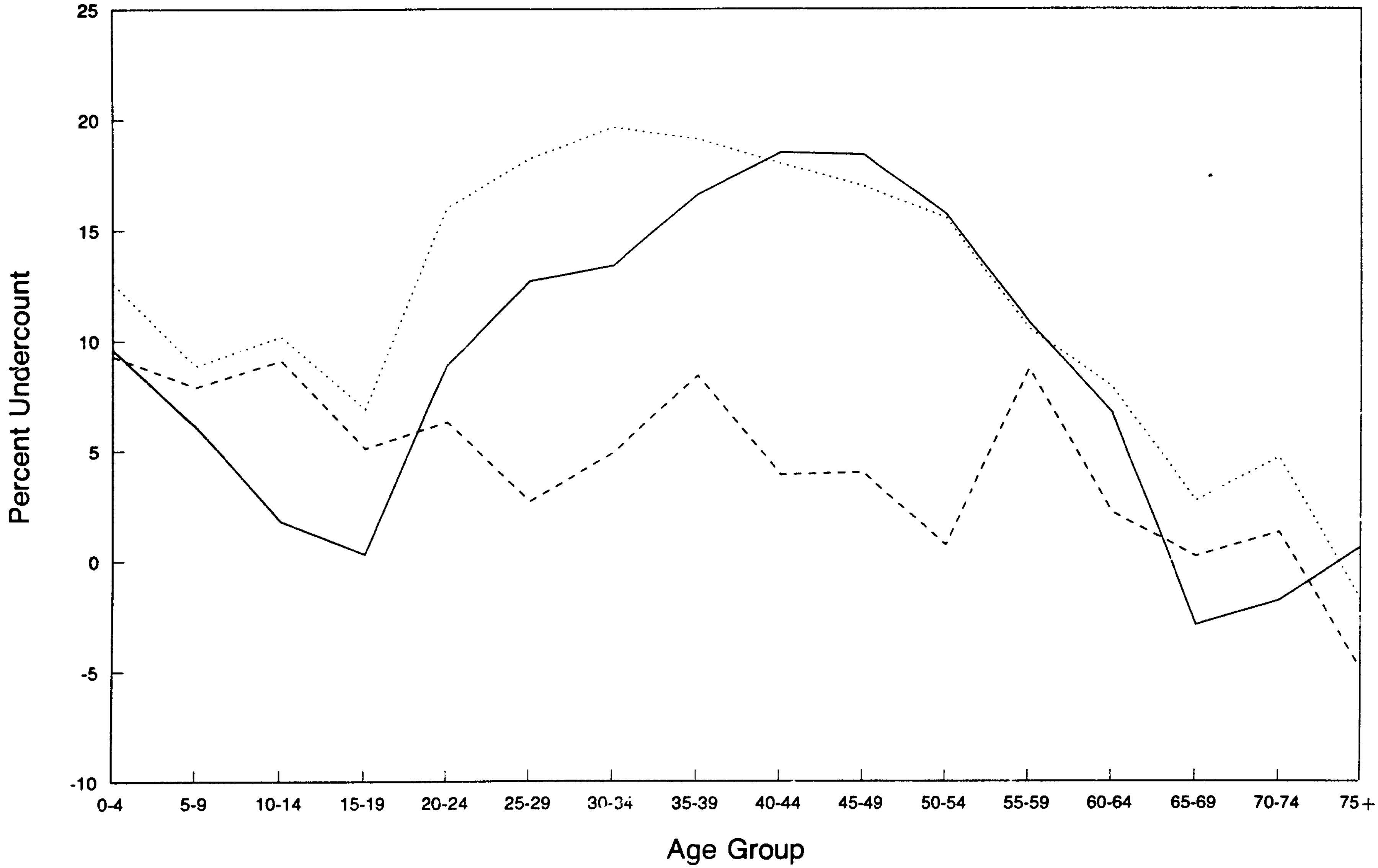


Figure 3
Percent Undercount for Black Females: 1980 Demographic Analysis, 1980 PEP 3-8 and Model 2