Combining the Nonresponse and Calibration Adjustments in Weighting

Stephen Ash
U.S. Census Bureau, 4700 Silverhill Road, Washington D.C. 20233

Abstract

There are two basic methods for the adjustment of non-interviews. In both you adjust the interviews for the non-interviews by applying a factor that is sometimes described as the inverse of the probability of completing an interview. One way to calculate the factor is directly as the ratio of the weighted count of eligible units (both completed interviews and non-interviews) with the weighted count of completed interviews. This is often done within cells to reduce bias, where cells are groups of similar units.

A different method is to model the probability of completing an interview using a general linear model. One problem with a model is that the results may not be tied to the weighted totals of eligible units that includes both the completed interviews and non-interviews. That is, the resultant probability may produce weighted counts that are less than or greater than the weighted total of eligible units.

This paper discusses using a logistic regression model that is constrained: either to the weighted totals of the eligible units, totals of the variables used in the model, or totals of other variables that normally would be used in calibration or a ratio adjustment.

Key Words: calibration, noninterview adjustment, logistic regression.

1. Introduction

This paper considers using a propensity model to account for noninterviews in survey weights. We add calibration constraints to the optimization so that the weighting adjustment accounts for noninterviews and is calibrated to known totals. To deal with large weights that can result from a model, we consider inequality constraints and a modified link function to limit the magnitude of the weights.

Our motivation is two-fold. First, can we simplify the weighting and have one adjustment instead of two separate adjustments. A single adjustment would be operationally easier. Second, would a single weighting adjustment better reduce nonresponse bias. Here, we assume that if two weighting adjustments are used, the impact of the first adjustment is diminished by the application of second.

Several authors have considered using calibration and/or propensity models to account for nonresponse in surveys. Folsom (1991), Folsom and Singh (2000) adjust for nonresponse with a model and calibrate to the variables used in the model. Bethlehem (1988) and Lündstrom and Särndal (1999) and Särndal and Lündstrom (2005) have suggested calibration as a method of adjusting for nonresponse. Kim and Kim (2007) and Kim and Riddles (2012) have explored logistic regression models with calibration and Kott (2004, 2006), Chang and Kott (2008), Kott and Chang (2010), Kott and Liao

(2012) have a series of papers that examine calibration as a means of accounting for nonresponse.

1.1 Notation

First we define *s* as the sample and *r* as the set of respondents from the sample ($r \subseteq s$). We define I_k as a Bernoulli random variable that represents that unit *k* is in the sample, i.e.,

$$I_k = \begin{cases} 1, & k \in s \\ 0, & k \notin s \end{cases}.$$

We define the probability of selection for unit *k* as $\pi_k = E(I_k) = P(I_k = 1)$ and the design weight as $d_k = \pi_k^{-1}$. We also define R_k as a Bernoulli random variable that represents that unit *k* completes an interview, i.e.,

$$R_k = \begin{cases} 1, & k \in r \\ 0, & k \notin r \end{cases}.$$

Then in terms of R_k and I_k , we define the probability of the k^{th} unit completing an interview given that it is in the sample as $\theta_k = E(R_k \mid I_k = 1) = P(k \in r \mid k \in s)$. The parameter θ_k , is often called the "response probability" or the probability of completing an interview.

1.2 Review of Weighting Class Method

The method often used to adjust survey weights for noninterviews is the weighting class method. Auxiliary information $\mathbf{x}_k$ is available for all sample units and has some association with response R_k. Groups of sample units, often called cells, with similar probabilities of response θ_k are defined using $\mathbf{x}_k$. Within each cell, an estimate of θ_k is calculated as the ratio of weighted completed interviews or respondents with the sum of weighted completed interviews and noninterviews, i.e., $\theta_k = \sum_{k \in r} d_k / \sum_{k \in s} d_k$.

1.3 Propensity Methods

An alternative method for estimating θ_k is to model it directly with a propensity or logistic regression model. Here instead of using the variables $\mathbf{x}_k$ to form cells, we use $\mathbf{x}_k$ in the model. To estimate θ_k , we calculate the maximum likelihood estimator (MLE) of θ_k . To do this we maximize the log-likelihood as in McCullagh and Nelder (1989, p. 114) as

$$\ell(\boldsymbol{\beta} \mid \mathbf{x}_k) = \sum_{k \in r} \log(\theta_k) + \sum_{k \in s-r} \log(1 - \theta_k) \qquad (1)$$

where we relate the linear model $\mathbf{x}_k' \boldsymbol{\beta}$ to the response probability θ_k with the logit or logistic link function (McCullagh and Nelder 1989, p. 108), i.e.,

$$\theta_k = \frac{e^{\mathbf{x}_k'\boldsymbol{\beta}}}{1+e^{\mathbf{x}_k'\boldsymbol{\beta}}} . \qquad (2)$$

We know that $0 \le \theta_k \le 1$ and thereby $1 \le \frac{1}{\theta_k} \le +\infty$ since $-\infty \le e^{\mathbf{x}_k'\boldsymbol{\beta}} \le +\infty$.

The logit link is one of several link functions we can use with a binary response variable. Figure 1 is a graph of the three common link functions. The horizontal axis is the possible values of the link function on the interval (0,1) and the vertical axis is the value of the linear model $\mathbf{x}_k'\boldsymbol{\beta}$.

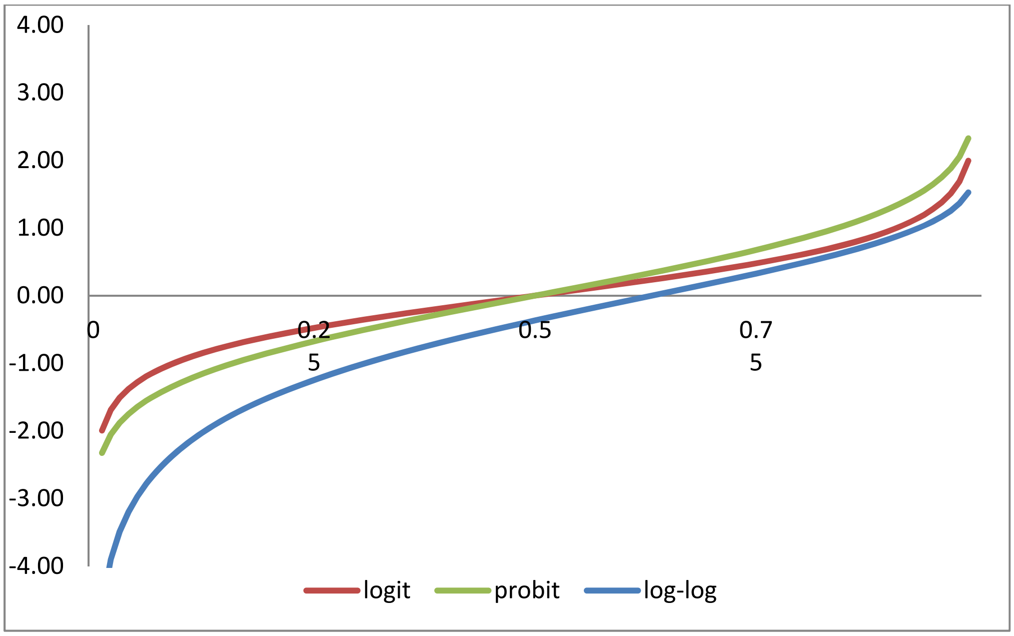

Figure 1: Link functions for binary response variables

Here $\mathbf{x}_k = [x_{k1}, x_{k2}, ..., x_{kj}, ..., x_{kp}]'$ is a p x 1 vector of auxiliary information that is associated with response/nonresponse for unit k. Also, $\boldsymbol{\beta} = [\beta_1, \beta_2, ..., \beta_j, ..., \beta_p]'$ is the p × 1 parameter vector.

2. Proposed Propensity Method

In this section, we show how we can add calibration constraints to a logistic regression model. We also discuss how to limit the extreme values predicted by the model by using either inequality constraints or a modified link function.

2.1 Noninterview Adjustments that Happen to be Calibrated

One of the properties of the weighting class method is that within cells

$$\sum_{k \in r} \frac{d_k}{\theta_k} = \sum_{k \in s} d_k . \qquad (3)$$

Having the sum of the weights before and after the nonresponse adjustment be equal to each other is not necessary, but many people are comfortable with this result. With a propensity model, this constraint is not usually met and the sum for the two totals can vary considerably. We can add (3) as a constraint to the solution of the optimization of (1) and then the solution finds the MLE that is also consistent with (3). We can generalize to calibration constraint involving a general auxiliary variable $\mathbf{z}_k^0$ in three different ways. We can calibrate the estimated totals of $\mathbf{z}_k^0$ using only the respondents to estimated totals of $\mathbf{z}_k^0$ using only the respondents, i.e.,

$$\sum_{k\in r}\frac{d_k}{\theta_k}\mathbf{z}_k^0=\sum_{k\in s}d_k\mathbf{z}_k^0 \tag{4}$$

or we can calibrate the estimated totals of $\mathbf{z}_k^0$ to the calculated known total of $\mathbf{z}_k^0$, i.e.,

$$\sum_{k\in r}\frac{d_k}{\theta_k}\mathbf{z}_k^0=\sum_{k\in U}\mathbf{z}_k^0 \tag{5}$$

or we can calibrate the estimated totals of $\mathbf{z}_k^0$ to a given total of $\mathbf{z}_k^0$, i.e.,

$$\sum_{k\in r}\frac{d_k}{\theta_k}\mathbf{z}_k^0=\mathbf{T}_{\mathbf{z}}\,. \tag{6}$$

The differences between the calibration constraints (4), (5), and (6) are how $\mathbf{z}_k^0$ is known. The calibration constraint (4) says that we know $\mathbf{z}_k^0$ for all sample units – both respondents and noninterviews. The calibration constraint (5) says that we know $\mathbf{z}_k^0$ for all units in the universe of interest and (6) refers to a known total or benchmark that is given where we do not need to know $\mathbf{z}_k^0$ for all units in the universe of interest. See also Särndal and Lündstrom (2005; section 6.2) for further discussion on the types of known totals.

To deal with the problem of extreme values θ_k, we can constrain the factors directly in the optimization of (1). This means that we add the inequality constraint $\theta_k, \geq b_k$ to the optimization problem for each $k\in r$. This insures that the estimated response probabilities do not become too small so that the adjustment factor is too large and thereby becomes an influential observation with respect to the variance. In the sections that follow, we will refer to this type of constraint as the extreme value constraint (EVC). To solve the optimization problem – maximizing (1) with the inequality constraint requires the estimation of a parameter for each $k\in r$. Computationally, the estimation of many parameters becomes burdensome. The next section considers an alternative solution for applying the EVC.

2.2 Methods with Built-in Bounds on the Noninterview Factor

To deal with the problem of estimating small values of θ_k, we can modify the link function so that it does not have extreme values. We define the probability of response as

$$\tilde{\theta}_k = \frac{(1/b_k) + (1/a_k)e^{\mathbf{x}_k'\boldsymbol{\beta}}}{1 + e^{\mathbf{x}_k'\boldsymbol{\beta}}}, \qquad (7)$$

With this new link function, we know that $0 \le \frac{1}{b_k} \le \tilde{\theta}_k \le \frac{1}{a_k} \le 1$ and thereby

$1 \le a_k \le \frac{1}{\tilde{\theta}_k} \le b_k \le +\infty$ since $-\infty \le \mathbf{x}_k'\boldsymbol{\beta} \le +\infty$. We note that θ_k is a special case of $\tilde{\theta}_k$, when $a_k = 1$ and $b_k = \infty$, $\forall k \in s$.

Figure 2 is similar to Figure 1, but it shows the values of the proposed $\tilde{\theta}_k$ for $a_k = 1$ and varying values of $b_k = \infty$, 100, 10, 5, and 4.

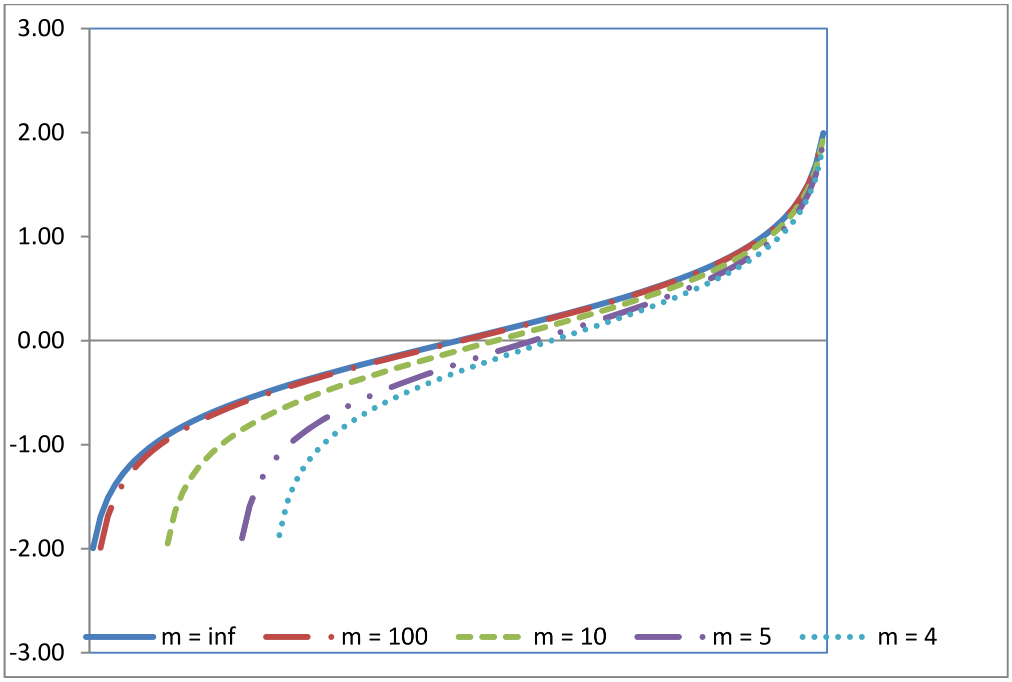

Figure 2: $\tilde{\theta}_k$ for Varying Values of b_k

Using $\tilde{\theta}_k$ instead of θ_k, we can deal with extreme values as in methods 2 and 3 but have less parameters in our objective function to minimize.

2.3 Calibration Adjustments that Happen to Account for Noninterviews

We now discuss two alternative methods that are similar to the methods already discussed that we will include in our comparison demonstration. Both methods use θ_k, as defined in (1) and both apply calibration constraints. The first is a special case of the second, but the two methods involve different solutions.

2.3.1 Folsom's Method (1991)
From Folsom (1991) and mentioned in Kim and Riddles (2012), find the parameter $\boldsymbol{\beta}$ with the calibration constraint

$$\sum_{k \in r} \frac{d_k \mathbf{x}_k}{\theta_k} = \sum_{k \in s} d_k \mathbf{x}_k \quad \text{and} \quad \sum_{k \in r} \frac{d_k}{\theta_k} = \sum_{k \in s} d_k \,. \tag{8}$$

The probability of response is defined as $\theta_k^{-1} = 1 + e^{-(\boldsymbol{\alpha} + \mathbf{x}_k' \boldsymbol{\beta})}$ which is the same as (2) except it includes an explicit intercept term α in the linear model. Folsom (1991) shows how (8) can be alternatively expressed as the *p* set of equations for which we solve for $\boldsymbol{\beta}$

$$\mathbf{U} = \sum_{k \in r} d_k e^{-(\mathbf{x}_k - \bar{\mathbf{x}}_{s\text{-}r})' \boldsymbol{\beta}} \left(\mathbf{x}_k - \bar{\mathbf{x}}_{s\text{-}r} \right) = \mathbf{0}$$

where the mean of the noninterviews is $\bar{\mathbf{x}}_{s\text{-}r} = \sum_{k \in s-r} d_k \mathbf{x}_k \Big/ \sum_{k \in s-r} d_k$ and the $p \times 1$ estimating equations are $\hat{\boldsymbol{\beta}}^{(a+1)} = \hat{\boldsymbol{\beta}}^{(a)} - \left[\mathbf{J}\left(\hat{\boldsymbol{\beta}}^{(a)} \right) \right]^{-1} \mathbf{U}\left(\hat{\boldsymbol{\beta}}^{(a)} \right)$ and $e^{-\boldsymbol{\alpha}^{(a+1)}} = \sum_{k \in s-r} d_k \Big/ \sum_{k \in r} d_k e^{-\mathbf{x}_k' \boldsymbol{\beta}^{(a+1)}}$, where the $p \times p$ Jacobian matrix is

$$\mathbf{J}(\boldsymbol{\beta}) = \frac{\partial}{\partial \boldsymbol{\beta}} \mathbf{U} = \sum_{k \in r} d_k e^{-(\mathbf{x}_k - \bar{\mathbf{x}}_{s\text{-}r})' \boldsymbol{\beta}} \left(\mathbf{x}_k - \bar{\mathbf{x}}_{s\text{-}r} \right) \left(\mathbf{x}_k - \bar{\mathbf{x}}_{s\text{-}r} \right)' = 0$$

Note that $\mathbf{x}_k$ is used in the model and in the calibration constraint and is defined as in (4) where $\mathbf{z}_k = \mathbf{x}_k$.

2.3.1 Chang and Kott (2008)
Chang and Kott (2008), which we will refer to as C&K from this point forward, generalize Folsom's method so that it has variable $\mathbf{x}_k$ related to nonresponse and variables $\mathbf{z}_k$ that are known totals or benchmark variables. The calibration constraint is as in (5). To estimate the parameter $\boldsymbol{\beta}$, C&K note that the Taylor series expansion of $\hat{\mathbf{T}}_{\mathbf{z}} - \mathbf{T}_{\mathbf{z}}$ can be expressed as $\hat{\mathbf{T}}_{\mathbf{z}} - \mathbf{T}_{\mathbf{z}} \cong \mathbf{H}(\boldsymbol{\beta})\left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta} \right) + \boldsymbol{\varepsilon}$ where $\mathbf{H}(\boldsymbol{\beta})$ is a $p \times q$ matrix with elements

$$\frac{\partial}{\partial \beta_i} \hat{\mathbf{T}}_{\mathbf{z}_j} = -\sum_{k \in r} d_k \left(\frac{1}{e^{\mathbf{x}_k' \boldsymbol{\beta}}} \right) x_{ki} z_{kj} \,.$$

An estimator that comes directly from the Taylor series expansion is

$$\hat{\boldsymbol{\beta}}^{(a+1)} = \hat{\boldsymbol{\beta}}^{(a)} - \left[\mathbf{H}\left(\hat{\boldsymbol{\beta}}^{(a)} \right)' \mathbf{H}\left(\hat{\boldsymbol{\beta}}^{(a)} \right) \right]^{-1} \mathbf{H}\left(\hat{\boldsymbol{\beta}}^{(a)} \right)' \left(\hat{\mathbf{T}}_{\mathbf{z}}\left(\hat{\boldsymbol{\beta}}^{(a)} \right) - \mathbf{T}_{\mathbf{z}} \right)$$

C&K suggest a “partial minimization” which starts by defining $\mathbf{U} = \hat{\mathbf{T}}_{\mathbf{z}} - \mathbf{T}_{\mathbf{z}} - \mathbf{H}(\hat{\boldsymbol{\beta}})(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})$ and the estimates the parameter $\boldsymbol{\beta}$ by minimizing $\mathbf{U}'(\mathbf{v}(\mathbf{U}))\mathbf{U}$ where we treat $\mathbf{v}(\mathbf{U})$ as a constant. The updating equation is given as

$$\hat{\boldsymbol{\beta}}^{(a+1)} = \hat{\boldsymbol{\beta}}^{(a)} - \left[\mathbf{H}\left(\hat{\boldsymbol{\beta}}^{(a)}\right)' \mathbf{v}\left(\mathbf{U}\left(\hat{\boldsymbol{\beta}}^{(a)}\right)\right) \mathbf{H}\left(\hat{\boldsymbol{\beta}}^{(a)}\right)\right]^{-1} \mathbf{H}\left(\hat{\boldsymbol{\beta}}^{(a)}\right)' \mathbf{v}\left(\mathbf{U}\left(\hat{\boldsymbol{\beta}}^{(a)}\right)\right)\left(\hat{\mathbf{T}}_{\mathbf{z}}\left(\hat{\boldsymbol{\beta}}^{(a)}\right) - \mathbf{T}_{\mathbf{z}}\right).$$

C&K discuss several ways of defining the covariance matrix $\mathbf{v}(\mathbf{U})$. For our comparison, we define it similar to (5) in C&K in terms of a two-phase variance of $\hat{\mathbf{T}}_{\mathbf{z}}$ as

$$\hat{\mathbf{v}}\left(\hat{\mathbf{T}}_{\mathbf{z}}\right) = \sum_{k,\ell \in r} \frac{\pi_{k\ell} - \pi_k \pi_\ell}{\pi_{k\ell} \pi_k \pi_\ell \theta_k \theta_\ell} \mathbf{z}_k \mathbf{z}_\ell{}' + \sum_{k \in r} \frac{1-\theta_k}{\theta_k^2 \pi_k} \mathbf{z}_k \mathbf{z}_k{}' \tag{8}$$

which we alternatively express in compact matrix notation as $\hat{\mathbf{v}}\left(\hat{\mathbf{T}}_{\mathbf{z}}\right) = \mathbf{Z}'[\boldsymbol{\Pi}_1 + \boldsymbol{\Pi}_2]\mathbf{Z}$, where the $p \times n_r$ data matrix is defined as concatenation of the vector $\mathbf{z}_k$ for each $k \in r$, i.e., $\mathbf{Z} = \mathbf{z}_1 \,|\, \mathbf{z}_2 \,|\, ... \,|\, \mathbf{z}_{n_r}$ and the $n_r \times n_r$ covariance matrices for the first and second phases of the sample design are $\boldsymbol{\Pi}_1$ and $\boldsymbol{\Pi}_2$, respectively. The k^{th} row and ℓ^{th} column of $\boldsymbol{\Pi}_1$ is defined as $(\pi_{k\ell}\pi_k\pi_\ell\theta_k\theta_\ell)^{-1}(\pi_{k\ell} - \pi_k\pi_\ell)$ and we can define the second phase covariance matrix simply as $\boldsymbol{\Pi}_2 = [\mathbf{diag}(\boldsymbol{\pi} \# \boldsymbol{\theta} \# \boldsymbol{\theta})]^{-1}[\mathbf{I} - \mathbf{diag}(\boldsymbol{\theta})]$. For the purpose of the demonstration, we define the second order selection probability $\pi_{k\ell}$ for unequal *sys* from a randomly ordered list as provided by Overton (1985).

3. Demonstration of Proposed Propensity Method

This section is devoted to a simple demonstration of the five methods using a small fictitious data set. The example involves $p = 3$ variables that are expected to be related to nonresponse. We also have $q = 2$ variables that we want to constrain to known totals. The data set had 32 sample units: 10 completed interviews and 22 noninterviews. The dataset was constructed such that two sample units had that $\theta_k^{-1} > 4.0$ when θ_k is estimated from a logistic regression model using $\mathbf{x}_k$ and no constraints.

We used the Newton-Raphson method to solve all of the optimization problems. We did this by finding first- and second-order derivatives which were then used to define the updating equations. The Newton-Raphson methods that we used can be found in a general text on optimization such as Nocedal and Wright (2006).

Results are provided for all of the noninterview ajdustments discussed. The results for the MLE with the built-in bounds is reviewed separately and after all the other methods.

3.1 Constrained MLE

Tables 1 and 2 provide the results of maximizing (1) with different constraints. Table 1 models with $\mathbf{x}_k$ and calibrates on $\mathbf{z}_k$, where Table 2 models and calibrates with $\mathbf{x}_k$. The columns of Table 1 and 2 vary by the set of constraints or whether Folsom or C&K was used. The first 10 rows of Tables 1-4 show the actual noninterview factor θ_k^{-1} for the 10 completed interviews. The next three rows show the estimated model parameters beta and the final two rows show the estimated totals of on $\mathbf{z}_k$. The estimated totals in blue bold indicate that the estimated totals are consistent with the known totals. When the EVC was applied, it constrained the factors such that $\theta_k^{-1} \leq 4.0$.

In Table 1, sample units 6 and 7 have weights larger than 4 for all methods except when the EVC was used. We also note that the parameter estimates were all in the same neighborhood except with the CC & EVC. It seems that using both sets of constraints can lead to an odd choice of parameter estimates and weights.

Table 1: Factors for Methods Involving θ_k -- Calibrating with $\mathbf{z}_k$ and Modeling with $\mathbf{x}_k$

Sample Unit	No	CC	EVC	CC & EVC	C&K
1	2.0000	2.0000	2.0000	1.6000	2.0000
2	2.1785	1.8873	2.1625	3.9161	2.4667
3	2.1785	1.8873	2.1625	3.9161	2.4667
4	1.5707	1.1876	1.5953	1.0000	1.1237
5	3.7336	6.4308	3.5807	4.0000	5.3403
6	**4.2215**	**5.8185**	**4.0000**	**4.0000**	**7.3660**
7	**4.2215**	**5.8185**	**4.0000**	**4.0000**	**7.3660**
8	2.3237	2.1484	2.3215	2.1565	1.3661
9	2.5600	2.0189	2.5362	3.9660	1.5370
10	2.5600	2.0189	2.5362	3.9660	1.5370
β_1	-1.0056	-1.6921	-0.9480	-26.3450	-1.4680
β_2	0.7252	1.5537	0.6693	24.4250	2.4727
β_3	-0.1642	0.1196	-0.1506	-4.9347	-0.3830
Est. Total z_1	19,555	**21,325**	19,141	**21,325**	**21,325**
Est. Total z_2	25,120	**28,675**	24,483	**28,675**	**28,675**

Table 2 adds Folsom's method so we need to both model and calibrate with x_k. Three of the methods have the same exact solution: CC, Folsom, and C&K. We do not know why this is so, but we think it is because using the same auxiliary vector in both the model and the CC locks the solution.

Table 2: Factors for Different Involving θ_k -- Calibrating with $\mathbf{x}_k$ and Modeling with $\mathbf{x}_k$

Sample Unit	No	CC	EVC	CC & EVC	Folsom	C&K
1	2.0000	2.0000	2.0000	1.6000	2.0000	2.0000
2	2.1785	2.5177	2.1625	2.5336	2.5177	2.5177
3	2.1785	2.5177	2.1625	2.5336	2.5177	2.5177
4	1.5707	1.7484	1.5953	1.4573	1.7484	1.7484
5	3.7336	2.8601	3.5807	2.9352	2.8601	2.8601
6	**4.2215**	**3.8231**	**4.0000**	**3.6513**	**3.8231**	**3.8231**
7	**4.2215**	**3.8231**	**4.0000**	**3.6513**	**3.8231**	**3.8231**
8	2.3237	1.9172	2.3215	1.7144	1.9172	1.9172
9	2.5600	2.3921	2.5362	2.6999	2.3921	2.3921
10	2.5600	2.3921	2.5362	2.6999	2.3921	2.3921
β_1	-1.0056	-0.6206	-0.9480	-1.9838	-0.6207	-0.6207
β_2	0.7252	0.7071	0.6693	1.7605	0.7071	0.7071
β_3	-0.1642	-0.1472	-0.1506	-1.4311	-0.1472	-0.1472
Est. Total x_1	4,974	**4,375**	4,823	**4,375**	**4,375**	**4,375**
Est. Total x_2	2,602	**2,400**	2,596	**2,400**	**2,400**	**2,400**
Est. Total x_3	4,437	**4,300**	4,316	**4,300**	**4,300**	**4,300**

3.2 Constrained Logit Model

Tables 3 and 4 provide the results of applying the constrained logit model of section 3.2. Table 3 only includes the constraint for extreme values and Table 4 constrains for extreme values and the calibration constrains on $\mathbf{z}_k$. The columns of Table 3 and 4 vary the values of b_k such that $\tilde{\theta}_k^{-1} \leq b_k$.

Table 3: Factors for $\tilde{\theta}_k$ and the Constrained Logit Model

Sample Unit	*Constrained so that $\tilde{\theta}_k^{-1} \leq \ldots$*				
	∞	100	10	5	4
1	2.0000	1.9802	1.8182	1.6667	1.6000
2	2.1785	2.1700	2.1158	3.5000	3.5000
3	2.1785	2.1700	2.1158	3.5000	3.5000
4	1.5707	1.5648	1.5033	1.0000	1.0000
5	3.7336	3.7156	3.6357	5.0000	4.0000
6	**4.2215**	**4.2305**	**4.3256**	**5.0000**	**4.0000**
7	**4.2215**	**4.2305**	**4.3256**	**5.0000**	**4.0000**
8	2.3237	2.3217	2.3295	5.0000	4.0000
9	2.5600	2.5766	2.7728	5.0000	4.0000
10	2.5600	2.5766	2.7728	5.0000	4.0000
β_1	-1.0056	-1.0369	-1.4210	-46.6732	-50.4049
β_2	0.7252	0.7345	0.8710	27.2002	26.7462
β_3	-0.1642	-0.1790	-0.3475	-2.1203	-2.9957
Est. Total z_1	19,555	19,561	19,790	30,100	21,325
Est. Total z_2	25,120	25,141	25,604	38,475	28,675

We see in Table 3 that applying more stringent constraints on the solution had an adverse impact – the

Table 4: Factors for $\tilde{\theta}_k$ and Constrained Logit Model with Calibration Constraints

	Constrained so that $\tilde{\theta}_k^{-1} \leq \ldots$				
Sample Unit	∞	100	10	5	4
1	2.0000	1.9802	1.8182	1.6667	1.6000
2	1.8873	2.5066	1.9555	2.2294	3.9161
3	1.8873	2.5066	1.9555	2.2294	3.9161
4	1.1876	1.7418	1.1231	1.0000	1.0000
5	6.4308	2.8464	5.7003	5.0000	4.0000
6	**5.8185**	**3.8247**	**6.0843**	**5.0000**	**4.0000**
7	**5.8185**	**3.8247**	**6.0843**	**5.0000**	**4.0000**
8	2.1484	1.9108	2.0189	2.3265	2.1565
9	2.0189	2.4005	2.1851	3.0960	3.9660
10	2.0189	2.4005	2.1851	3.0960	3.9660
β_1	-1.6921	-0.6421	-2.3917	-10.3212	-51.9975
β_2	1.5537	0.7163	2.1474	9.4127	51.0775
β_3	0.1196	-0.4352	-0.1721	-0.7972	-4.9347
Est. Total z_1	**21,325**	**21,325**	**21,325**	**21,325**	**21,325**
Est. Total z_2	**28,675**	**28,675**	**28,675**	**28,675**	**28,675**

In Table 4, we again see that that applying more stringent constraints on the solution had an adverse impact, but the calibration constrains seemed to help the optimization problem by helping the updating equations keep the solution close to the.

Table 2 adds Folsom's method so we need to both model and calibrate with $\mathbf{x}_k$. Two things are interesting. First, the additional CC seems to help the

4. Conclusions

Our goal was to combine the noninterview adjustment and calibration adjustment and thereby have a single weighting adjustment. We have shown how a propensity model can be applied with different types of calibration constraints. The calibrated propensity adjustment makes sense because it finds weighting factors that satisfy the calibration constraints and makes the most sense as a noninterview adjustment.

Any views expressed are those of the author(s) and not necessarily those of the U.S. Census Bureau.

References

Bethlehem, J.G. (1988). "Reduction of Nonresponse Bias Through Regression Estimation," Journal of Official Statistics, 4, 3, 251-260.

Chang, T. and Kott, P.S. (2008). "Using calibration weighting to adjust for nonresponse under a plausible model," Biometrika, 95, 3, 555-571.

Folsom, R.E. (1991). "Exponential and Logistic Weight Adjustments for Sampling and Nonresponse Error Reduction," Proceedings of the Section on Survey Research Methods.

Folsom, R. E., and Singh, A.C. (2000). "The generalized exponential model for sampling weight calibration for extreme values, nonresponse, and poststratification", Proceedings of the Survey Methods Section, 47-58.

Kim, J.K. and Kim, J.J. (2007). "Nonresponse weighting adjustment using estimated response probability," Canadian Journal of Statistics, 35, 4, 501-514.

Kim, J.K. and Riddles, M.K. (2012). "Some theory for propensity-score adjustment estimators in survey sampling," Survey Methodology, 38, 2, 157-165.

Kott, P.S. (2006). "Using Calibration Weighting to Adjust for Nonresponse and Coverage Errors," Survey Methodology, 32, 2, 133-142.

Kott, P.S. and Chang, T. (2010). "Using calibration weighting to adjust for Nonignorable Unit Nonresponse," Journal of the American Statistical Association, 105, 491, 1265-1275.

Kott, P.S. and Liao, D. (2012). "Providing double protection for unit nonresponse with nonlinear calibration-weighting routine," Survey Research Methods, 6, 2, 105-111.

Lündstrom, S. and Särndal, C.-E. (1999). "Calibration as a standard method for treatment of nonresponse," Journal of Official Statistics, 15, 305-327.

McCullagh, P. and Nelder, J.A. (1989). *Generalized Linear Models*, 2nd Edition, Chapman Hall/CRC.

Nocedal, J. and Wright, S.J. (2006). *Numerical Optimization*, 2nd Edition, Springer-Verlag.

Overton, W.S. (1985). "A Sampling Plan for Streams in the National Stream Survey," Technical Report 114, Department of Statistics, Oregon State University, Corvallis, Oregon, 97331.

Särndal, C.-E. and Lündstrom, S. (2005). *Estimation in Surveys with Nonresponse*, John Wiley & Sons.