Simultaneous Calibration and Nonresponse Adjustment

Eric V. Slud[1,2] & Yves Thibaudeau[1]
[1]Census Bureau, SRD, and [2]Univ. of Maryland

OUTLINE

I. Def'n of Constrained Optimum Problem for Loss plus Penalty
— loss for 2 types of adjustment (cf Deville & Särndal 1992)

II. Form of Linear Single-Stage Equations & Iterative Sol'ns
— comparison with special & limiting Cases

III. Superpopulation Properties of Solutions
— consistency & linearized variance formulas

IV. Numerical Illustration with SIPP 96 Data

Notations & Formulation

Frame $\mathcal{U}$, Sample $\mathcal{S}$, Initial (inclusion) weights $w_k^o = \frac{1}{\pi_k}$

Unit response indicators $r_k = 0, 1$, observe $(y_k : \; k \in \mathcal{S}, \; r_k = 1)$

$\mathbf{x}_k$ nonresponse calibration variables, usually $(I_{[k \in C_1]}, \ldots, I_{[k \in C_m]})$

$\mathbf{z}_k$ population calibration/control variables

fixed 'known' totals $t^*_{\mathbf{x}} \approx \sum_{k \in \mathcal{U}} \mathbf{x}_k$, $t^*_{\mathbf{z}} \approx \sum_{k \in \mathcal{U}} \mathbf{z}_k$

Note: totals $t^*_{\mathbf{z}}$ generally derived from external knowledge (e.g., the updated census)

BUT

the totals $t^*_{\mathbf{x}}$ may involve $\sum_s w_k^o \, r_k \, \mathbf{x}_k$ or other estimates.

Background: Standard 'Two-Stage Method'

$\mathbf{x}_k = (I_{[k \in C_1]}, \ldots, I_{[k \in C_m]})$ indicators of *adjustment cells*

$\mathbf{z}_k$ totals for population controls

1st stage: ratio adjust, for $i \in C_j$, $w_i \equiv r_i \, w_i^o \, \frac{\sum_s w_k^o \, I_{[k \in C_j]}}{\sum_s r_k \, w_k^o \, I_{[k \in C_j]}}$

2nd stage: calibrate: $\hat{w}_i = w_i \{1 + \mathbf{z}_i' (\sum_s w_k \mathbf{z}_k \mathbf{z}_k')^{-1} (t_{\mathbf{z}}^* - \sum_s w_k z_k)\}$

Notes: (i) in US, often do 2nd stage by raking instead,

(ii) inclusion prob. based variance formula given in Särndal & Lundström (2005, Sec. 11.4)

(iii) consistent (Fuller 2002) when adjustment cell response model ($P(r_k = 1)$ constant for k in each C_j) is correct and $t_{\mathbf{z}}^*$ is exact frame **z** total.

Other Approaches

Single-stage calibration: Särndal & Lundström (2005) methods would immediately allow simultaneous (linear) calibration to:

$$\sum_{k \in S} \widehat{w}_k \, r_k \binom{\mathbf{x}_k}{\mathbf{z}_k} = \binom{t^*_{\mathbf{x}}}{t^*_{\mathbf{z}}} , \quad \widehat{w}_k = w^o_k \ (1 + \lambda' \mathbf{x}_k + \mu' \mathbf{z}_k)$$

Chang & Kott (2008) instead of a constraint use a nonresponse model $P(r_k = 1) = p(\mathbf{x}'_k \beta)$, setting to 0 or minimizing a quadratic form in $\Delta = \sum_s r_k \, w^o_k \, \mathbf{z}_k / p(\widehat{\beta}' \mathbf{x}_k) - t^*_{\mathbf{z}}$

Thus Chang & Kott (2008) relax the $t_{\mathbf{z}}$ constraint and do without the $t^*_{\mathbf{x}}$ constraint.

Generally: can enforce these constraints exactly or relax them.

Our approach: relax the $t^*_{\mathbf{x}}$ constraint, but do not remove it. (Implicitly assume $t^*_{\mathbf{x}}$ less accurate or less vital.)

Relaxed nonresp. adjusted weights w_k , **final weights** $\widehat{w}_k$

Measure distances between 1, w_k/w^o_k and $\widehat{w}_k/w^o_k$ subj. to

$$\sum_{k \in S} r_k \begin{pmatrix} w_k \, \mathbf{x}_k \\ \widehat{w}_k \, \mathbf{z}_k \end{pmatrix} = \begin{pmatrix} t^*_{\mathbf{x}} \\ t^*_{\mathbf{z}} \end{pmatrix}$$

Survey totals $\sum_{k \in \mathcal{U}} y_k$ estimated by $\widehat{t}_{y,adj} = \sum_{k \in \mathcal{S}} r_k \, \widehat{w}_k \, y_k$

Convex Penalty function $Q(\cdot)$ will be applied to $\widehat{w}_k/w^o_k$

Single Stage Weight Adjustment Method

- New approach does all adjustments in **single stage**

$$\min_{\mathbf{w},\,\hat{\mathbf{w}}} \sum_{k \in \mathcal{S}} r_k\, w_k^o \,\{\, G_1\Big(\frac{w_k - w_k^o}{w_k^o}\Big) \,+\, \alpha\, G_2\Big(\frac{\widehat{w}_k - w_k}{w_k^o}\Big) \,+\, Q\Big(\frac{\widehat{w}_k}{w_k^o}\Big) \,\}$$

subject to: $$\sum_{k \in \mathcal{S}} r_k\, w_k\, \mathbf{x}_k \;=\; t^*_{\mathbf{x}} \quad , \qquad \sum_{k \in \mathcal{S}} r_k\, \widehat{w}_k\, \mathbf{z}_k \;=\; t^*_{\mathbf{z}}$$

w_k are approx. nonresp-adjusted weights, $\widehat{w}_k$ are **final weights**

$G_1(u)$, $G_2(u)$ *loss fcns* $= \begin{cases} u^2/2 & \text{for linear calibration} \\ (1+u)\log(1+u) - u & \text{for raking} \end{cases}$

$Q(u)$ convex penalty function $\equiv 0$ on interval , e.g. [0.6, 2] and finite only on larger interval, e.g. [0.1, 10)

Other Weight Compression Methods

- Deville and Särndal (1992): modify loss functions on weights to ∞ when $a = \widehat{w}_k / w_k^o$ outside fixed range $[L, U]$

- Singh and Mohl (1996), Théberge (2000): loss function for nonresponse adjusted weights includes large-weight penalty.

Deville & Särndal 'Case 6': for $a \in [L, U]$:

$$G(a-1) \; =$$

$$\frac{(1-L)(U-1)}{U-L} \{(a-L) \log \frac{a-L}{1-L} + (U-a) \log \frac{U-a}{U-1}\}$$

Examples of Penalty Functions Q

Fix $L < c_1 < 1 < c_2 < U$: define for $a \in [L, U]$:

Penalty Function $Q(a)$ remaining finite at L, U:

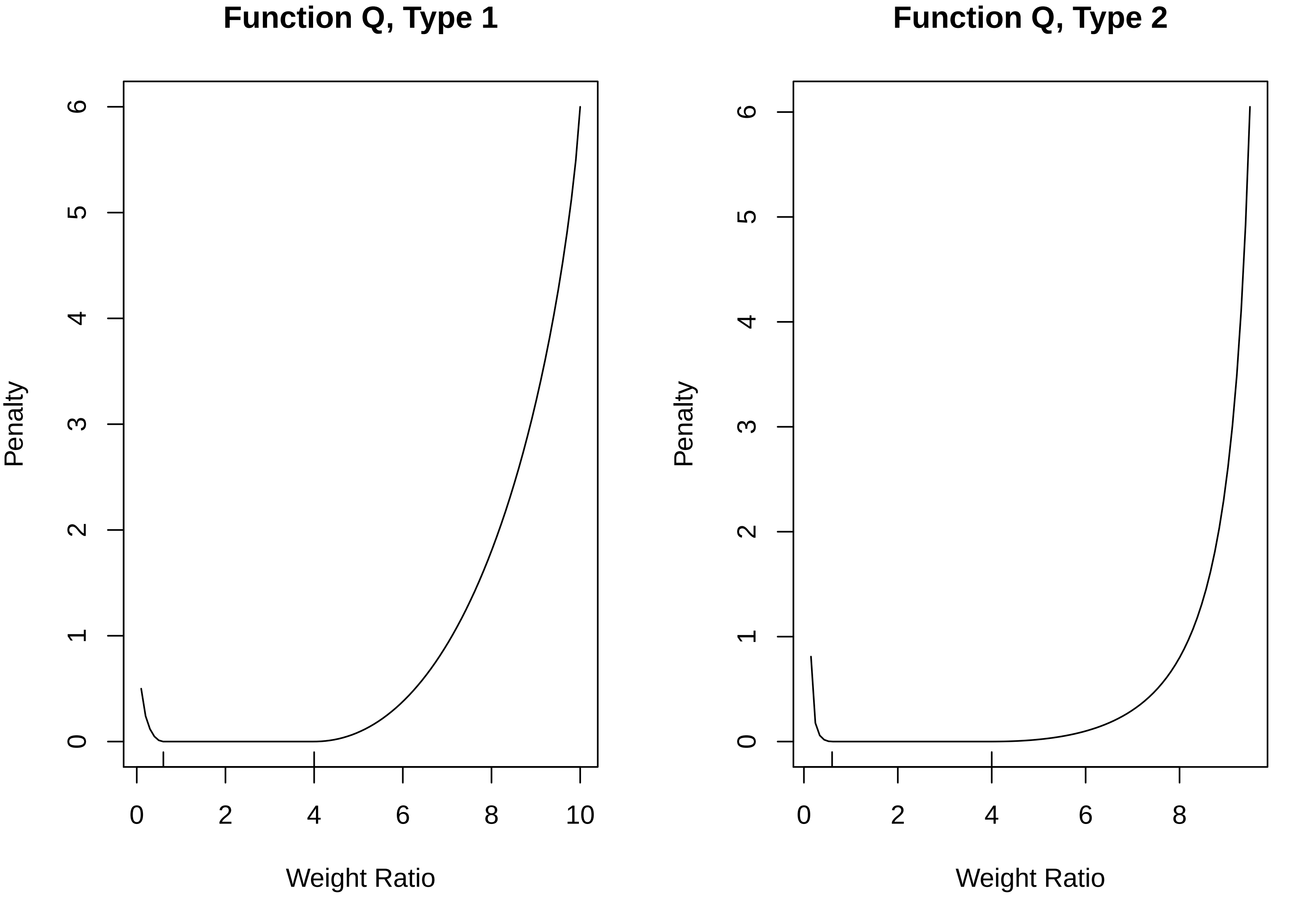
$$Q(a) \;=\; I_{[a \le c_1]} \left[(a - L) \log \frac{a - L}{c_1 - L} + c_1 - a \right]$$

$$+ I_{[c_2 \le a]} \left[(U - a) \log \frac{U - a}{U - c_2} + a - c_2 \right]$$

Penalty Function $Q(a)$ becoming infinite at L, U:

for $a \in (L, U)$:

$$Q(a) \;=\; A_1 \, I_{[a \le c_1]} \frac{(c_1 - a)^2}{a - L} + A_2 \, I_{[a \ge c_2]} \frac{(a - c_2)^2}{U - a}$$



Function Q, Type 1
Penalty
Weight Ratio
0
1
2
3
4
5
6
0
2
4
6
8
10
Function Q, Type 2
Penalty
Weight Ratio
0
1
2
3
4
5
6
0
2
4
6
8

Recall Motivation

- Current methods usually start with nonresponse weight adjustment, then impose population controls (eg by raking to demographic-cell census counts) with some weight trimming.

- Methods based on linearization, HT variance formulas require joint inclusion probabilities, but these are available only *before* weight adjustments !

- Ambiguous role of nonresponse adjustment: are the adjustment cell proportions to be maintained or not ?

Equations for Lagrange Mult's & $\widehat{w}_k$ for $G_j(u) \equiv u^2/2$

$$\widehat{w}_k + \frac{1+\alpha}{\alpha} w_k^o Q'(\frac{\widehat{w}_k}{w_k^o}) = w_k^o \{1 + \frac{1+\alpha}{\alpha} \mu' \mathbf{z}_k + \lambda' \mathbf{x}_k\}$$

Lagrange multipliers λ, μ for $t^*_{\mathbf{x}}, t^*_{\mathbf{z}}$ constraints determined by

$$M_\alpha = \sum_{k \in \mathcal{S}} r_k w_k^o \begin{pmatrix} \mathbf{x}_k \mathbf{x}_k' & \mathbf{x}_k \mathbf{z}_k' \\ \mathbf{z}_k \mathbf{x}_k' & (1+\alpha^{-1}) \mathbf{z}_k \mathbf{z}_k' \end{pmatrix}$$

$$M_\alpha \binom{\lambda}{\mu} = \binom{t^*_{\mathbf{x}}}{t^*_{\mathbf{z}}} - \sum_{k \in \mathcal{S}} r_k w_k^o \binom{\mathbf{x}_k}{\mathbf{z}_k} + \sum_{k \in \mathcal{S}} r_k w_k^o Q'(\frac{\widehat{w}_k}{w_k^o}) \binom{\mathbf{x}_k}{(1+\alpha^{-1}) \mathbf{z}_k}$$

NB. $\widehat{w}_k - w_k = \frac{w_k^o}{\alpha} (\mu' \mathbf{z}_k - Q'(\widehat{w}_k / w_k^o))$ `small when` α `large`

Numerical Solution

Solution is iterative based on:

- $w_k^o = \widehat{w}_k^{(0)}$ initially

- solve for multipliers $\lambda^{(j)}$, $\mu^{(j)}$ from $\{\widehat{w}_k^{(j)}\}_k$ using bottom displayed equation, for each $j = 0, 1, 2, \ldots$

- for each j, solve for $\widehat{w}_k^{(j+1)}$ using top displayed equation

- Convergence can be proved if $Q'(\widehat{w}_k^{(j)}/w_k^o)$ remain bounded.

`This iteration & related theory is less easy to reproduce -- and not yet known -- for general` G_j `.`

Advantages of New Method

- Ease of documentation of adustment/controls/trimming
- Tuning parameters (α and c_1, c_2, L, U in Q)
  - — small α roughly approximates '2-stage method'
  - — large α method gives ($w_k \approx \widehat{w}_k$ and) simultaneous one-stage calibration to $(\mathbf{x}_k, \mathbf{z}_k)$.
- Variance formulas based on inclusion prob's & linearization for given α, Q, allow estimated $t^*_{\mathbf{x}}$
- New method does not dramatically change estimates but allows approx. calibration to enhance consistency when some calibration totals are off.

Superpopulations and Large-Sample Consistency

`Superpopulation model:` sequence of large frames, $N = |\mathcal{U}| \to \infty$

Populations non-random, r_k indep. $\{0, 1\}$ resp. indicators

Design and r model Consistency for survey est's based on 2-stage *and* 1-stage methods when

(i) $1/P(r_k = 1) \equiv 1 + \lambda' \mathbf{x}_k$, (ii) $\lim \frac{\sqrt{n}}{N} \begin{pmatrix} t^*_{\mathbf{x}} - \sum_{\mathcal{U}} \mathbf{x}_k \\ t^*_{\mathbf{z}} - \sum_{\mathcal{U}} \mathbf{z}_k \end{pmatrix}$ bdd.

(iii) $Q(1 + \lambda' \mathbf{x}_k) = 0$ for all but a negligible fraction of $k \in \mathcal{U}$.

But otherwise: consistency depends on (ii), (iii) along with 'model-based' assumption that for some β the residuals $y_k - \beta' \mathbf{z}_k$ are approx. orthogonal to lin. comb's of $\mathbf{z}_k$ entries.

SIPP Background

Stratified national Survey on Income & Program Participation

217 Strata, of which 112 are Self-Representing (SR)

Longitudinal design, consider only 94444 Wave 1 responders.

Nonresponse adjustment by 149 demographic & economic Adjustment Cells; here $t^*_{\mathbf{x},j} = (N/\sum_{\mathcal{S}} r_k\, w^o_k)\ \sum_{C_j \cap \mathcal{S}} r_k\, w^o_k$.

Population controls to updated census estimates of 126 linearly independent demographic-cell population totals.

Numerical Results from Adjusting SIPP 96 Weights

Multipliers from alpha=1 Adjustments

	Min.	Q1	Median	Mean	Q3	Max.
Lambda	0.079	0.177	0.216	0.200	0.231	0.256
Mu	-1.126	-0.025	0.048	0.032	0.114	0.723
Final/Base	0.681	1.097	1.184	1.212	1.285	3.773

Multipliers from alpha=100 Adjustments

	Min.	Q1	Median	Mean	Q3	Max.
Lambda	-0.355	-0.052	0.292	0.171	0.335	0.404
Mu	-0.859	0.028	0.187	0.174	0.379	1.773
Final/Base	0.622	1.093	1.193	1.208	1.298	4.076

Correlation between Calibrated (pop-controlled) and new weights is : 0.995 for $\alpha = 1$ and 0.968 for $\alpha = 100$.

SIPP 96 Est'd Totals & SE's, in 1000's

	Totals			Std Errors		
Item	2stg	$\alpha = 1$	$\alpha = 100$	VPLX	$\alpha = 1$	$\alpha = 100$
FOODST	27541	27454	26930	687	318	301
SOCSEC	36994	37071	37240	470	157	157
HEINS	194216	194475	195030	1625	439	423
POV	41951	41978	41475	747	360	357
EMP	190871	190733	190097	1477	255	240

1-Stg SE's reflect assumed **known** nonresp. adj. totals $t^*_{\mathbf{x}}$

SE's increase 0–50% when $t^*_{\mathbf{x}}$ estimated

Discussion / Summary

- New single-stage weight adjustment methods allow design-consistent survey estimators on same basis as standard two-stage methods. (Need calibration totals and nonresponse model to be valid.)

- Incorporating penalty function for large & small weights into the single-stage adjustments enhances numerical stability when calibration totals and nonresponse mode are incorrect.

- The new methods with large α allow slightly greater ‘model-based’ protection against incorrect model and pop-controls.

- Generalization to other G_j loss-functions including those associated with raking is a topic of further research.

References

- Chang, T. and Kott, P. (2008), *Biometrika*
- Deville, J.-C. and Särndal, C.-E. (1992), Calibration estimators in survey sampling, *JASA*.
- Fuller, W. (2002), Regression estimation for survey samples. *Surv. Meth.*
- Oh, H. and Scheuren, F. (1983) Weighting adjustment for unit nonresponse.
- Särndal, C.-E. and Lundström, S. (2005) *Estimation in Surveys with Nonresponse*: Wiley.
- Singh, A. and Mohl, C. (1996), *Surv. Meth.*.
- Théberge, A. (2000) Calibration and restricted weights. *Surv. Meth.*