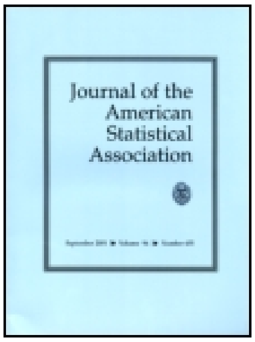

Journal of the American Statistical Association

ISSN: 0162-1459 (Print) 1537-274X (Online) Journal homepage: http://www.tandfonline.com/loi/uasa20

A Bayesian Approach to 2000 Census Evaluation Using ACE Survey Data and Demographic Analysis

Michael R Elliott & Roderick J. A Little

To cite this article: Michael R Elliott & Roderick J. A Little (2005) A Bayesian Approach to 2000 Census Evaluation Using ACE Survey Data and Demographic Analysis, Journal of the American Statistical Association, 100:470, 380-388, DOI: 10.1198/016214504000001826

To link to this article: https://doi.org/10.1198/016214504000001826

Published online: 01 Jan 2012.

Submit your article to this journal

Article views: 63

View related articles

Citing articles: 4 View citing articles

Full Terms & Conditions of access and use can be found at
http://www.tandfonline.com/action/journalInformation?journalCode=uasa20

Download by: [Universita degli Studi di Torino] **Date:** 09 January 2018, At: 01:46

A Bayesian Approach to 2000 Census Evaluation Using ACE Survey Data and Demographic Analysis

Michael R. ELLIOTT and Roderick J. A. LITTLE

Demographic analysis of data on births, deaths, and migration, together with coverage measurement surveys that use capture-recapture methods, have established that U.S. Census counts are flawed for certain subpopulations. Previous work using 1990 Census data in African–Americans age 30–49 proposed a hierarchical Bayesian model that assembled Census, follow-up survey, and demographic data, providing a principled solution to the problem of negative estimated counts in some subpopulations, smoothing highly variable estimates across subpopulations, and providing estimates of precision that incorporate uncertainty in the demographic analysis estimates. This article extends that effort by refining the hierarchical model design, expanding the set of models considered, considering the presence of bias in the Census or follow-up survey counts, obtaining Bayes factors for use in model selection, and applying the methods to the entire 2000 U.S. Census. Comparisons with the 1990 U.S. Census results are included as well.

KEY WORDS: Bayes factors; Capture-recapture; Gibbs sampling; Postenumeration survey.

Downloaded by [Universita degli Studi di Torino] at 01:46 09 January 2018

1. INTRODUCTION

Despite enormous efforts, the U.S. Census missed enumerating some U.S. residents and counted others twice or more as part of Census 2000. This gross undercount subtracted from the gross overcount yields the net undercount or overcount, which can be considered at a national or subnational level (i.e., by geographic region, gender, or race/ethnicity). Evaluations of previous Censuses have found that older Caucasians tend to be somewhat overcounted, whereas young and middle-aged minority adults tend to be substantially undercounted (U.S. Bureau of the Census 1988; Hogan 1993).

To evaluate the degree of net overcount/undercount, capture-recapture techniques similar to those used in wildlife population estimation (Seber 1982) can be used. These techniques typically rely on behavioral assumptions—namely, independence of capture and recapture—that are untenable in human populations. However, in human populations demographers can supply alternative sources of population data that allow Census evaluation under less restrictive assumptions (Wolter 1990; Bell 1993; Bell et al. 1996). Incorporating demographic analysis (DA) data into Census evaluation poses numerous difficulties. DA data relies in part on immigration and emigration estimates and consequently is available at only the national level, but evaluation is desired at various small-area and subpopulation estimates. Various assumptions can be made that yield different estimates of undercount at subpopulation level yet are difficult to distinguish empirically; in addition, previous techniques have yielded unstable subpopulation estimates and do not incorporate uncertainty in demographic analysis data.

The issue of evaluating national census estimates, both in the United States and elsewhere, has been an active area of research for at least 20 years (Bell 1993; Choi, Steel, and Skinner 1988; Datta et al. 1992; Elliott and Little 2000; Isaki, Huang, and Tsay 1991; Isaki and Schultz 1986; Isaki, Tsay, and Fuller 2000). This article extends a Bayesian hierarchical model, first outlined by Elliott and Little (2000) and applied to a small subset of 1990 U.S. Census data (African–Americans age 30–49), that incorporates capture-recapture and demographic data to estimate Census overcount/undercount at subnational levels. Our methods (1) use small-area smoothing techniques (Ghosh and Meeden 1986; Ghosh, Natarajan, Stroud, and Carlin 1998) to shrink outlying estimates of Census overcount/undercount, (2) avoid spurious "negative cell counts" in overcount/undercount estimation, (3) use regression techniques to estimate common main effects of subnational covariates, (4) provide inference about overcount/undercount estimates, (5) provide simple techniques for model checking, (6) allow some degree of model selection across a set of plausible behavioral models, and (7) allow for incorporation of uncertainty in DA data. Section 2 describes the Census, follow-up survey, and demographic data and reviews previous approaches to combining them. Section 3 recasts these approaches in a Bayesian framework, with particular attention to three models that make different assumptions about human behavior. Section 4 applies our methodology to the 2000 U.S. Census. Section 5 discusses some implications of our results.

2. BACKGROUND: CAPTURE–RECAPTURE METHODS IN THE 2000 CENSUS

Since 1970, the U.S. Census has used capture-recapture methods to estimate net overcount/undercount in the U.S. Census by supplementing the Census enumeration with a coverage measurement survey, in 2000 termed the *Accuracy and Coverage Evaluation* (ACE) Survey. The ACE began by drawing a probability sample of Census clusters, stratified by size and American Indian reservation status and sorted by race/ethnicity and tenure (owner/renter) makeup within stratum. A detailed independent enumeration of households in these sampled clusters was then conducted in April–August 2000, immediately after the actual Census. To deal with the population that moved between Census day (April 1, 2000) and the ACE enumeration day, this enumeration attempted to ascertain the status of the household as of Census day (U.S. Bureau of the Census 2000).

Michael Elliott is Assistant Professor, Center for Clinical Epidemiology and Biostatistics, Department of Biostatistics and Epidemiology, University of Pennsylvania Medical Center, Philadelphia, PA 19104 (E-mail: *melliott@cceb.upenn.edu*). Roderick Little is Richard D. Remington Collegiate Professor of Biostatistics, Department of Biostatistics, University of Michigan, Ann Arbor, MI 48109 (E-mail: *rlittle@umich.edu*). This research was supported by Bureau of the Census Assignment Agreement and by Census Contract No. 50-YABC-8-00004 to NORC, subcontract to Datametrics Research, Inc. The authors thank Bruce Spencer of Northwestern University, Lawrence Brown of the University of Pennsylvania, and Rita Petroni, William Bell, Katie Bench, and Eric Schindler at the U.S. Census Bureau for their invaluable assistance. The authors also thank the editor and associate editor, as well as two anonymous reviewers, for their excellent suggestions that improved the article.

© 2005 American Statistical Association
Journal of the American Statistical Association
June 2005, Vol. 100, No. 470, Applications and Case Studies
DOI 10.1198/016214504000001826

Downloaded by [Universita degli Studi di Torino] at 01:46 09 January 2018

To combine the Census and ACE data, individuals in the Census listing in the ACE clusters—the "E-sample"—who were imputed or for whom insufficient information existed to match with the ACE were removed from the Census total. In addition, Census and ACE listings were compared, and persons appearing only on the Census list were rechecked to determine whether they were "erroneous enumerations." The E-sample total was then deflated by the fraction estimated to be erroneously enumerated. Similarly, persons in the ACE—the "P-sample"—were cross-checked against the E-sample and assigned to either an In-Census or Out-Census category. This matching process compared names and person characteristics from the P- and E-samples, ranked pairs by probability of match, and assigned the P-sample cases to match (In-Census), nonmatch (Out-Census), or possible match. Follow-up P-sample interviews were conducted with selected possible match cases (e.g., partial households matches) to attempt to more accurately assign match and nonmatch status. After inflating the counts in the sampled clusters by the inverse of their probability of selection, a 2×2 table (In–Out-Census; In–Out-ACE) can be formed (Table 1; see U.S. Bureau of the Census 2000 for more details). Note that this description should be viewed as an approximation to the actual procedure used to generate the summary data in Table 1, the details of which are beyond the scope of this article.

Because of a large degree of duplication in the master address file (MAF) containing the housing unit addresses used for the decennial Census, as well as a failure to detect nonresident household members in the P-sample, the original ACE results were deemed to have an unacceptably high level of error to be used for directly assessing undercount (U.S. Bureau of the Census 2001). Thus the results presented in this article are from the ACE Revision II (U.S. Bureau of the Census 2003), which attempted to remove previously undetected MAF duplicates and to determine other erroneous enumerations, such as adult children being counted at both a dormitory and a household residence, by using a nationwide rather than a local-block search for duplicated records.

If we estimate the In-Census In-ACE cell in Table 1 by $\hat{\psi}_{11} = y_{11}$ and the Out-Census In-ACE cell by $\hat{\psi}_{01} = y_{01}$, then the natural estimate of the In-Census Out-ACE cell is $\hat{\psi}_{10} = z_{1.} - y_{11}$. To estimate the Out-Census Out-ACE cell, however, additional assumptions are needed. The most straightforward assumption is independence of "capture" and "recapture." The odds of being missed in the Census for those not in the ACE can be estimated by the odds of being missed in the Census for those in the ACE. After some simple algebra, this yields $\hat{\psi}^{I}_{..} = \hat{\psi}_{11} + \hat{\psi}_{10} + \hat{\psi}_{01} + \hat{\psi}_{00} = z_{1.}y_{.1}/y_{11}$. As Sekar and Deming (1949) pointed out, if the probability of capture in the ACE conditional on being captured in the Census is unequal to the probability of capture in the ACE conditional on being missed in the Census (i.e., if $\psi_{11}/\psi_{1.} \neq \psi_{01}/\psi_{0.}$), then the independence estimate of the Out-Census Out-ACE cell given by $\hat{\psi}_{00} = \hat{\psi}_{10}\hat{\psi}_{01}/\hat{\psi}_{11}$ consistently overestimates or underestimates the true cell count ψ_{00} by the factor $\frac{\psi_{10}\psi_{01}}{\psi_{11}\psi_{00}}$. In particular, if the probability of capture in the ACE conditional on being captured in the Census is greater that the probability of capture in the ACE conditional on being missed in the Census, then the independence estimate underestimates the Out-Census Out-ACE cell. However, as Sekar and Deming (1949, p. 106) noted, "correlation signifies heterogeneity in the population for it implies that events that fail to be detected do not form a random sample from the whole population of events." In particular, individual-level independence of capture and recapture is not sufficient for the independence estimate of the total population $\hat{\psi}^{I}_{..}$ to consistently estimate $\psi_{..}$. Consider a population of N individuals consisting of G latent groups, $2 \leq G \leq N$, where independence of capture and recapture holds within each group. Denoting the population proportion of the groups by p_g and the capture and recapture proportions by $\pi_{1.g}$ and $\pi_{.1g}$, the marginal probabilities of capture, recapture, and capture-recapture in the population are given by $\sum_g p_g\pi_{1.g} = \psi_{1.}/\psi_{..}$, $\sum_g p_g\pi_{.1g} = \psi_{.1}/\psi_{..}$, and $\sum_g p_g\pi_{1.g}\pi_{.1g} = \psi_{11}/\psi_{..}$. Then the independence model will hold if and only if

$$\sum_g p_g\pi_{1.g} \sum_g p_g\pi_{.1g} = \sum_g p_g\pi_{1.g}\pi_{.1g}. \tag{1}$$

Equality of probability of capture or equality of the probability of recapture for all subgroups is clearly a sufficient condition for (1) to hold. In the case where $G = 2$, equality of probability of capture or probability of recapture for both subgroups is also necessary; $(p_1\pi_{1.1} + p_2\pi_{1.2})(p_1\pi_{.11} + p_2\pi_{.12}) = p_1\pi_{1.1}\pi_{.11} + p_2\pi_{1.2}\pi_{.12}$ implies that $(p_1^2 - p_1)(\pi_{1.1} - \pi_{1.2})(\pi_{.11} - \pi_{.12}) = 0$, which of course implies that $\pi_{1.1} = \pi_{1.2}$ or $\pi_{.11} = \pi_{.12}$. Otherwise, (1) will hold only under special circumstances, such as when the group sizes are equal and distribution of the recapture probabilities is the same across a set of distinct capture probabilities, for example, if $G = 4$, and $p_g = 1/4$ for all g, and $\pi_{1.1} = \pi_{1.2}$, $\pi_{1.3} = \pi_{1.4}$, $\pi_{.11} = \pi_{.13}$, and $\pi_{.12} = \pi_{.14}$.

The U.S. Census Bureau attempts to ensure independence of capture and recapture through independent enumeration of households and residents during the ACE; however, it can do little about heterogeneity in the capture probabilities, which is due to persons having different tendencies to respond to the Census and ACE survey attempts. To deal with this problem, Sekar and Deming (1949) suggested poststratifying the 2×2 table into P poststrata with similar capture-recapture profiles. The total population in each poststratum is then estimated under the assumption of independence and summed across the P poststrata to obtain estimates for the total population. This is the standard undercount adjustment made by the Census and is called *dual system estimation* (DSE).

In the 2000 Census, these poststrata were based on classification by gender, age, race/ethnicity, region of residence, urbanicity of residence, tenure (owner/renter status), and whether

Table 1. Observed Data and Associated Underlying Population Parameters From the Census and ACE Survey

		Observed data ACE			*Underlying parameters ACE*		
		In	*Out*		*In*	*Out*	
Census	In	y_{11}		$z_{1.}$	ψ_{11}	ψ_{10}	$\psi_{1.}$
	Out	y_{01}			ψ_{01}	ψ_{00}	$\psi_{0.}$
		$y_{.1}$			$\psi_{.1}$	$\psi_{.0}$	$\psi_{..}$

NOTE: y_{11} and y_{01} are estimated counts of individuals in and out of the Census on the basis of the ACE interviewing and follow-up (estimated from the P-sample); $z_{1.}$ is the Census count, minus imputations and an estimate of erroneous enumerations (estimated from the E-sample); and ψ_{ij} is the population that would reside in the ijth cell if the ACE had been a complete census.

Downloaded by [Universita degli Studi di Torino] at 01:46 09 January 2018

or not the subject resided in a "low return" block, where "low return" rate is defined by the 25th percentile nationally for block return rates of Census mailings among subjects in a race/ethnicity-by-tenure domain. A total of 448 poststrata were created (U.S. Bureau of the Census 2000). However, because the DA data models implicitly rely on different behavioral assumptions for males and females, an assumption that is probably untenable for those under age 18, this analysis is restricted to the 384 poststrata for those age 18 and older. We modify the notation in Table 1 to accommodate this stratification by superscripting the cell counts by gender, $S = M, F$, and subscripting by age-by-region-by-tenure-by-urbanicity-by-return-rate index $k = 1, \dots, 196$.

2.1 Incorporating Data From Demographic Analysis

The plausibility of the assumption of zero correlation bias within poststrata can be assessed by comparing the sex ratios (i.e., ratio of number of men to number of women) in subpopulations obtained under the DSE model with estimates from DA. The evidence for the 2000 Census is that correlation bias exists, because the DSE model yields a pattern of implausibly low values of the male–female sex ratio for some subpopulations. For example, in the 2000 Census, the DA estimate of the sex ratio among the noninstitutionalized African–American population age 30–49 was .89, compared with the unadjusted Census sex ratio estimate of .80 and the DSE-adjusted estimate of .81. To correct for this apparent undercount of men relative to women, Wolter (1990) suggested attributing the low observed male–female ratio to an undercount of males. Specifically, he assumed that the ratio of the odds of enumeration in the Census for those included in the follow-up survey to the odds of enumeration in the Census for those not in the follow-up was 1 for all female poststrata and an unknown parameter $\theta^M \in (0, \infty)$ for all male poststrata. Under this model, the estimated number of males in the population is increased to match the overall male–female ratio ρ (assumed known from DA) subject to this constraint; that is, the total population $\hat{\psi}^{DA}_{\dots}$ is estimated by $\hat{\psi}^M_{\dots} + \hat{\psi}^F_{\dots}$, where $\hat{\psi}^M_{\dots} = \rho\hat{\psi}^F_{\dots}$ and $\hat{\psi}^F_{\dots}$ is estimated under the independence model. Wolter (1990) showed that the maximum likelihood estimate of θ under a multinomial model in a single stratum is given by $\hat{\theta} = 1 + (\hat{\psi}^{DA}_{..} - \hat{\psi}^I_{..})/\hat{\psi}^I_{00}$, where $\hat{\psi}^I_{..}$ is the total population estimated under the independence model. Bell (1993) extended this idea to the stratified case by assuming that

$$\text{FOR:} \quad \theta^M_k = \frac{\psi^M_{k11}/\psi^M_{k10}}{\psi^M_{k01}/\psi^M_{k00}} = \theta^M \qquad \text{and}$$
$$\theta^F_k = \theta^F = 1 \quad \text{for all } k. \tag{2}$$

We term this the *fixed odds ratio* (FOR) model. Under (1), Bell estimated the total population size within each poststratum by $\hat{\psi}^\theta_{k..} = \hat{\psi}^I_{k..} + (\hat{\theta} - 1)\hat{\psi}^I_{k00} = \hat{\psi}^I_{k..} + (\hat{\psi}^{DA}_{\dots} - \hat{\psi}^I_{\dots})(\hat{\psi}^I_{k00}/\hat{\psi}^I_{.00})$.

Wolter and Bell's approach can be applied to alternative models of behavior. We consider two others. The *fixed relative-risk* (FRR) model (Bell 1993) assumes a constant relative risk for enumeration in ACE among those captured relative to those missed in the Census for males and independence for females,

$$\text{FRR:} \quad \gamma^M_k = \frac{\psi^M_{k11}/\psi^M_{k1.}}{\psi^M_{k01}/\psi^M_{k0.}} = \gamma^M \qquad \text{and}$$
$$\gamma^F_k = 1 \quad \text{for all } k, \tag{3}$$

whereas the *two group* (TG) model (U.S. Bureau of the Census 1999) assumes a constant proportion in the Census among those captured in the ACE relative to the proportion in the Census among the entire population,

$$\text{TG:} \quad \eta^M_k = \frac{\psi^M_{k11}/\psi^M_{k.1}}{\psi^M_{k1.}/\psi^M_{k..}} = \eta^M \qquad \text{and} \qquad \eta^F_k = 1 \quad \text{for all } k. \tag{4}$$

The TG model can be derived by assuming that each poststratum contains two (or more) unobserved groups within which independence holds; the model then assumes that the relative inclusion probabilities of the Census, the ACE, and the total population into these groups are equal across the poststrata.

Under the FRR and TG models, the total population, $\hat{\psi}^{DA}_{\dots}$, is estimated under the same demographic constraints as under the FOR model. The total population size within each poststratum is then given by $\hat{\psi}^\gamma_{k..} = \hat{\psi}^I_{k..} + (\hat{\psi}^{DA}_{\dots} - \hat{\psi}^I_{\dots})([\hat{\psi}_{k01} + \hat{\psi}^I_{k00}]/[(\hat{\psi}_{.01} + \hat{\psi}^I_{.00}])$ under the FRR model and by $\hat{\psi}^\eta_{k..} = \frac{\hat{\psi}^{DA}_{\dots}}{\hat{\psi}^I_{\dots}}\hat{\psi}^I_{k..}$ under the TG model. Thus the TG model has the property that the discrepancy between the DA estimate and the independence estimate is allocated to the poststrata in proportion to the independence estimate. Note that each of these models contains the independence model as a special case (Elliott and Little 2000); the independence model implies that $\gamma^S_k = \theta^S_k = \eta^S_k = 1$ for all k. Also, by formulating these correlation bias parameters as a function of each other and the undercount or enumeration proportion parameters, it can be shown that these models are nonnested; that is, $\psi^\theta_{k..} \neq \psi^\gamma_{k..} \neq \psi^\eta_{k..}$, unless independence holds.

Because ρ is estimated using demographic analysis within an age–race group, the models are applied separately to each of the six age–race domains: 18–29, 30–40, and 50 or older, by African–American versus non–African–American. We suppress an age–race domain subscript for notational simplicity.

Initial DA estimates suggested that there was substantial overcount and thus *negative* correlation bias among non–African–Americans age 18–29 (Robinson 2001). This degree of negative correlation bias was implausible, especially given that estimation of unauthorized immigration is a key component of the sex ratio estimate in this age–race group. The DA estimates used in this analysis assume a 15% underestimation in unauthorized immigration relative to the baseline unauthorized immigration estimates.

Other model have been proposed to combine Census, ACE, and demographic data (Bell 1993; Bell et al. 1996). For example, a fixed sex ratio model would assume that the sex ratio in the Out-ACE Out-Census cell is constant across poststrata. More general models could be identified by assuming equal correlation bias in the male and female poststrata: however, there is evidence that different degrees of correlation bias exist among males and females, with the independence assumption more tenable among females than males (Bell 1993); also, estimates under these models tend to be very unstable (Mallet, Rivest, and Bell 1994). Elliott and Little (2000) argued in favor of (2)–(4) because they included the independence model as a special case; however, this limitation is to some degree arbitrary. The model comparisons described next could be extended to a larger class of models if desired.

Downloaded by [Universita degli Studi di Torino] at 01:46 09 January 2018

3. BAYESIAN MODELS FOR CENSUS EVALUATION

Because the cell counts are sums, we model the data from the kth postratum as $(y_{k11}^S\ y_{k01}^S\ z_{k1.}^S)^T \sim \mathrm{N}_3((\psi_{k11}^S\ \psi_{k01}^S\ \psi_{k1.}^S)^T, \boldsymbol{\Sigma}_k)$. The covariance $\boldsymbol{\Sigma}_k$ is estimated via a jackknife procedure and is treated as known.

Assuming the FOR model given by (2), we reparameterized the eight population counts in poststratum k, $\psi_k = \{\psi_{kij}^S : i = 0, 1; j = 0, 1; S = M, F\}$, as $\boldsymbol{\psi}_k^* = (\psi_{k..}, \rho_k, \delta_k^M, \delta_k^F, \phi_k^M, \phi_k^F, \theta_k^M, \theta_k^F)$, where

- $\psi_{k..}$, the total population count in poststratum k
- $(\psi_{k..}^M)/(\psi_{k..}^F) = \rho_k$, the sex ratio $(\psi_{k..}^M + \psi_{k..}^F = \psi_{k..})$
- $\delta_k^S = \Phi^{-1}(\psi_{k1.}^S/\psi_{k..}^S)$, the linear predictor of the probit regression on the proportion of the total population captured in the adjusted Census counts ("E-total") for sex S
- $\phi_k^S = \Phi^{-1}(\psi_{k11}^S/\psi_{k1.}^S)$, the linear predictor of the probit regression on the proportion of the Census cases enumerated in the ACE ("match total") for sex S
- $\frac{\psi_{k11}^S/\psi_{k10}^S}{\psi_{k01}^S/\psi_{k00}^S} = \theta_k^S$, a measure of correlation bias given by the odds among sex S of being included in the ACE among Census respondents relative to odds of being included in the ACE among Census nonrespondents for sex S.

We then selected the following independent priors for each parameter, with weakly informative hyperpriors:

- $\psi_{k..} \sim \mathrm{N}(\hat{\psi}_{k..}^I, (\hat{\psi}_{k..}^I)^4)$, a nearly flat prior corresponding to our lack of knowledge about the total population counts in each poststratum
- $\rho_k \sim \mathrm{N}(\rho, 1)$, subject to the constraint that $\sum_k w_k \rho_k = \rho$, where $w_k = (\psi_{k..})/(\sum_k \psi_{k..})$ and ρ is the DA-estimated nationwide sex ratio
- $\delta_k^S \sim \mathrm{N}((\mathbf{x}_k^S)'\boldsymbol{\beta}_\delta^S, (\sigma^2)_\delta^S)$ and $\phi_k^S \sim \mathrm{N}((\mathbf{x}_k^S)'\boldsymbol{\beta}_\phi^S, (\sigma^2)_\phi^S)$. The probit regression models smooth the proportion of Census undercounts and the proportion of Census cases enumerated in the ACE across poststrata and allows us to estimate common mean effects of poststratum covariates $\mathbf{x}_k^S$. For the African–American age–race groupings, $\mathbf{x}_k^S$ includes an intercept and dummy variables for owner versus renter and high versus low return-rate households; for the non–African–American age–race groupings, $\mathbf{x}_k^S$ includes an intercept and dummy variables for ownership, return rate, and ethnicity (Hispanic, Native Hawaiian or Pacific Islander, Asian, and American Indian or Alaskan Native). We further assume that $\boldsymbol{\beta}_\delta^S \sim \mathrm{N}(\mathbf{0}, \mathrm{diag}\,1{,}000)$, $\boldsymbol{\beta}_\phi^S \sim \mathrm{N}(\mathbf{0}, \mathrm{diag}\,1{,}000)$, $\sigma_\delta^2 \sim inv - \mathrm{Gamma}(.01, .01)$, and $\sigma_\phi^2 \sim inv - \mathrm{Gamma}(.01, .01)$.
- $\theta_k^M = \theta^M \sim \mathrm{Gamma}(\alpha, \beta)$ for all k; $\theta_k^F = 1$ for all k. These priors assume that the correlation bias is a constant across poststrata (likely > 1) for all males and is constant and known to be equal to 1 (under the independence assumption) for females. We choose α and β to maximize the correlation bias prior variance $(\alpha\beta^2)$ under the constraint that 95% of the prior distribution lies within a reasonable interval, $P(1.0 < \theta^M < 11.0) = .95$.

The FRR model (3) replaces θ_k^S with $\gamma_k^S = \frac{\psi_{k11}^S/\psi_{k1.}^S}{\psi_{k01}^S/\psi_{k0.}^S}$, the proportion in the ACE among Census respondents relative to the proportion in the ACE among Census non-respondents for sex S; the TG model (4) replaces θ_k^S with $\eta_k^S = \frac{\psi_{k11}^S/\psi_{k.1}^S}{\psi_{k1.}^S/\psi_{k..}^S}$, the proportion in the Census among ACE respondents relative to the proportion in the Census among the entire population. The remainder of the model formulation is that same, except that the correlation bias prior Gamma(α, β) parameters are changed so that $P(.5 < \gamma^M < 2.0) = P(.8 < \eta^M < 1.2) = .95$, which have roughly similar interpretations of correlation bias.

We used Gibbs sampling to draw estimates of the population parameters from their joint posterior distribution. We obtained preliminary results by drawing from nonstandard distributions via the inverse-cdf method; we estimated normalizing constants by numerical integration using 32-point Gauss–Hermite quadrature on each of 1,000 subintervals of the support. Results were similar to those from simple Monte Carlo integration, so we used Monte Carlo integration for full "production." To determine that convergence of the Gibbs sampler had been achieved, we ran five chains of 1,000 draws with different startpoints and calculated the between- and within-sequence posterior variances and for each (scalar) parameter after dropping the first 100 draws as "burn-in." The estimated ratio of the marginal posterior variance for all sequences to the mean within-sequence posterior variance, $\widehat{R}$, was determined to be < 1.1 for all parameters (Gelman, Carlin, Stern, and Rubin 2004, pp. 296–297).

3.1 Model Checking/Model Selection

Regarding model fit, we can consider how well posterior predictive distributions match the observed cell counts or, more formally, the posterior probability that the model is "correct" among the models considered. We obtain the posterior predictive distribution of the $3K$ cell counts in each age–race domain,

$$f\big((y_{k11}^S)^{\mathrm{pred}}, (y_{k01}^S)^{\mathrm{pred}}, (z_{k1.}^S)^{\mathrm{pred}} \big| \mathbf{y}, \mathbf{z}\big) = \int f(y_{k11}^S, y_{k01}^S, z_{k1.}^S | \boldsymbol{\psi}_k^*) p(\boldsymbol{\psi}_k^* | \mathbf{y}, \mathbf{z})\, d\boldsymbol{\psi}_k^*, \quad (5)$$

by using the Gibbs sampling draws as $\boldsymbol{\psi}_k^{*\mathrm{rep}}$, backtransforming to $\boldsymbol{\psi}_k^{\mathrm{rep}}$, and then generating $((y_{k11}^S)^{\mathrm{pred}}\ (y_{k01}^S))^{\mathrm{pred}}\ (z_{k1.}^S))^{\mathrm{pred}})^T$. We consider two types of posterior predictive distributions,

$$\big((y_{k11}^S\ y_{k01}^S\ z_{k1.}^S) - (\psi_{k11}^S\ \psi_{k01}^S\ \psi_{k1.}^S)\big)^T \times \boldsymbol{\Sigma}_k^{-1}\big((y_{k11}^S\ y_{k01}^S\ z_{k1.}^S) - (\psi_{k11}^S\ \psi_{k01}^S\ \psi_{k1.}^S)\big) \quad (6)$$

and

$$z_{k1.}^S - y_{k11}^S. \quad (7)$$

By comparing the distributions of (6) and (7) obtained under the posterior predictive distribution (5) with those obtained using the observed data, we obtain *posterior predictive distribution* (PPD) p values (Gelman, Meng, and Stern 1996) that indicate whether the observed data deviate seriously from what we would have expected to observe under the model. In particular, (6) gives an omnibus measure of model fit, whereas (7) is designed to detect the presence of poststrata with match counts y_{k11} larger than the marginal adjusted E-sample totals $z_{k1.}$ that may not be due entirely to sampling variability.

To choose among the three models under consideration, we can consider Bayes factors. The Bayes factor comparing

Downloaded by [Universita degli Studi di Torino] at 01:46 09 January 2018

model D_j with model $D_{j'}$ among the models $D = \{D_j : FRR, FOR, TG\}$ is given by

$$BF = \frac{p(\mathbf{y}, \mathbf{z}|D = D_j)}{p(\mathbf{y}, \mathbf{z}|D = D_j')},$$

where $p(\mathbf{y}, \mathbf{z}|D = D_j) = \int f(\mathbf{y}, \mathbf{z}|\boldsymbol{\psi}^{*D_j}) p(\boldsymbol{\psi}^{*D_j})\, d\boldsymbol{\psi}^{*D_j}$, where $\boldsymbol{\psi}^{*D_j}$ is the parameterization of $\boldsymbol{\psi}^*$ associated with model $D = D_j$ and

$$\log p(\mathbf{y}, \mathbf{z}|D = D_j) \\ = \log f(\mathbf{y}, \mathbf{z}|\boldsymbol{\psi}^{*D_j}) + \log p(\boldsymbol{\psi}^{*D_j}) - \log p(\boldsymbol{\psi}^{*D_j}|\mathbf{y}, \mathbf{z}) \quad (8)$$

for all $\boldsymbol{\psi}^{*D_j}$. Following the method of Chib (1995), we then estimate (8) at a value $\boldsymbol{\psi}^{*D_j} = \tilde{\boldsymbol{\psi}}^{*D_j}$ with high posterior probability, here the posterior median.

4. RESULTS

4.1 Total Population Estimates

Figure 1 plots the posterior means of the undercount estimates under the FOR model against the DSE undercount estimates, the posterior means of the undercount estimates under the FRR model, the posterior means of the undercount estimates under the TG model, and the undercount estimates under the FRR model against the undercount estimates under the TG model. Examining these plots together with the specific poststrata associated with each point indicates that undercount tends to be greatest among male African–Americans and Hispanics living in rental housing units. Here the FOR, FRR, and TG models are usually estimating positive correlation bias, resulting in a higher estimated undercount than that under the DSE model. The female estimates of undercount are, of course, more similar under the four models due to the common assumption of independence, although two poorly estimated Pacific Islander poststrata (females age 50+ in rental households; males age 18–29 in owner-occupied households) with high DSE undercount estimates are pulled toward 0. A greater degree of undercount is suggested under the FOR model than under the FRR or TG model in male minority strata, particularly among Native Americans and Pacific Islanders. The FRR and TG models tend to give more similar estimates.

Figure 2 shows the relationship of the capture parameters δ_k^S and ϕ_k^S between the models. The TG model tends to assign higher ACE capture rates to predominately Hispanic and Asian male poststrata, and consequently to produce smaller estimates of undercount in these poststrata.

Table 2 gives the DSE estimates and approximate 95% confidence intervals of undercount/overcount, together with the posterior mean and 95% posterior predictive intervals under the FOR, FRR, and TG models, for the nationwide age–

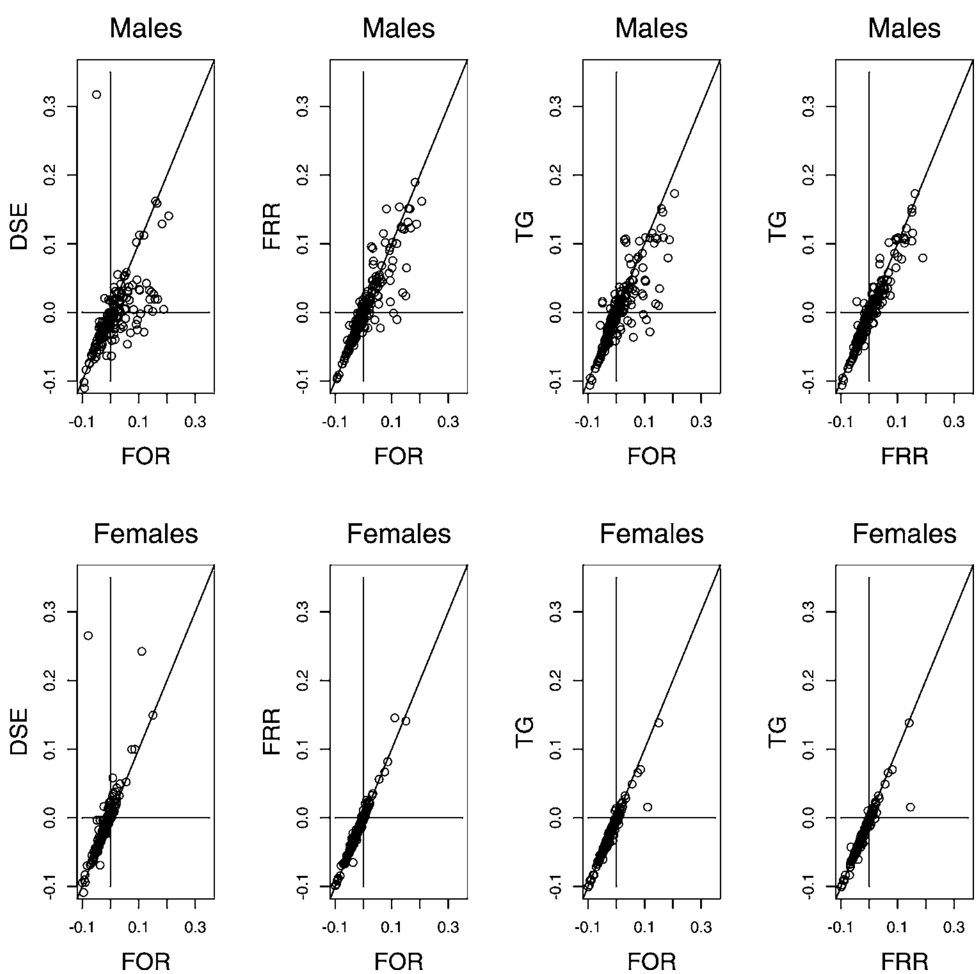

Figure 1. DSE Estimates and Posterior Mean Estimates of Undercount/Overcount in 384 Poststrata Under the FOR, FRR, and TG Models: FOR versus DSE, FRR; and TG and FRR versus TG, Separately for Males and Females.

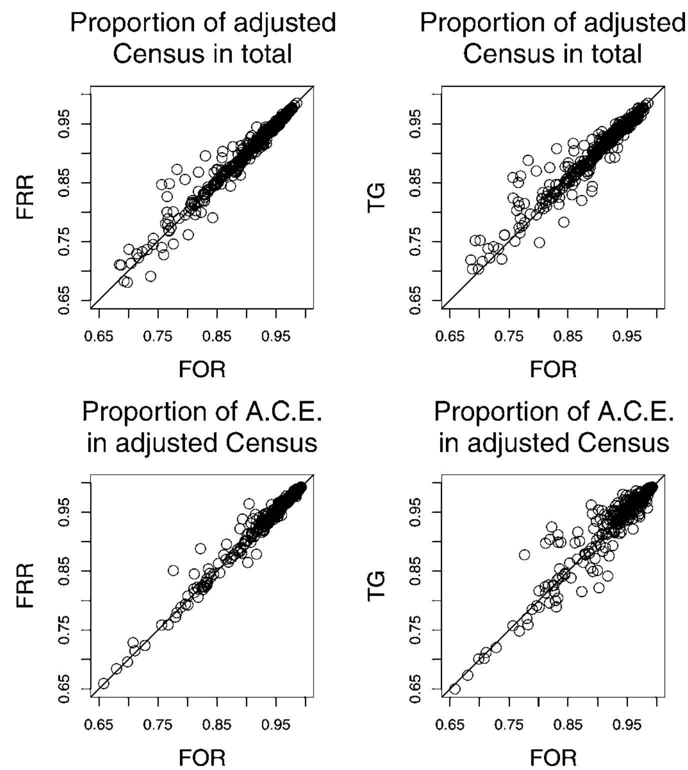

Figure 2. Posterior Means of δ_k^S (proportion of adjusted Census count in total) and ϕ_k^S (proportion of ACE count in adjusted Census count) and Under the FOR, FRR, and TG Models: FOR versus FRR and TG.

race/ethnicity–gender groupings. The estimates of undercount among African–Americans are similar for the three models. The TG model tended to give smaller estimates of undercount (larger estimates of overcount) among Caucasians age 18–29 and Asians, Hispanics, and Native Americans; the FOR model tended to give larger estimates of undercount (smaller estimates of overcount) among Asian and Hispanic males of all ages.

4.2 Effect of Poststratum Characteristics on Undercount

Examining the probit regression parameters, it appears that, after adjusting for race/ethnicity, sex, and ownership status, older persons and females tended to have higher capture and recapture rates, whereas African–Americans, Hispanics, and American Indians tended to have lower capture and recapture rates. The effect of ethnicity was more pronounced among the recapture rates than among the capture rates. Those in owner-occupied dwelling units tended to have higher capture rates only, whereas high postal return rates were associated with increases in recapture rates only, and only in non–African–American poststrata. Recapture rates tended to be higher than capture rates for a given set of poststratum predictors.

4.3 Correlation Bias Estimates

Table 3 gives the posterior means and credible intervals for the correlation bias parameters for the FOR, FRR, and TG models, for each of the six age–race/ethnicity groups. Positive correlation bias appear to be strongest among African–Americans age 30 and older and are present among all age–race domains except non–African–Americans age 18–29. In this age–race domain, the models disagree. The FOR and FRR models are not

Table 2. Percent Undercount Estimates by Age, Race, and Gender in the 2000 Census: DSE Estimate and FOR, FRR, and TG Posterior Means

	DSE	*FOR*	*FRR*	*TG*
Males				
African–American age 18–29	−.1 (−2.0,1.8)	7.4 (5.8,8.9)	7.4 (5.8,9.1)	7.4 (5.8,9.0)
African–American age 30–49	.1 (−1.1,1.2)	10.1 (9.1,11.0)	10.1 (9.2,11.1)	10.2 (9.3,11.0)
African–American age 50+	−2.8 (−3.9,−1.7)	2.2 (1.2,3.1)	2.2 (1.3,3.3)	2.2 (1.1,3.3)
Caucasian age 18–29	−1.1 (−1.8,−.4)	−1.6 (−2.2,−1.0)	−1.7 (−2.3,−1.1)	−2.1 (−2.8,−1.5)
Caucasian age 30–49	−.8 (−1.2,−.4)	−.1 (−.5,.4)	.3 (−.0,.7)	.4 (.0,.7)
Caucasian age 50+	−2.3 (−2.7,−1.9)	−1.7 (−2.0,−1.3)	−1.5 (−1.8,−1.1)	−1.6 (−1.9,−1.3)
Hispanic age 18–29	4.8 (3.2,6.4)	4.5 (2.8,6.5)	4.2 (2.5,6.0)	4.1 (2.5,5.7)
Hispanic age 30–49	1.2 (.1,2.2)	5.0 (2.4,7.5)	3.5 (2.2,4.7)	2.5 (1.5,3.5)
Hispanic age 50+	−1.6 (−3.0,−.3)	.8 (−1.8,4.6)	−.1 (−1.5,1.3)	−1.0 (−2.2,.2)
Asian age 18–29	−2.0 (−5.1,.9)	−2.5 (−5.2,.6)	−2.7 (−5.4,−.0)	−3.0 (−5.6,−.1)
Asian age 30–49	.2 (−1.7,2.2)	1.6 (−1.1,6.3)	1.9 (−.2,4.2)	1.5 (−.4,3.4)
Asian age 50+	−.5 (−2.3,1.3)	2.1 (−1.4,7.9)	.8 (−1.2,3.0)	.2 (−1.5,2.0)
Females				
African–American age 18–29	−.1 (−1.6,1.5)	−.5 (−2.0,.9)	−.5 (−2.0,1.1)	−.5 (−2.0,1.0)
African–American age 30–49	−.4 (−1.2,.5)	−.6 (−1.4,.2)	−.6 (−1.4,.3)	−.5 (−1.3,.3)
African–American 50+	−2.7 (−3.7,−1.6)	−3.0 (−3.9,−2.1)	−2.9 (−3.8,−1.9)	−2.8 (−3.4,−2.3)
Caucasian age 18–29	−2.3 (−2.9,−1.7)	−2.4 (−2.9,−1.9)	−2.5 (−3.1,−1.9)	−2.8 (−3.4,−2.3)
Caucasian age 30–49	−.7 (−1.0,−.3)	−1.0 (−1.3,−.6)	−.8 (−1.2,−.5)	−1.0 (−1.3,−.7)
Caucasian age 50+	−2.7 (−3.0,−2.4)	−2.8 (−3.1,−2.5)	−2.8 (−3.1,−2.4)	−2.9 (−3.2,−2.6)
Hispanic age 18–29	.8 (−.5,2.1)	.9 (−.5,2.3)	.7 (−.7,2.0)	.3 (−.9,1.6)
Hispanic age 30–49	−.5 (−1.4,.4)	−.8 (−1.6,.1)	−.7 (−1.6,.2)	−.8 (−1.7,.0)
Hispanic age 50+	−1.5 (−2.6,−.2)	−1.8 (−3.0,−.7)	−1.7 (−2.9,.2)	−2.1 (−3.2,−1.0)
Asian age 18–29	−2.1 (−5.2,.8)	−2.2 (−5.0,.7)	−2.4 (−5.1,.3)	−2.8 (−5.4,−.0)
Asian age 30–49	−.9 (−2.3,.5)	−1.1 (−2.5,.3)	−1.0 (−2.4,.3)	−1.2 (−2.4,.1)
Asian age 50+	−.4 (−2.1,1.3)	−.7 (−2.3,.9)	−.6 (−2.2,1.0)	−.8 (−2.5,.9)

NOTE: The 95% credible intervals are in parentheses. Negative values indicate Census overcount.

Downloaded by [Universita degli Studi di Torino] at 01:46 09 January 2018

Table 3. Posterior Mean of Correlation Bias Parameter and Associated 95% Credible Interval Under the FOR, FRR, and TG Models, by Age–Race Domain

Age–race domain	FOR (θ^M)	FRR (γ^M)	TG (η^M)
African–American age 18–29	$3.70_{(2.52,5.32)}$	$1.44_{(1.28,1.64)}$	$1.075_{(1.050,1.100)}$
African–American age 30–49	$6.43_{(4.60,9.01)}$	$1.73_{(1.58,1.88)}$	$1.101_{(1.086,1.116)}$
African–American age 50+	$6.34_{(3.96,9.36)}$	$1.55_{(1.37,1.76)}$	$1.050_{(1.035,1.067)}$
Non–African–American age 18–29	$1.12_{(.57,1.90)}$	$.98_{(.92,1.05)}$	$.990_{(.982,.998)}$
Non–African–American age 30–49	$5.43_{(3.53,8.07)}$	$1.22_{(1.15,1.29)}$	$1.012_{(1.007,1.017)}$
Non–African–American age 50+	$6.28_{(3.94,9.19)}$	$1.21_{(1.13,1.30)}$	$1.008_{(1.003,1.012)}$

NOTE: Parameters are scaled differently under each model except that lack of correlation bias corresponds to 1 under all models.

suggestive of either positive or correlation bias; however, the TG model suggests a negative correlation bias; that is, that the Census capture rate for those in the ACE is lower than the Census capture rate among the entire population.

4.4 Model Checking/Model Selection Results

4.4.1 Posterior Predictive Distributions. Among the male poststrata, the posterior predictive p values of the FOR model were generally the highest, followed by the FRR model and then the TG model. Among the female poststrata, no clear pattern was present. PPD p values were generally consistent with model fit; only 6 of the 384 poststrata had PPD p values $<.10$ for all 3 models. The worst was for Caucasian females age 18–29 in owner-occupied dwelling units from high–return rate neighborhoods in large MSAs in the Northeast U.S., which yielded a PPD value for $(z^S)_{k1.} - (y^S)_{k11}$ of .005–.010 under the three models. This poststratum was with only one that also failed the overall model fit χ^2_3 statistic, with values of .07–.09 under the three models. Also problematic were Caucasian females age 30–49 in owner-occupied dwelling units from high–return rate neighborhoods in large MSAs in the Northeast U.S. [PPD value for $(z^S)_{k1.} - (y^S)_{k11}$ of .07–.09 under the three models]; because these two age–sex categories share the same set of geographic blocks, there is some evidence to suggest that overestimation of the number of matches from the ACE or overestimation of the rate of erroneous enumerations from the Census may be present in at least some of these blocks that make up the poststrata for Caucasian females residing in owner-occupied dwelling units from high–return rate neighborhoods in large MSAs in the Northeast U.S. The other poststrata with evidence of overestimation of the number of matches from the ACE or overestimation of the rate of erroneous enumerations from the Census are Caucasian males age 30–49 (PPD p values .05–.06) and 50 and older (PPD p values .06–.07) in owner-occupied dwellings units in small MSA regions in the Northeast in low–return rate areas, African–American females age 30–49 in owner-occupied dwellings units in small MSA regions in low–return rate areas (PPD p values .06–.07), and Native American females age 50 and older in rental dwelling units off reservation (PPD p values .07–.09).

4.4.2 Model Selection. Bell (1993) argued that because the various behavioral models allocate the discrepancy between the DSE estimate and the DA estimate of the male population (under the DSE assumption for females) according to untestable assumptions, there is no information in the data that favors one model over the other. In contrast to Bell, however, we use information from the covariance matrix of the data $\boldsymbol{\Sigma}_k$ to allocate estimates of both the male and female population. As a simple example, in the case of the negative In-Census Out-ACE cells, if the variance of the marginal Census count $z_{k1.}$ is small relative to the variance of the In-Census In-ACE cell y_{k11}, then the estimate of ψ_{k11} will be pulled toward $z_{k1.}$ to accommodate the fact that $z_{k1.}$ is a good estimate of $\psi_{k1.}$; the reverse will be true if the variance of $z_{k1.}$ is large relative to y_{k11}. Thus, to the extent that the allocations of the population under the constraints imposed by the DA-estimated sex ratios and the behavioral models differ between the models, this data-derived covariance matrix of the cell counts should provide information on which to base a choice between the different models. Table 4 gives the Bayes factors comparing the FOR, FRR, and TG models within each of the six age–race domains. The FOR model is strongly favored among non–African–Americans age 30–49 and 50 and older and is moderately favored among African–Americans age 18–29. The FRR model is strongly favored among non–African–Americans age 18–29. The FOR and FRR models have approximate parity among African–American age 30–49, and all three models have approximate parity among African–Americans age 50+.

Table 4. Bayes Factors (posterior odds that the model is correct) Comparing FRR, TG, and FOR Models Within Each Age–Sex Domain

	Bayes factor under gamma prior for correlation bias		Bayes factor under uniform prior for correlation bias	
Age–sex domain	FRR vs. TG	FRR vs. FOR	FRR vs. TG	FRR vs. FOR
African–American age 18–29	2.91	.012	27.6	1.78
African–American age 30–49	544	.73	15.6	.17
African–American age 50+	1.26	1.42	2.69	.17
Non–African–American age 18–29	1.85×10^4	1.32×10^3	3.15×10^7	71.0
Non–African–American age 30–49	20.9	9.10×10^{-10}	.31	7.10×10^{-10}
Non–African–American age 50+	1.20×10^{-4}	1.23×10^{-9}	2.45×10^{-3}	5.37×10^{-9}

Downloaded by [Universita degli Studi di Torino] at 01:46 09 January 2018

Downloaded by [Universita degli Studi di Torino] at 01:46 09 January 2018

Because Bayes factors function as checks of the model, including both its distributional and prior assumptions, they can be sensitive to the choice of priors distributions. Thus we conducted a sensitivity analysis by computing the Bayes factors estimated under uniform priors for the correlation bias parameters, with bounds given by those used to determine the gamma hyperprior distributions, $\theta^M \sim \text{uni}(.5, 11)$, $\gamma^M \sim \text{uni}(.5, 2)$, and $\eta^M \sim \text{uni}(.8, 1.2)$. We use importance weights to recompute posterior medians under the uniform correlation bias priors, and proceed using the method of Chib (1995) as before. It appears that the results are somewhat sensitive to the choice of prior, with the FOR model no longer favored over the FRR model in African–Americans age 18–29 and the FRR model not as heavily favored over the FOR model among non–African–Americans age 18–29. Hence the Bayes factors presented in Table 4 should be interpreted with a substantial degree of caution; inferior models may fit better with a different prior assignment.

Table 5. DSE Undercount Estimates and Posterior Mean of Undercount Estimates Under the Bayesian Relative Risk Model: Application to 1990 and 2000 Census Data

	1990		2000	
	DSE	*Bayes FRR*	*DSE*	*Bayes FRR*
African–American age 18–29	5.1	7.3	−.1	3.1
African–American age 30–49	5.1	8.0	−.2	4.2
African–American age 50+	−.9	1.6	−2.6	−.8
Caucasian age 18–29	2.0	1.6	−1.6	−2.3
Caucasian age 30–49	.5	1.2	−.7	−.3
Caucasian age 50+	−1.1	−1.0	−2.4	−1.9
Asian age 18–29	6.7	3.9	−2.0	−3.0
Asian age 30–49	.9	1.6	−.3	.1
Asian age 50+	−1.9	−3.1	−.4	1.9
Hispanic age 18–29	7.0	6.6	3.1	2.4
Hispanic age 30–49	5.3	3.6	.4	1.0
Hispanic age 50+	2.5	−.2	−1.5	−.9

NOTE: Negative values indicate Census overcount.

5. DISCUSSION

Our Bayesian approach to incorporating information from the follow-up ACE survey and demographic analysis addresses the problem of "negative cell counts" in In-Census Out-ACE cells and provides inference about population total estimates that account for all modeled sources of uncertainty. By supplying proper priors and assuming exchangeability, the proposed model shrinks outlying estimates of Census undercount toward overall means. As an example, consider PS 58, female Hawaiian and Pacific Islanders age 50 and older living in rental housing units. The DSE for this poststratum is 19,583, and the Bayes FRR posterior mean is 14,395 (the actual Census count is 15,475). However, the underlying cell counts for the DSE are unstable, and the raw In-Census percentage of 68% and In-ACE percentage of 85% are the second-lowest and sixth-lowest among all 56 poststrata of non–African–American women age 50 and older, and compare with the raw values of 98% and 103% for Pacific Islander females age 50 and older in owner-occupied households. The Bayesian FRR model smoothes both the In-Census and In-ACE percentages to 94%. Recognizing that some the poststratum estimates are unstable, the Census Bureau collapsed 48 of the poststrata across age categories, including Hawaiian and Pacific Islanders living in rental housing units. Computing the DSEs and the "synthetically uncollapsing" the results using the Census estimates to allocate the DSE-estimated undercount (W. Bell, personal communication) yields a total population estimate of 16,258 for female Hawaiian and Pacific Islanders age 50 and older living in rental housing units, which is closer to the Bayes posterior mean.

There was some evidence of bias in match rate or erroneous enumeration estimates that would lead to excessively large match counts or too-small marginal adjusted E-sample total counts among several poststrata, including those derived from in owner-occupied dwelling units from high–return rate neighborhoods in large MSAs in the Northeast U.S., and from owner-occupied dwellings units in small MSA regions in the Northeast in low–return rate areas. Review of these blocks that make up this poststratum may be warranted.

Computation of the posterior predictive distribution p values and Bayes factors provides some evidence that the TG model tends to not fit as well as the FOR and FRR models. The FRR model tends to fit best when correlation bias is low, whereas the FOR model tends to fit best when correlation bias is high. This is consistent with the FRR model tending to give undercount estimates between the larger undercount estimates of FOR model and the small undercount estimates of the TG model. However, these results appear to be also influenced in part by our formulation of the model priors, so these comparisons should be viewed with some caution. Finally, the correlations of the capture parameters δ_k^S and ϕ_k^S are .44, .45, and .45 for females and .47, .39, and .62 for males under the FRR, TG, and FOR models, providing further evidence in favor of the FOR model and against the TG model.

Table 5 briefly summarizes undercount estimates by age and race/ethnicity in 1990 and 2000. Comparing the DSE estimates, it appears that the 2000 Census reduced or eliminated undercount among African–Americans, Asians, and Hispanics age 30 and older. Comparing the Bayes FRR models suggests that undercount remains a problem for African–Americans under age 50, albeit reduced from 1990. Overcount appears to be an increasing problem, especially among Caucasians and Asians under age 30.

Census adjustment is a difficult problem, and we believe that a good method for adjusting the Census should make use of all available information that is of sufficiently high quality to improve the raw Census counts. The current mode of operation is to use the postenumeration survey to adjust the Census, under the unrealistic assumption of zero correlation bias, and then use demographic information as an assessment tool. The Bayesian approach to Census adjustment adopted in this article allows incorporation of information from demographic analysis directly into the statistical adjustment process based on the postenumeration survey, yielding unified estimates, together with estimates of uncertainty, that capitalize on all sources of information.

[Received December 2002. Revised May 2004.]

REFERENCES

Bell, W. R. (1993), "Using Information From Demographic Analysis in Postenumeration Survey Estimation," *Journal of the American Statistical Association*, 88, 1106–1118.

Bell, W., Gibson, C., Das Gupta, P., Spenser, G., Robinson, G., Mulry, M., Vacca, A., Far, R., and Leggieri, C. (1996), "Report of the Working Group on the Use of Demographic Analysis in Census 2000," U.S. Department of Commerce, Bureau of the Census.

Chib, S. (1995), "Marginal Likelihood From the Gibbs Output," *Journal of the American Statistical Association*, 90, 1313–1321.

Choi, C. Y., Steel, D. G., and Skinner, T. J. (1988), "Adjusting the 1986 Australian Census Count for Underenumeration," *Survey Methodology*, 14, 173–190.

Datta, G. S., Ghosh, M., Huang, E. T., Isaki, C. T., Schultz, L. K., and Tsay, J. H. (1992), "Hierarchical and Empirical Bayes Methods for Adjustment of Census Undercount: The 1988 Missouri Dress Rehearsal Data," *Survey Methodology*, 18, 95–108.

Elliott, M. R., and Little, R. J. A. (2000), "A Bayesian Approach to Combining Information From a Census, a Coverage Measurement Survey and Demographic Analysis," *Journal of the American Statistical Association*, 95, 351–362.

Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2004), *Bayesian Data Analysis* (2nd ed.), New York: Chapman & Hall.

Gelman, A., Meng, X.-L., and Stern, H. S. (1996), "Posterior Predictive Assessment of Model Fitness via Realized Discrepancies" (with discussion), *Statistica Sinica*, 6, 733–807.

Ghosh, M., and Meeden, G. (1986), "Empirical Bayes Estimation in Finite Population Sampling," *Journal of the American Statistical Association*, 81, 1058–1062.

Ghosh, M., Natarajan, K., Stroud, T. F. W., and Carlin, B. P. (1998), "Generalized Linear Models for Small-Area Estimation," *Journal of the American Statistical Association*, 93, 273–282.

Hogan, H. (1993), "The 1990 Postenumeration Survey: Operations and Results," *Journal of the American Statistical Association*, 88, 1047–1060.

Isaki, C. T., Huang, E. T., and Tsay, J. H. (1991), "Smoothing Adjustment Factors From the 1990 Postenumeration Survey," in *Proceedings of the Social Statistics Section*, American Statistical Association, pp. 338–343.

Isaki, C. T., and Schultz, L. K. (1986), "Dual System Estimation Using Demographic Analysis Data," *Journal of Official Statistics*, 2, 169–179.

Isaki, C. T., Tsay, J. H., and Fuller, W. A. (2000), "Estimation of Census Adjustment Factors," *Survey Methodology*, 26, 31–42.

Mallet, M., Rivest, L. P., and Bell, W. R. (1994), "Capture-Recapture Estimation With Known Sex Ratio," *Biometrics*, 50, 1219–1223.

Office of the Inspector General (2001), "Actions to Address the Impact on the Accuracy and Coverage Evaluation of Suspected Duplicate Persons in the 2000 Decennial Census," Inspection Report No. OSE-13812, Washington, DC: U.S. Department of Commerce, available at *http://www.oig.doc.gov/reports/2001-3/2001-3-13812-01.pdf*.

Robinson, J. G. (2001), "ESCAP II: Demographic Analysis Results," Executive Steering Committee for ACE Policy II, Report No. 1, Washington, DC: U.S. Bureau of the Census, U.S. Department of Commerce, available at *http://www.census.gov/dmd/www/pdf/Report1.PDF*.

Seber, G. A. F. (1982), *The Estimation of Animal Abundance and Related Parameters*, New York: Macmillan.

Sekar, C. C., and Deming, W. E. (1949), "On a Method of Estimating Birth and Death Rates and the Extent of Registration," *Journal of the American Statistical Association*, 44, 101–115.

U.S. Bureau of the Census (1988), "The Coverage of the Population in the 1980 Census," PHC80-E4, Washington, DC: U.S. Department of Commerce.

——— (2000), "Census 2000: ACE Methodology," Washington, DC: U.S. Department of Commerce, available at *http://www.census.gov/dmd/www/DetailSpec.htm*.

——— (2001), "Report of the Executive Steering Committee for Accuracy and Coverage Evaluation Policy on Adjustment for Non-Redistricting Uses," Washington, DC: U.S. Department of Commerce, available at *http://www.census.gov/dmd/www/pdf/Recommend2.pdf*.

——— (2003), "Technical Assessment of ACE Revision II," Washington, DC: U.S. Department of Commerce, available at *http://www.census.gov/dmd/www/pdf/ACETechAssess.pdf*.

Wolter, K. M. (1990), "Capture-Recapture Estimation in the Presence of a Known Sex Ratio," *Biometrics*, 46, 157–162.

Downloaded by [Universita degli Studi di Torino] at 01:46 09 January 2018