

Evaluation and Selection of Models for Attrition Nonresponse Adjustment

Eric Slud and Leroy Bailey

Statistical Research Div., Census Bureau

FCSM Research Conference, 2007

Outline

- 1 Longitudinal Nonresponse Weight Adjustment
 - Previous Evaluation Methods
 - Model-based Adjustments
- 2 Adjustment Metrics
 - Subdomain Bias & 1st Metric
 - Whole-Pop Bias vs. Bound
 - 2nd Metric: Preserving Cells
- 3 Computed Results on SIPP 96

Definitions motivated by SIPP

Survey of Income & Program Participation

- Individuals i observed in successive 'Waves'
- Sample \mathbf{s} here denotes Wave-1 **responders**
- Weights w_i are inverse inclusion prob's
re-weighted for Wave-1 response & raked.
- Survey items y_i e.g. socSec income indicator
are **only Wave 1** data values.

Definitions motivated by SIPP

Survey of Income & Program Participation

- Individuals i observed in successive 'Waves'
- Sample \mathbf{s} here denotes Wave-1 **responders**
- Weights w_i are inverse inclusion prob's
re-weighted for Wave-1 response & raked.
- Survey items y_i e.g. socSec income indicator
are **only Wave 1** data values.

- $r_i =$ indicator of later-wave (say Wave-4) response
treated as random beyond probability sample
- \hat{p}_i model-based estimator of $p_i = \Pr(r_i = 1 \mid i \in \mathbf{s})$

Tracking of time-changes in weights

Tracking of time-changes in weights

- *Dufour et al. 2001 study adj. weights $w_i r_i / \hat{p}_i$
simulated using Canadian Labor Force survey.*

Tracking of time-changes in weights

- *Dufour et al. 2001 study adj. weights $w_i r_i / \hat{p}_i$ simulated using Canadian Labor Force survey.*

Present Methodology – Internal Evaluation

Target $t_y = \sum_{i \in U} y_i$ wave 1 responder total

Estimates *initial* $\sum_{i \in S} w_i y_i$ (Horvitz-Thompson)

at later wave, adjusted: $\sum_{i \in S} w_i y_i r_i / \hat{p}_i$

Adj. Bias $\sum_{i \in S} w_i y_i \left\{ r_i / \hat{p}_i - 1 \right\}$

SIPP *Implemented in Bailey (2004), with SE calculation in Slud & Bailey (2006).*

Model-based Adjustments

- **Cell-based Method:** for i in cell $C \subset U$,

$$\hat{p}_i = \sum_{j \in C} r_j w_j / \sum_{j \in C} w_j$$

Model-based Adjustments

- **Cell-based Method:** for i in cell $C \subset U$,

$$\hat{p}_i = \sum_{j \in C} r_j w_j / \sum_{j \in C} w_j$$

- **Logistic Regression Method:** $\hat{p}_i = (1 + e^{-\mathbf{x}'_i \hat{\beta}})^{-1}$
fitted from weighted ML equation, Wave 1 covariates \mathbf{x}_i
and later-Wave indicator r_i as response.

Model-based Adjustments

- **Cell-based Method:** for i in cell $C \subset U$,

$$\hat{p}_i = \sum_{j \in C} r_j w_j / \sum_{j \in C} w_j$$

- **Logistic Regression Method:** $\hat{p}_i = (1 + e^{-\mathbf{x}'_i \hat{\beta}})^{-1}$
fitted from weighted ML equation, Wave 1 covariates \mathbf{x}_i
and later-Wave indicator r_i as response.
- **Score-statistic model comparisons:** Kim and Kim (2007).

Model-based Adjustments

- **Cell-based Method:** for i in cell $C \subset U$,

$$\hat{p}_i = \sum_{j \in C} r_j w_j / \sum_{j \in C} w_j$$

- **Logistic Regression Method:** $\hat{p}_i = (1 + e^{-\mathbf{x}'_i \hat{\beta}})^{-1}$
fitted from weighted ML equation, Wave 1 covariates \mathbf{x}_i
and later-Wave indicator r_i as response.
- **Score-statistic model comparisons:** Kim and Kim (2007).
- **Model comparisons via Subdomain Bias Estimates**

Subdomain Bias & 1st Metric

- **Bias over Subdomain** \mathcal{D} for k 'th survey item

$$\hat{B}_k(\mathcal{D}) = \sum_{i \in \mathcal{D} \cap \mathcal{S}} \left(\frac{r_i}{\hat{p}_i} - 1 \right) w_i y_i^{(k)}$$

Subdomain Bias & 1st Metric

- **Bias over Subdomain** \mathcal{D} for k 'th survey item

$$\hat{B}_k(\mathcal{D}) = \sum_{i \in \mathcal{D} \cap \mathcal{S}} \left(\frac{r_i}{\hat{p}_i} - 1 \right) w_i y_i^{(k)}$$

- **Metric def'n:** max Relative Bias over consecutive subsets after random re-ordering $\tau = (\tau(1), \tau(2), \dots, \tau(n))$ of \mathcal{S} : then averaged over permutations τ :

$$m_k = E_{\tau} \left(\max_{1 \leq a \leq n} |\hat{B}_k(\{\tau(1), \dots, \tau(a)\})| \right) / \hat{t}_{y^{(k)}}$$

Subdomain Bias & 1st Metric

- **Bias over Subdomain** \mathcal{D} for k 'th survey item

$$\hat{B}_k(\mathcal{D}) = \sum_{i \in \mathcal{D} \cap \mathcal{S}} \left(\frac{r_i}{\hat{p}_i} - 1 \right) w_i y_i^{(k)}$$

- **Metric def'n**: max Relative Bias over consecutive subsets after random re-ordering $\tau = (\tau(1), \tau(2), \dots, \tau(n))$ of \mathcal{S} : then averaged over permutations τ :

$$m_k = E_{\tau} \left(\max_{1 \leq a \leq n} |\hat{B}_k(\{\tau(1), \dots, \tau(a)\})| \right) / \hat{t}_{y^{(k)}}$$

- **Monte Carlo Estimate** (over permutations τ_c):

$$\hat{m}_k = \sum_{c=1}^R \max_{1 \leq a \leq n} |\hat{B}_k(\{\tau_c(1), \dots, \tau_c(a)\})| / (R \hat{t}_{y^{(k)}})$$

Whole-Pop Bias vs. Bound

- Random permutations imply cannot adjust average max bias to 0 !!

$$\begin{aligned}
 \left| m_k - \frac{|\hat{B}_k(\mathcal{U})|}{\hat{t}_{y^{(k)}}} \right| &\leq \frac{1.229}{\hat{t}_{y^{(k)}}} \left(\sum_{i \in \mathcal{S}} \left(\frac{r_i}{\hat{p}_i} - 1 \right)^2 \left(\frac{y_i^{(k)}}{\pi_i} \right)^2 \right)^{1/2} \\
 &= \hat{b}_k
 \end{aligned}$$

- **Dominant term** in m_k & \hat{m}_k , when large $= \frac{|\hat{B}_k(\mathcal{U})|}{\hat{t}_{y^{(k)}}$
- Bound \hat{b}_k on RHS is correct magnitude due to maximum over subdomains if no overall bias.

2nd Metric: Preserving Cells

Subdomain biases most interesting in cells used in current cell-based adjustment:

- Random permutations σ of indices now permute cells A_j , $j = 1, \dots, J$, and **individuals within** consecutively indexed **cells**:

$$m_k^* \equiv E_\sigma \left(\max_{1 \leq q \leq n} |\hat{B}_k(\{\sigma(1), \dots, \sigma(q)\})| \right) / \hat{t}_{y(k)}$$

2nd Metric: Preserving Cells

Subdomain biases most interesting in cells used in current cell-based adjustment:

- Random permutations σ of indices now permute cells A_j , $j = 1, \dots, J$, and **individuals within** consecutively indexed **cells**:

$$m_k^* \equiv E_{\sigma} \left(\max_{1 \leq q \leq n} |\hat{B}_k(\{\sigma(1), \dots, \sigma(q)\})| \right) / \hat{t}_{y^{(k)}}$$

- Monte Carlo estimated version of this metric

$$\hat{m}_k^* \equiv \frac{1}{R} \sum_{r=1}^R \max_{1 \leq q \leq n} |\hat{B}_k(\{\sigma_r(1), \dots, \sigma_r(q)\})| / \hat{t}_{y^{(k)}}$$

Adjustment Models Used in SIPP

- **149 Adjustment Cells** used in SIPP production:
defined in terms of : Education, Income-Level,
Labor-force status, Self-employment,
Race/Hispanic, Assets, ...
- **Logistic regression** using variables:
Renter, race, 'ref-person', education,
poverty, some pairwise interactions, and
Survey Items (AFDC, SocSec, Unemp, ...)

Wave 4, 12 response fitted separately to Wave 1 predictors.

Main Issue. Previously (2006): including Poverty mostly removed Poverty adjustment bias over whole pop'n.

- **Subdomain analyses** used to study remaining biases
- add more survey items to logistic model ?

Method based on Metrics.

Main Issue. Previously (2006): including `Poverty` mostly removed Poverty adjustment bias over whole pop'n.

- **Subdomain analyses** used to study remaining biases
- add more survey items to logistic model ?

Method based on Metrics.

- Items with bad adjustment bias flagged by

$$\hat{m}_k \gg \hat{b}_k$$

- Metric value always of order $\geq \hat{b}_k$.

Logistic Models Compared on SIPP 96

	Df	Variables	Dev
A	7	Wnotsp Renter College RefPer Blk Renter*College Blk*College	76558
B	8	same as A , plus Pov	76545
C	13	same as B , plus Foodst Mdcd Heins UnEmp Div	76299
D	13	same as B , minus Blk*College + Mdcd Heins UnEmp Pov*Heins Mdcd*Heins Heins*College	76242
E	17	same as D + hisp + Famtyp	76017
F	18	C plus Afdc SocSec Emp Mar	76280

Model B vs Adjustment Cell

- Small Logistic Pov biases small, $SocSec$, $Heins$, $UnEmp$ large.
- Bounds \hat{b}_k approx. same for all models.

Item	\hat{m}^{4C}	\hat{m}^{4L}	$\hat{b}_{4,k}$
AFDC	.0067	.0248	.0078
SocSec	.0191	.0116	.0041
Heins	.0085	.0065	.0019
Pov	.0187	.0033	.0047
Emp	.0016	.0017	.0020
UnEmp	.0534	.0594	.0131

Metric for Models A, D, F, Wave 4

- Successive improvement with more Survey-item predictors.

Item	$\hat{m}^{A,A}$	$\hat{m}^{A,D}$	$\hat{m}^{A,F}$	$b_{4,k}^D$
AFDC	.0175	.0067	.0053	.0077
SocSec	.0117	.0125	.0027	.0041
Pov	.0123	.0032	.0032	.0047
UnEmp	.0626	.0095	.0098	.0139

Metrics \hat{m} , all Models

- **1st metric**, averaged over 12 survey items
- Steady improvement in models with more terms

Model	Wave-4	Wave-12
Adj.Cell	0.01228	0.04741
LReg, A	0.01451	0.03942
LReg, B	0.01504	0.03893
LReg, C	0.00426	0.02475
LReg, D	0.00571	0.02812
LReg, E	0.00481	0.02654
LReg, F	0.00342	0.00782

Metrics \hat{m}_k^* Across Models

- **2nd metric**, by survey item
- Richer models often slightly better,
but now Models C and F virtually tied !!

item	ModB	ModC	ModD	ModF	Adj.Cell
AFDC	.0840	.0728	.0734	.0722	.0697
SocS	.0300	.0305	.0310	.0287	.0332
Pov	.0632	.0574	.0576	.0566	.0572
Emp	.0281	.0263	.0261	.0259	.0227

Metrics \hat{m}_k^* Across Models

- **2nd metric**, by survey item
- Richer models often slightly better,
but now Models C and F virtually tied !!

item	ModB	ModC	ModD	ModF	Adj.Cell
AFDC	.0840	.0728	.0734	.0722	.0697
SocS	.0300	.0305	.0310	.0287	.0332
Pov	.0632	.0574	.0576	.0566	.0572
Emp	.0281	.0263	.0261	.0259	.0227

- **Low-dimensional Model C as good as Adjustment Cells.**

Summary

- **Internal evaluation** of attrition nonresponse adjustment.
- Introduced **2 metrics** for largest absolute relative bias over subsequence in random ordering of survey items.





Summary

- **Internal evaluation** of attrition nonresponse adjustment.
- Introduced **2 metrics** for largest absolute relative bias over subsequence in random ordering of survey items.
- Metrics prevent 'adjusting away' biases at whole-pop level.
- Metrics rate models **differently from Deviance**: 1st favors greater adjustment. 2nd Metric favors less.

Summary

- **Internal evaluation** of attrition nonresponse adjustment.
- Introduced **2 metrics** for largest absolute relative bias over subsequence in random ordering of survey items.
- Metrics prevent 'adjusting away' biases at whole-pop level.
- Metrics rate models **differently from Deviance**: 1st favors greater adjustment. 2nd Metric favors less.
- **Future step.** Re-do metric model-comparisons in SIPP after implementing pop-control '2nd stage adjustment'.

References

-  Dufour, J., Gagnon, F., Morin, Y., Renaud, M. and Särndal, C.-E. (2001)
Survey Methodology **27**, 97-108.
-  Bailey, Leroy (2004)
Census Bureau internal report, Nov. 16, 2004.
-  Slud, E. and Bailey, L. (2006)
Estimation of attrition biases in SIPP.
ASA Surv. Res. Meth. Proceedings, JSM 2006.
-  Kim, Jae-Kwang and Kim, Jay (2007)
Canadian Journal of Statistics, to appear.