Optimal Estimators for Two-Phase Sample Designs

Stephen Ash
U.S. Census Bureau, 4600 Silver Hill Road, Room 7H141, Washington D.C. 20233

Abstract

This paper extends the general result of the optimal estimator (Montanari 1987) to the two-phase sample design. We do this for several different combinations of auxiliary variables that can be available with the two-phase sample design. For each combination of auxiliary variables, we define an appropriate model and associated estimator. We then find parameters which minimize the survey variances, i.e., which make the estimators "optimal."

Keywords: Auxiliary information, generalized regression estimation, model-assisted.

1. Introduction

The application of auxiliary information in survey estimation is important because its role in reducing the variance of survey estimators. This paper discusses a general method to improve estimators of two-phase surveys when auxiliary information is available. The paper extends the optimal estimator (Montanari 1987) to the two-phase sample design (s.d.). Within the context of the generalized regression estimator [GRE] (Cassel et.al. 1976) for the one-stage s.d., there are several choices for defining the regression parameters. The optimal estimator is defined as the solution that minimizes the sample variance.

We begin by discussing similar research and developing notation. Next we discuss the types of auxiliary variables and the construction of the estimators that use them. Once the estimators are developed we discuss their interpretation, bias and estimation. The paper concludes with an empirical example that compares the optimal estimators we propose with their calibration counterparts.

1.1 Background

Solutions to the regression parameters of the general regression estimator can be found in several ways.

Calibration estimators that take advantage of available auxiliary information for two-phase s.d.s are discussed by Särndal et.al. (1992; p. 359), Hidroglou and Särndal (1995, 1998), Dupont (1998), and Ash (2003). Estevao and Särndal (2002), which we will refer to as E&S, apply calibration with auxiliary variables for ten different combinations of the two-phase s.d. Wu and Luan (2003) present optimal calibration estimators for a two-phase s.d. as calibrating with known totals of a modeled value of the variable of interest.

Kiregyera (1980, 1984), Roy (2003) and others use ratio or chain estimators to take advantage of auxiliary information in conjunction with two-phase s.d.s. For an overview of chain estimators see Singh, Upadhyaya, and Chandra (2004).

Tripathi and Ahmed (1995) discuss the optimal estimator for the case we will later call (2c) or "All and Second". Sahoo and Panda (1999) discuss the optimal estimator for the case we will refer to as (2b) or "Overall and Second". Ahmad et.al. (1995) also discuss the optimal estimator in the context of multi-phase sampling.

1.2 Notation

Let U be the universe which includes N units. We index the units of U as k or ℓ and select a sample s_1 of n_1 units from U. The first phase selection probabilities are defined as $\pi_{1k} = P(k \in s_1)$ and $\pi_{1k\ell} = P(k \,\&\, \ell \in s_1)$. Similarly we select the second phase sample s_2 of n_2 units from s_1 and define the selection probabilities for the second phase as $\pi_{2k} = P(k \in s_2 \mid k \in s_1)$ and $\pi_{2k\ell} = P(k \,\&\, \ell \in s_2 \mid k \,\&\, \ell \in s_1)$. We are interested in estimating the total of some variable of interest y, i.e., $T_y = \sum_{k \in U} y_k$ which we can estimate as $\hat{T}_y = \sum_{k \in s} \pi_{1k}^{-1} \pi_{2k}^{-1} y_k$.

Since our discussion concerns two-phase s.d.s we have two indexes for the expectation, variance and covariance. Let E_1, v_1 and cov_1, be the expectation, variance and covariance, respectively, of the first stage s.d. Similarly let E_2, v_2 and cov_2 be the expectation, variance and covariance, respectively, of the second phase s.d. If a subscript is not used, then the expectation, variance or covariance is with respect to both phases or simply the overall s.d. We also assume that the inverse of the probabilities of selection can be used to produce unbiased estimators of the totals.

1.3 Auxiliary Information

In general, auxiliary information is something we know for all units of the universe. Similarly, an auxiliary variable (AV) is a variable that is known for all units in the universe. We denote AVs as either **x**, **z**, or **v**. Hidroglou and Särndal (1995) characterized auxiliary information for a two-phase s.d. as being specific to one of three different levels: (1) $\sum_{k \in U} \mathbf{x}_k$ is known , (2) $\mathbf{x}_k$ is known for all $k \in s_1$ or (3) $\mathbf{x}_k$ is known for all $k \in s$. A given AV can be known for any combination of these three levels. The combinations we examine are listed in Table 1.

Table 1: Combinations of Auxiliary Information

Code	Name	$\sum_{k \in U} \mathbf{x}_k$ is known	$\mathbf{x}_k$ is known for all $k \in s_1$	$\mathbf{x}_k$ is known for all $k \in s$
111	All	1	1	1
101	Overall	1	0	1
110	First	1	1	0
011	Second	0	1	1

As in E&S, the coding in the first column of Table 1 refers to the availability of each of the three levels of auxiliary information, where 1 indicates "yes", the auxiliary information is available, and 0 indicates that "no", it is not. The second column is a shorthand name for the combination that we sometimes use. For example, if $\mathbf{x}_k$ is an overall (or equivalently 101) AV of a two-phase s.d., we know $\sum_{k \in U} \mathbf{x}_k$ and $\mathbf{x}_k$ for all $k \in s$.

Table 2: Two-Phase Sample Designs that Use Auxiliary Information

Case	Code	Label from E&S
(0) None	000	C4
(1a) All	111	A2
(1b) Overall	101	B2
(1c) First	110	A4
(1d) Second	011 [000/011, 011/011]	C2 [C3, C1]
(2a) All & First	111/110	
(2b) Overall & First	101/011	
(2c) All & Second	111/011	A1
(2d) Overall & Second	101/011	B1
(2e) First and Second	110/011	A3

Several combinations of auxiliary information are not mentioned in Table 2 – the trivial case of having no auxiliary information (000) and having an AV with only one level (001, 010 and 100). These cases cannot lead to regression estimators because an AV is only useful in constructing a GRE when at least two levels are known.

With multiple AVs there are several combinations of AVs we can consider. These combinations, or cases as E&S referred to them, are listed in Table 2.
We only list one combination involving three different AVs of different types – the complete list is too long and we only provide the solution for this case. The third column of Table 2 identifies the ten cases from E&S and relates them to the organization of our paper.

2. Constructing the Estimators

To begin the construction of our two-phase optimal estimator, we first assume a model μ_k. The model is based on the available AVs. In all the models we assume that the random error of the model, denoted as e_k, has a mean of zero and a constant positive variance. The general form of the GRE (Cassel et.al.1976) for the two-phase s.d. is $\hat{T}_{y,reg} = \sum_U \mu_k + \sum_{s_2} \pi_{1k}^{-1}\pi_{2k}^{-1} e_k$ where $e_k = y_k - \mu_k$. We assume that μ_k is a linear function of the AVs available.

Table 3 specifies the model for several cases of the two-phase s.d. The first column of Table 3 identifies the estimators and the second column defines the appropriate model. The GREs based on the model are listed in the third column of Table 3. This construction makes the estimator model-assisted, in that we use a regression model, but the estimator remains design unbiased. With the regression estimators established, the third column defines the regression residuals e_{1k} and e_{2k} for each case.

The next task is the choice of regression parameters. One solution is to find parameters that minimize the sample variance as Montanari (1987) suggested for the one-stage s.d. To define our optimal regression parameters we mimic the result of Lemma 1 of Montanari (1987; p.196) and as in Montanari we get estimators which have recognizable values and interesting interpretations.

The variance of $\hat{T}_{y,reg}$ may be written as $\operatorname{var}\left(\hat{T}_{y,reg}\right) = V_1 + V_2$ where Axelson (2000) notes $V_1 = \operatorname{var}_1\left(E_2\left(\hat{T}_{y,reg}\right)\middle| s_1\right)$ and $V_2 = E_1\left(\operatorname{var}_2\left(\hat{T}_{y,reg}\right)\middle| s_1\right)$ are often referred to as the first and second phase variance components. Using the regression residuals e_{1k} and e_{2k}, an approximate variance for the two-phase s.d. (Särndal and Swensson 1987), (Särndal et.al. 1992; eq. 9.3.6) is

$$v\left(\hat{T}_{y,reg}\right) = \sum\sum_{k\ell \in U} \left(\pi_{1k\ell} - \pi_{1k}\pi_{1\ell}\right)\left(e_{1k}/\pi_{1k}\right)\left(e_{1\ell}/\pi_{1\ell}\right)$$

$$+ E_1\left[\sum\sum_{k\ell \in s_1} \left(\pi_{2k\ell} - \pi_{2k}\pi_{2\ell}\right)\left(e_{2k}/\pi_{1k}\pi_{2k}\right)\left(e_{2\ell}/\pi_{1\ell}\,\pi_{2\ell}\right)\right].$$

The next result will help simplify the solution of the regression parameters by allowing us to express the two-phase variance as a quadratic form.

Result:

$$E_1\left[\sum\sum_{k\ell\in s}\left(\pi_{2k\ell}-\pi_{2k}\pi_{2\ell}\right)\left(e_{2k}/\pi_{2k}\right)\left(e_{2\ell}/\pi_{2\ell}\right)\right]=\mathbf{e}_2'\mathbf{W}_2\mathbf{e}_2$$

where $\mathbf{W}_2$ is an $N \times N$ symmetric matrix and $\mathbf{e}_2$ is the N x 1 vector of second phase residuals e_{2k}.

We begin by defining $\mathbf{W}_{s_1}$ as an $N \times N$ diagonal matrix which we define for each sample s_1. The elements of $\mathbf{W}_{s_1}$ are $\pi_{2k}^{-1}\pi_{2\ell}^{-1}\left(\pi_{2k\ell}-\pi_{2k}\pi_{2\ell}\right)$ if k and $\ell \in s_1$, and zero otherwise. Then

$$E_1\left[\sum\sum_{k\ell\in s}\left(\pi_{2k\ell}-\pi_{2k}\pi_{2\ell}\right)\left(e_{2k}/\pi_{2k}\right)\left(e_{2\ell}/\pi_{2\ell}\right)\right].$$

$$=E_1\left(\mathbf{e}_2'\mathbf{W}_{s_1}\mathbf{e}_2\right)=\mathbf{e}_2'\left[E_1\left(\mathbf{W}_{s_1}\right)\right]\mathbf{e}_2$$

Given every $\mathbf{W}_{s_1}$ is symmetrical and $E_1\left[\mathbf{W}_{s_1}\right]=\sum_{s_1}p(s_1)\mathbf{W}_{s_1}$, the result follows because the sum of diagonal matrices is symmetrical.

With this result we can write the two-phase variance without the first phase expectation as $v\left(\hat{T}_{y,reg}\right)=\mathbf{e}_1'\mathbf{W}_1\mathbf{e}_1+\mathbf{e}_2'\mathbf{W}_2\mathbf{e}_2$ where $\mathbf{W}_1$ is a $N \times N$ matrix with $\left(\pi_{1k\ell}-\pi_{1k}\pi_{1\ell}\right)\pi_{1k}^{-1}\pi_{1\ell}^{-1}$ $\left(\pi_{1kk}\equiv\pi_{1k}\right)$. Now the regression parameters that minimize the sample variance are the solution to $\frac{\partial}{\partial\beta_x}v\left(\hat{T}_y\right)=0$.

Example: If we have an AV that is 101, we know $\sum_{k\in U}\mathbf{x}_k$ and $\mathbf{x}_k$ is known for all $k \in s_2$. We can use the model $\mu_k=\mathbf{x}_k\beta_{\mathbf{x}}+e_k$ and write the regression estimator as $\hat{T}_y^{(1b)}=\mathbf{T}_{\mathbf{x}}'\beta_{\mathbf{x}}+\sum_{i\in s}\pi_{1k}^{-1}\pi_{2k}^{-1}\left(y_k-\mathbf{x}_k'\beta_{\mathbf{x}}\right)$ which has estimated regression residuals $\hat{e}_{1k}=y_k-\mathbf{x}_k'\hat{\beta}_{\mathbf{x}}$ and $\hat{e}_{2k}=y_k-\mathbf{x}_k'\hat{\beta}_{\mathbf{x}}$. With the residuals we can write the variance as the quadratic form

$$v\left(\hat{T}_y^{(1b)}\right)=\left(\mathbf{Y}-\mathbf{X}\beta_{\mathbf{x}}\right)'\mathbf{W_1}\left(\mathbf{Y}-\mathbf{X}\beta_{\mathbf{x}}\right)+\left(\mathbf{Y}-\mathbf{X}\beta_{\mathbf{x}}\right)'\mathbf{W_2}\left(\mathbf{Y}-\mathbf{X}\beta_{\mathbf{x}}\right)$$

, where $\mathbf{X}$ is defined as the $N \times p$ matrix of $\mathbf{x}_k$ and $\mathbf{Y}$ are defined as the N x 1 matrix of y_k. Then solving $\frac{\partial}{\partial\beta_x}v\left(\hat{T}_y^{(1)}\right)=0$ we get

$\beta_{\mathbf{x}}=\left[\mathbf{X}'\mathbf{W}_1\mathbf{X}+\mathbf{X}'\mathbf{W}_2\mathbf{X}\right]^{-1}\left[\mathbf{X}'\mathbf{W}_1\mathbf{Y}+\mathbf{X}'\mathbf{W}_2\mathbf{Y}\right]$. Within the result we recognize the two-stage variance and covariances, i.e.,

$$\beta_{\mathbf{x}}=\left[v_1\left(\hat{\mathbf{T}}_{\mathbf{x}}\right)+v_2\left(\hat{\mathbf{T}}_{\mathbf{x}}\right)\right]^{-1}\left[\mathrm{cov}_1\left(\hat{\mathbf{T}}_{\mathbf{x}},\hat{T}_y\right)+\mathrm{cov}_2\left(\hat{\mathbf{T}}_{\mathbf{x}},\hat{T}_y\right)\right].$$

The parameters for the other cases in Table 4 were found with the same method.

2.1 Bias of Estimators

The optimal regression estimator is not unbiased because the estimator of the parameter β is nonlinear. We can express the bias of the one phase s.d. optimal estimator and this is useful to understanding the bias of the estimators (1a), (1b), (1c) and (1d).

In the one stage s.d., the optimal estimator is $\hat{T}_{y,opt}=\sum_{k\in s}\pi_k^{-1}y_k+\left(\sum_{k\in U}\mathbf{x}_k-\sum_{k\in s}\pi_k^{-1}\mathbf{x}_k\right)\beta_x$ where we estimate the regression parameter as $\hat{\beta}_{\mathbf{x}}=\left[\hat{v}\left(\hat{\mathbf{T}}_{\mathbf{x}}\right)\right]^{-1}\hat{\mathrm{cov}}\left(\hat{\mathbf{T}}_{\mathbf{x}},\hat{T}_y\right)$. We can express the bias of this estimator in terms of the first and second order terms of its Taylor series expansion, i.e.,

$$\begin{aligned}bias\left(\hat{T}_{y,opt}\right)&=E\left(\hat{T}_y-T_y\right)-E\left(\hat{\mathbf{T}}_{\mathbf{x}}-\mathbf{T}_{\mathbf{x}}\right)\left[v\left(\hat{\mathbf{T}}_{\mathbf{x}}\right)\right]^{-1}\mathrm{cov}\left(\hat{\mathbf{T}}_{\mathbf{x}},\hat{T}_y\right)\\&-2\left[v\left(\hat{\mathbf{T}}_{\mathbf{x}}\right)\right]^{-1}\mathrm{cov}\left(\hat{\mathbf{T}}_{\mathbf{x}},\hat{\mathrm{cov}}\left(\hat{\mathbf{T}}_{\mathbf{x}},\hat{T}_y\right)\right)\\&+2\,\mathrm{cov}\left(\hat{\mathbf{T}}_{\mathbf{x}},\hat{v}\left(\hat{\mathbf{T}}_{\mathbf{x}}\right)\right)\left[v\left(\hat{\mathbf{T}}_{\mathbf{x}}\right)\right]^{-2}\mathrm{cov}\left(\hat{\mathbf{T}}_{\mathbf{x}},\hat{T}_y\right)+R(3)\end{aligned}\qquad(1)$$

The first line of (1) represents the first order terms and likewise the first two terms of the second line represent the second order terms. $R(3)$ represents the third and higher order terms. To consider the bias more closely we note the following two conditions.

Condition 1: $E\left(\hat{T}_y\right)=T_y$ and $E\left(\hat{\mathbf{T}}_{\mathbf{x}}\right)=\mathbf{T}_{\mathbf{x}}$

Condition 2:

$$\left[v\left(\hat{\mathbf{T}}_{\mathbf{x}}\right)\right]^{-1}\mathrm{cov}\left(\hat{\mathbf{T}}_{\mathbf{x}},\hat{T}_y\right)=\left[\mathrm{cov}\left(\hat{\mathbf{T}}_{\mathbf{x}},\hat{v}\left(\hat{\mathbf{T}}_{\mathbf{x}}\right)\right)\right]^{-1}\mathrm{cov}\left(\hat{\mathbf{T}}_{\mathbf{x}},\hat{\mathrm{cov}}\left(\hat{\mathbf{T}}_{\mathbf{x}},\hat{T}_y\right)\right)$$

Under Condition 1, the first order terms of the bias is equal to zero and under Condition 2 the second order terms of the bias are equal to zero. Since we expect our estimators to be design unbiased, we know that the bias does not include any first order terms. We cannot claim that Condition 2 is true for all cases, because it depends on 3rd and 4th order selection probabilities which are difficult to evaluate. However when $y_k = \boldsymbol{\alpha}\,\mathbf{x}_k$, for some vector of constants $\boldsymbol{\alpha}$, we can show that Condition 2 is also true and in that case we only have third order and higher terms contributing to the bias.

Given what we have learned from Conditions 1 and 2, we say that the optimal estimator has a small bias when the estimator is design unbiased and μ_k is a strong model of y_k.

2.2 More on the Interpretation of the Optimal Regression Parameters

Similar to Montanari (1987) we can interpret the parameters of each of the optimal regression estimators as the covariance of y and the AV(s) **x**, **z** or **v** divided by the variance of the AV(s). For example, the parameter for case (1a) is $\beta_{\mathbf{x}} = \left[v\left(\hat{\mathbf{T}}_{\mathbf{x}}\right)\right]^{-1} \operatorname{cov}\left(\hat{\mathbf{T}}_{\mathbf{x}}, \hat{T}_y\right)$, a function of the overall two-phase variance and covariance of **x**.

The regression parameters for cases (2b), (2d) and (2e) do not have simple interpretations, but can be interpreted similarly. Consider case (2b) - "overall and first," where we use overall and first stage AVs. If we define the first phase correlation as $\rho_1\left(\hat{\mathbf{T}}_{\mathbf{z}}, \hat{T}_y\right) = \left[v_1\left(\hat{\mathbf{T}}_{\mathbf{z}}\right)\right]^{-1/2} \operatorname{cov}_1\left(\hat{\mathbf{T}}_{\mathbf{z}}, \hat{T}_y\right)\left[v_1\left(\hat{T}_y\right)\right]^{-1/2}$. We then can rewrite the parameter $\beta_{\mathbf{x}}$ from Table 4 as

$$\left[v_1\left(\hat{\mathbf{T}}_x\right)\left(1-\rho_1^2\left(\hat{\mathbf{T}}_x, \hat{\mathbf{T}}_z\right)\right)+v_2\left(\hat{\mathbf{T}}_x\right)\right]^{-1} \times$$
$$\left[\operatorname{cov}_1\left(\hat{\mathbf{T}}_x, \hat{T}_y\right)\left(1-\left[\rho_1\left(\hat{\mathbf{T}}_x, \hat{T}_y\right)\right]^{-1} \rho_1\left(\hat{\mathbf{T}}_x, \hat{T}_z\right)\rho_1\left(\hat{\mathbf{T}}_z, \hat{T}_y\right)\right)+\operatorname{cov}_2\left(\hat{\mathbf{T}}_x, \hat{T}_y\right)\right]$$

We now see that the denominator of $\beta_{\mathbf{x}}$ is an expression for the overall "effective" variance of **x**. We say that $v_1\left(\hat{\mathbf{T}}_{\mathbf{x}}\right)\left(1-\rho_1^2\left(\hat{\mathbf{T}}_x, \hat{\mathbf{T}}_z\right)\right)$ is the "effective" first stage variance, since if **x** has an exact linear relationship with **z**, a variable that has no first stage variance, then **x** also has no first stage variance. The numerator of $\beta_{\mathbf{x}}$ has a similar interpretation. The effective first phase covariance of **x** and y is reduced by the term $\left(1-\left[\rho_1\left(\hat{\mathbf{T}}_x, \hat{T}_y\right)\right]^{-1} \rho_1\left(\hat{\mathbf{T}}_x, \hat{T}_z\right)\rho_1\left(\hat{\mathbf{T}}_z, \hat{T}_y\right)\right)$ as **x**, **z** and y become more linearly associated with one another.

Next we interpret $\beta_{\mathbf{z}}$. We start by defining the two modified correlations $\bar{\rho}_1\left(\hat{\mathbf{T}}_{\mathbf{z}}, \hat{T}_y\right) = \left[v\left(\hat{\mathbf{T}}_{\mathbf{z}}\right)\right]^{-1/2} \operatorname{cov}_1\left(\hat{\mathbf{T}}_{\mathbf{z}}, \hat{T}_y\right)\left[v\left(\hat{T}_y\right)\right]^{-1/2}$ and $\breve{\rho}_1\left(\tilde{\mathbf{T}}_{\mathbf{z}}, \hat{\mathbf{T}}_{\mathbf{x}}\right) = \left[v_1\left(\tilde{\mathbf{T}}_{\mathbf{z}}\right)\right]^{-1/2} \operatorname{cov}_1\left(\tilde{\mathbf{T}}_{\mathbf{z}}, \hat{\mathbf{T}}_{\mathbf{x}}\right)\left[v\left(\hat{\mathbf{T}}_{\mathbf{x}}\right)\right]^{-1/2}$. Here we use the tilda (~) to define a total based only on the first phase sample, i.e., $\tilde{\mathbf{T}}_{\mathbf{z}} = \sum_{k \in s_1} d_{1k}\mathbf{z}_k$. The modified correlations are smaller than the usual correlation because they use the overall variances in their denominators. With the modified correlations $\beta_{\mathbf{z}}$ becomes

$$\left[v_1\left(\hat{\mathbf{T}}_z\right)\right]^{-1} \times$$
$$\left[\operatorname{cov}_1\left(\hat{\mathbf{T}}_z, \hat{T}_y\right)\left(1-\left[\bar{\rho}_1\left(\hat{\mathbf{T}}_z, \hat{T}_y\right)\right]^{-1} \bar{\rho}_1\left(\hat{\mathbf{T}}_z, \hat{\mathbf{T}}_{\mathbf{x}}\right)\bar{\rho}_1\left(\hat{\mathbf{T}}_{\mathbf{x}}, \hat{T}_y\right)\right)\right].$$

First note that the modified correlations are less than or equal to 1 and strictly less than 1 if the second stage variance of **x** is greater than zero. This means that the correlations are less than 1 and therefore the effective covariance of **z** and y cannot be equal zero.

The interpretation is reasonable: the full knowledge of **x** does not translate into full knowledge of **z**, even if they are completely correlated. So the "effective" first phase variance and covariance cannot be completely eliminated because **x** and **z** are completely correlated, i.e., **x** = α **z** for some vector of constants α and no error.

3. Estimators of the Regression Parameters

3.1 Estimation

We suggest estimating the parameters using the unbiased estimators of the sample covariances. Based on Särndal et.al. (1992; p. 348) an estimator of the first and second phase sample covariances of the totals of y and **x** are, respectively,

$$\widehat{\operatorname{cov}}_1\left(\hat{T}_y, \hat{\mathbf{T}}_{\mathbf{x}}\right) = \sum\sum_{k\ell \in s_1} \pi_{1k\ell}^{-1}\left(\pi_{1k\ell} - \pi_{1k}\pi_{1\ell}\right)\left(y_{1k}/\pi_{1k}\right)\left(\mathbf{x}_{1\ell}/\pi_{1\ell}\right) \quad (1)$$

and

$$\widehat{\operatorname{cov}}_2\left(\hat{T}_y, \hat{\mathbf{T}}_{\mathbf{x}}\right) = \sum\sum_{k\ell \in s_2} \pi_{2k\ell}^{-1}\left(\pi_{2k\ell} - \pi_{2k}\pi_{2\ell}\right)\left(y_{2k}/\pi_{1k}\pi_{2k}\right)\left(\mathbf{x}_{2\ell}/\pi_{1\ell}\pi_{2\ell}\right) \quad (2)$$

Note that the overall covariances are equal to the sum of the first and second phase covariances, i.e., $\widehat{\operatorname{cov}}\left(\hat{T}_y, \hat{\mathbf{T}}_{\mathbf{x}}\right) = \widehat{\operatorname{cov}}_1\left(\hat{T}_y, \hat{\mathbf{T}}_{\mathbf{x}}\right) + \widehat{\operatorname{cov}}_2\left(\hat{T}_y, \hat{\mathbf{T}}_{\mathbf{x}}\right)$.

With the residuals in Table 3, we suggest estimating the variance of $\hat{T}_{y,opt}$ using the weighted residual technique (Särndal 1987) applied to the two-phase s.d. Applying equations (1) and (2), the variance estimator for the two-phase s.d. is

$$\hat{v}\left(\hat{T}_{y,reg}\right) = \sum\sum_{k\ell \in s} \pi_{1k\ell}^{-1}\left(\pi_{1k\ell} - \pi_{1k}\pi_{1\ell}\right)\left(e_{1k}/\pi_{1k}\right)\left(e_{1\ell}/\pi_{1\ell}\right) + \sum\sum_{k\ell \in s} \pi_{2k\ell}^{-1}\left(\pi_{2k\ell} - \pi_{2k}\pi_{2\ell}\right)\left(e_{2k}/\pi_{1k}\pi_{2k}\right)\left(e_{2\ell}/\pi_{1\ell}\pi_{2\ell}\right).$$

3.2 Estimators for the Cases Involving 111

Tables 3 and 4 do not list any of the cases that involve 111. We could provide their unique solution, but instead note that they would be equivalent to applying the solutions we have already provided. All we need to do is understand that a AV that is 111 can be treated as two different AVs that are either 101, 110 or 011. With this knowledge we can use the estimators we already have.

4. Empirical Example

We now consider an example for the two-phase s.d. that will allow us to examine the following questions:

Q1. Are the optimal estimators "optimal," with respect to the calibration estimators of E&S? We constructed them by minimizing the variance, so they should have smaller variances than the calibration estimators of E&S?
Q2. What kind of auxiliary information is best to have and use?
Q3. Are there any cases where including an AV twice in the model improves the estimator?
Q4. Are the estimators of the variances reasonable?

We compare the estimators using the populations of E&S. We constructed the three populations as in E&S, each with N = 1,000 units. The three populations were generated so that the two AVs x_{1k} and x_{2k} were independent of each other. We generated the AVs x_{1k} and x_{2k} from a gamma (9,10), where a gamma(a, b) distribution has density $f(x) = [\Gamma(a)\, b^a]^{-1} x^{a-1} \exp(-x/b)$ for $x > 0$, with $E(x) = ab$ and $\mathrm{var}(x) = ab^2$. The error term e_k was generated from a normal distribution with mean zero and variance 25. The variable of interest y_k for the first set of population was a different function of x_{1k}, x_{2k} and e_k, i.e., Pop 12(0): $y_k = x_{1k} + x_{2k} + e_k$, Pop 1(0): $y_k = \sqrt{2}x_{1k} + e_k$, and Pop 2(0): $y_k = \sqrt{2}x_{2k} + e_k$.

We selected 100,000 independent samples of 100 units from each population. The first phase sample had n_1 = 500 units and the second phase sample had n_2 = 200 units. The s.d. of both phases was simple random sampling without replacement (srswor). Estimates of the total and variance of the total were calculated. For each combination of AVs we applied two estimators: the optimal estimators of Table 3 and calibration estimators of E&S.

Table 5 describes the empirical results. The first two columns of Table 5 list the estimators we applied and the third column identifies whether the estimator was a calibration (cal) or optimal (opt) estimator. The column "SimVar" denotes the simulated variance of the estimator or more simply the variance calculated using the estimated totals. We did not include the simulated MSE because the estimators, both the optimal and the calibration, were all approximately unbiased.

The next three columns of Table 5, "Est Var," "Est VE," and "Est EV"; respectively, denote the overall, first and second phase estimates of the variance. The values of Est VE and Est EV are the mean of the estimated first and second phase variances [using equations (1) and (2) and the residuals defined in Table 3] over all the simulations. Est Var is the mean of the sum of "Est VE + Est EV", over all simulations.

We now discuss the results of the example with respect to our three questions.

A1. In most cases the variances of the optimal estimators were smaller than the calibration estimators. The two exceptions were cases (1d) and (2d). To understand this result we note that for cases (1d) and (2d) the calibration solution for the regression parameters is the same as used by (1b). This solution treats both AVs as "overall" AVs, i.e., $\beta_{x1} = \left(\sum_{k\in s} d_{1k} d_{2k} \mathbf{x}'_{1k}\mathbf{x}_{1k}\right)^{-1} \sum_{k\in s} d_{1k} d_{2k} \mathbf{x}_{1k} y_k$ and $\beta_{x2} = \left(\sum_{k\in s} d_{1k} d_{2k} \mathbf{x}'_{2k}\mathbf{x}_{2k}\right)^{-1} \sum_{k\in s} d_{1k} d_{2k} \mathbf{x}_{2k} y_k$. This is advantageous since the s.d. of the example is simple – both phases are srswor. In comparison, the optimal estimator limits itself to using only the covariances related to the type of AVs available. For (1d) the optimal estimator only uses the second phase covariances and (2d) uses both the overall and second phase covariances.

Although the calibration estimator for (1d) and (2d) might do well in this example, we expect that the optimal estimator would do a better than the calibration estimators with a more complicated s.d. since it uses all the information about the sample design via the covariances.

A2. In general the example confirmed what is readily accepted about using AVs in regression estimators.

- Using models that employ AVs associated with the variable of interest reduce the variance of the estimator. The more information an AV contains, and thereby adds to the model, the more it can reduce the sample variance. Using an AV that is 111 was better than 101, which was better than both 110 or 011.
- Additional AVs that are also associated with the variable of interest can reduce the variance further, if they are not collinear with the other AVs in the model.
- The model assisted estimators are resistant to model misspecification; however, using models with AVs

that have no association with the variable of interest can increase the variance – making it larger than doing nothing at all, i.e., estimator (0).

A3. Using an AV that is 111 and therefore includes the AV twice in the estimator, reduces the variance as compared with an AV that is 101. When we consider population 12(0) where y_k is defined in terms of both AVs, we see that estimators (2a) and (2c) do better than estimators (2b) and (2d), respectively. The exception is the case of one AV, where the estimation is not improved. Both cases (1a) and (1b) have the same estimator of the regression parameter, so knowledge of all three levels of an AV does not improve the estimator.

A4. The estimators of the variance, both the overall and the separate estimators for each of the two-phases, appear to be reasonable, since the estimates of Est Var are close to the estimates of SimVar.

This report is released to inform interested parties of (ongoing) research and to encourage discussion (of work in progress).

References

Ahmad, M.S., Khan, S.U., and Tripathi, T.P. (1995), "Model based regression estimator of a finite population mean using multi-phase sampling," The Aligarh Journal of Statistics, 15, 69-74.

Andersson, P.G. and Thorburn, D. (2005). An Optimal Calibration Distance Leading to the Optimal Regression Estimator, Survey Methodology, 31, 1, 95-99.

Ash, S. (2003). "Simultaneous calibration estimators for two-phase samples," Proceedings of the American Statistical Association, Section on Survey Research Methods, 395-400.

Axelson, M. (2000). "Variance Estimation for the Two-Phase Regression Estimator - A Calibration Approach," Proceedings of the American Statistical Association, Section on Survey Research Methods, 799-804.

Cassel, M.C., Särndal, C.-E., and Wretman, J.H. (1976). "Some results on generalized difference estimation and generalized regression estimation for finite populations," Biometrika, 63, 615-620.

Dupont, F. (1995)."Alternative Adjustments Where There Are Several Levels of Auxiliary Information," Survey Methodology, 24, 11-20.

Estevao, V.M., and Särndal, C.-E. (2003). "The Ten Cases of Auxiliary Informtion for Calibration in Two-Phase Sampling," Journal of Official Statistics, 18, 233-255.

Hidroglou, M.A. and Särndal, C.-E. (1995). "Use of Auxiliary Information for Two-Phase Sampling," Proceedings of the Section on Survey Research Methods, Joint Statistical Meetings, 873-878.

Hidroglou, M.A. and Särndal, C.-E. (1998). "Use of Auxiliary Information for Two-Phase Sampling," Survey Methodology, 24, 11-20.

Kiregyera, B. (1980) "A Chain Ratio-Type Estimator in Finite Population Double Sampling Using Two Auxiliary Variables," Metrika, 27, 217-223.

Kiregyera, B. (1984) "Regression-Type Estimators Using Two Auxiliary Variables and the Model of Double Sampling from Finite Populations," Metrika, 31, 215-226.

Montanari, G.E. (1987). "Post-Sampling Efficient QR-Prediction in Large-Sample Surveys," International Statistical Review, 55, 191-202.

Roy, D.C. (2003). "A Regression-Type Estimator in Two-Phase Sampling Using Two Auxiliary Variables," Pakistan Journal of Statistics, 19, 3, 281-290.

Sahoo, L.N. and Panda, P. (1999). "A Class of Estimators Using Auxiliary Information in Two-Stage Sampling," Australian and New Zealand J. Statistics, 41, 4, 405-410.

Särndal, C.E., Swensson, B., and Wretman, J. (1989). "The weighted residual technique for estimating the variance of the general regression estimator of the finite population total," Biometrika, 76, 527-537.

Särndal, C.E., Swensson, B., and Wretman, J. (1992). Model Assisted Survey Sampling, New York, Springer-Verlag.

Singh, H.P., Upadhyaya, L.N. and Chandra, P.(2004). "A General Family of Estimators for Estimating Population Mean Using Two Auxiliary Variables in Two-Phase Sampling," Statistics in Transition, 6, 7, 1055-1077.

Tripathi, T. P. , and Ahmed, M. S. (1995), "A class of estimators for a finite population mean based on multivariate information and general two phase sampling," Calcutta Statistical Association Bulletin, 45 , 203-218.

Wu, C. and Luan, Y. (2003). "Optimal Calibration Estimators Under Two-Phase Sampling," Journal of Official Statistics, 19, 119-131.

Table 3: Optimal Regression Parameters for Two-Phase Sample Designs

	Model	Regression Estimator	Residuals
(0) None		$\sum_{i\in s}\pi_{1k}^{-1}\pi_{2k}^{-1}y_k$	$e_{1k}=y_k$, $e_{2k}=y_k$
(1b) Overall (**x**)	$\mu_k=\mathbf{x}_k\beta_\mathbf{x}$	$\mathbf{T}'_\mathbf{x}\beta_\mathbf{x}+\sum_{i\in s}\pi_{1k}^{-1}\pi_{2k}^{-1}\left(y_k-\mathbf{x}'_k\beta_\mathbf{x}\right)$	$e_{1k}=y_k-\mathbf{x}'_k\beta_\mathbf{x}$ $e_{2k}=y_k-\mathbf{x}'_k\beta_\mathbf{x}$
(1c) First (**z**)	$\mu_k=\mathbf{z}_k\beta_\mathbf{z}$	$\mathbf{T}'_\mathbf{z}\beta_\mathbf{z}+\sum_{k\in s}\pi_{1k}^{-1}\pi_{2k}^{-1}y_k-\sum_{k\in s}\pi_{1k}^{-1}\mathbf{z}'_k\beta_\mathbf{z}$	$e_{1k}=y_k-\mathbf{z}'_k\beta_\mathbf{z}$ $e_k=y_k$
(1d) Second (**v**)	$\mu_k=\mathbf{v}_k\beta_\mathbf{v}$	$\sum_{k\in s_1}\pi_{1k}^{-1}\mathbf{v}'_k\beta_\mathbf{v}+\sum_{k\in s}\pi_{1k}^{-1}\pi_{2k}^{-1}\left(y_k-\mathbf{v}'_k\beta_\mathbf{v}\right)$	$e_{1k}=y_k$ $e_{2k}=y_k-\mathbf{v}'_k\beta_\mathbf{v}$
(2b) Overall (**x**) & First (**z**)	$\mu_k=\mathbf{x}_k\beta_\mathbf{x}+\mathbf{z}_k\beta_\mathbf{z}$	$\mathbf{T}'_\mathbf{x}\beta_\mathbf{x}+\mathbf{T}'_\mathbf{z}\beta_\mathbf{z}+\sum_{k\in s}\pi_{1k}^{-1}\pi_{2k}^{-1}\left(y_k-\mathbf{x}'_k\beta_\mathbf{x}\right)-\sum_{k\in s_1}\pi_{1k}^{-1}\mathbf{z}'_k\beta_\mathbf{z}$	$e_{1k}=y_k-\mathbf{x}'_k\beta_\mathbf{x}-\mathbf{z}'_k\beta_\mathbf{z}$ $\hat{e}_{2k}=y_k-\mathbf{x}'_k\hat{\beta}_\mathbf{x}$
(2d) Overall (**x**) & Second (**v**)	$\mu_k=\mathbf{x}_k\beta_\mathbf{x}+\mathbf{v}_k\beta_\mathbf{v}$	$\mathbf{T}'_\mathbf{x}\beta_\mathbf{x}+\sum_{k\in s_1}\pi_{1k}^{-1}\mathbf{v}'_k\beta_\mathbf{v}+\sum_{k\in s}\pi_{1k}^{-1}\pi_{2k}^{-1}\left(y_k-\mathbf{x}'_k\beta_\mathbf{x}-\mathbf{v}'_k\beta_\mathbf{v}\right)$	$e_{1k}=y_k-\mathbf{x}'_k\beta_\mathbf{x}$ $e_{2k}=y_k-\mathbf{x}'_k\beta_\mathbf{x}-\mathbf{v}'_k\beta_\mathbf{v}$
(2e) First (**z**) & Second (**v**)	$\mu_k=+\mathbf{z}_k\beta_\mathbf{z}+\mathbf{v}_k\beta_\mathbf{v}$	$\mathbf{T}'_\mathbf{z}\beta_\mathbf{z}+\sum_{k\in s_1}\pi_{1k}^{-1}\mathbf{v}'_k\beta_\mathbf{v}+\sum_{k\in s}\pi_{1k}^{-1}\pi_{2k}^{-1}\left(y_k-\mathbf{v}'_k\beta_\mathbf{v}\right)-\sum_{k\in s_1}\pi_{1k}^{-1}\mathbf{z}'_k\beta_\mathbf{z}$	$e_{1k}=y_k-\mathbf{z}'_k\beta_\mathbf{z}$ $e_{2k}=y_k-\mathbf{x}'_k\beta_\mathbf{x}$
(3) Overall (**x**), First (**z**) & Second (**v**)	$\mu_k=\mathbf{x}_k\beta_\mathbf{x}+\mathbf{z}_k\beta_\mathbf{z}+\mathbf{v}_k\beta_\mathbf{v}$	$\mathbf{T}_\mathbf{x}\beta_\mathbf{x}+\mathbf{T}_\mathbf{z}\beta_\mathbf{z}$ $+\sum_{k\in s_1}\pi_{1k}^{-1}\mathbf{v}'_k\beta_\mathbf{v}+\sum_{k\in s}\pi_{1k}^{-1}\pi_{2k}^{-1}\left(y_k-\mathbf{x}'_k\beta_\mathbf{x}-\mathbf{v}'_k\beta_\mathbf{v}\right)-\sum_{k\in s_1}\pi_{1k}^{-1}\mathbf{z}'_k\beta_\mathbf{z}$	$e_{1k}=y_k-\mathbf{z}'_k\beta_\mathbf{z}-\hat{\mathbf{x}}'_k\beta_\mathbf{x}$ $e_{2k}=y_k-\mathbf{x}'_k\beta_\mathbf{x}-\mathbf{v}'_k\beta_\mathbf{v}$

Table 4: Optimal Regression Parameters for Two-Phase Sample Designs

Estimators	Model parameters
(1b) Overall (**x**)	$\beta_\mathbf{x}=\left[v\left(\hat{\mathbf{T}}_\mathbf{x}\right)\right]^{-1}\mathrm{cov}\left(\hat{\mathbf{T}}_\mathbf{x},\hat{T}_y\right)$
(1c) First (**z**)	$\beta_\mathbf{z}=\left[v_1\left(\hat{\mathbf{T}}_\mathbf{z}\right)\right]^{-1}\mathrm{cov}_1\left(\hat{\mathbf{T}}_\mathbf{z},\hat{T}_y\right)$
(1d) Second (**v**)	$\beta_\mathbf{v}=\left[v_2\left(\hat{\mathbf{T}}_\mathbf{v}\right)\right]^{-1}\mathrm{cov}_2\left(\hat{\mathbf{T}}_\mathbf{v},\hat{T}_y\right)$
(2b) Overall (**x**) & First (**z**)	$\beta_\mathbf{x}=\left[v\left(\hat{\mathbf{T}}_\mathbf{x}\right)-\mathrm{cov}_1\left(\hat{\mathbf{T}}_\mathbf{x},\hat{\mathbf{T}}_\mathbf{z}\right)\left[v_1\left(\hat{\mathbf{T}}_\mathbf{z}\right)\right]^{-1}\mathrm{cov}_1\left(\hat{\mathbf{T}}_\mathbf{z},\hat{\mathbf{T}}_\mathbf{x}\right)\right]^{-1}\left[\mathrm{cov}\left(\hat{\mathbf{T}}_\mathbf{x},\hat{T}_y\right)-\mathrm{cov}_1\left(\hat{\mathbf{T}}_\mathbf{x},\hat{\mathbf{T}}_\mathbf{z}\right)\left[v_1\left(\hat{\mathbf{T}}_\mathbf{z}\right)\right]^{-1}\mathrm{cov}_1\left(\hat{\mathbf{T}}_\mathbf{z},\hat{T}_y\right)\right]$ $\beta_\mathbf{z}=\left[v_1\left(\hat{\mathbf{T}}_\mathbf{z}\right)-\mathrm{cov}_1\left(\hat{\mathbf{T}}_\mathbf{z},\hat{\mathbf{T}}_\mathbf{x}\right)\left[v\left(\hat{\mathbf{T}}_\mathbf{x}\right)\right]^{-1}\mathrm{cov}_1\left(\hat{\mathbf{T}}_\mathbf{x},\hat{\mathbf{T}}_\mathbf{z}\right)\right]^{-1}\left[\mathrm{cov}_1\left(\hat{\mathbf{T}}_\mathbf{z},\hat{T}_y\right)-\mathrm{cov}_1\left(\hat{\mathbf{T}}_\mathbf{z},\hat{\mathbf{T}}_\mathbf{x}\right)\left[v\left(\hat{\mathbf{T}}_\mathbf{x}\right)\right]^{-1}\mathrm{cov}\left(\hat{\mathbf{T}}_\mathbf{x},\hat{T}_y\right)\right]$
(2d) Overall (**x**) & Second (**v**)	$\beta_\mathbf{x}=\left[v\left(\hat{\mathbf{T}}_\mathbf{x}\right)-\mathrm{cov}_2\left(\hat{\mathbf{T}}_\mathbf{x},\hat{\mathbf{T}}_\mathbf{v}\right)\left[v_2\left(\hat{\mathbf{T}}_\mathbf{v}\right)\right]^{-1}\mathrm{cov}_2\left(\hat{\mathbf{T}}_\mathbf{v},\hat{\mathbf{T}}_\mathbf{x}\right)\right]^{-1}\left[\mathrm{cov}\left(\hat{\mathbf{T}}_\mathbf{x},\hat{T}_y\right)-\mathrm{cov}_2\left(\hat{\mathbf{T}}_\mathbf{x},\hat{\mathbf{T}}_\mathbf{z}\right)\left[v_2\left(\hat{\mathbf{T}}_\mathbf{z}\right)\right]^{-1}\mathrm{cov}_2\left(\hat{\mathbf{T}}_\mathbf{z},\hat{T}_y\right)\right]$ $\beta_\mathbf{v}=\left[v_2\left(\hat{\mathbf{T}}_\mathbf{v}\right)-\mathrm{cov}_2\left(\hat{\mathbf{T}}_\mathbf{v},\hat{\mathbf{T}}_x\right)\left[v\left(\hat{\mathbf{T}}_x\right)\right]^{-1}\mathrm{cov}_2\left(\hat{\mathbf{T}}_\mathbf{x},\hat{\mathbf{T}}_\mathbf{v}\right)\right]^{-1}\left[\mathrm{cov}_2\left(\hat{\mathbf{T}}_\mathbf{v},\hat{T}_y\right)-\mathrm{cov}_2\left(\hat{\mathbf{T}}_\mathbf{v},\hat{\mathbf{T}}_\mathbf{x}\right)\left[v\left(\hat{\mathbf{T}}_\mathbf{x}\right)\right]^{-1}\mathrm{cov}\left(\hat{\mathbf{T}}_\mathbf{x},\hat{T}_y\right)\right]$
(2e) First (**z**) & Second (**v**)	$\beta_\mathbf{z}=\left[v_1\left(\hat{\mathbf{T}}_\mathbf{z}\right)\right]^{-1}\mathrm{cov}_1\left(\hat{\mathbf{T}}_\mathbf{z},\hat{T}_y\right)$, $\beta_\mathbf{v}=\left[v_2\left(\hat{\mathbf{T}}_\mathbf{v}\right)\right]^{-1}\mathrm{cov}_2\left(\hat{\mathbf{T}}_\mathbf{v},\hat{T}_y\right)$
(3) Overall (**x**), First (**z**) & Second (**v**)	$\beta_\mathbf{x}=\left[\left(v_1\left(\hat{\mathbf{T}}_\mathbf{x}\right)-\mathrm{cov}_1\left(\hat{\mathbf{T}}_\mathbf{x},\hat{\mathbf{T}}_\mathbf{z}\right)\left[v_1\left(\hat{\mathbf{T}}_\mathbf{z}\right)\right]^{-1}\mathrm{cov}_1\left(\hat{\mathbf{T}}_\mathbf{z},\hat{\mathbf{T}}_\mathbf{x}\right)\right)+\left(v_2\left(\hat{\mathbf{T}}_\mathbf{x}\right)-\mathrm{cov}_2\left(\hat{\mathbf{T}}_\mathbf{x},\hat{\mathbf{T}}_\mathbf{v}\right)\left[v_2\left(\hat{\mathbf{T}}_\mathbf{v}\right)\right]^{-1}\mathrm{cov}_2\left(\hat{\mathbf{T}}_\mathbf{v},\hat{\mathbf{T}}_\mathbf{x}\right)\right)\right]^{-1}$ $\times\left[\left(\mathrm{cov}_1\left(\hat{\mathbf{T}}_\mathbf{x},T_y\right)-\mathrm{cov}_1\left(\hat{\mathbf{T}}_\mathbf{x},\hat{\mathbf{T}}_\mathbf{z}\right)\left[v_1\left(\hat{\mathbf{T}}_\mathbf{z}\right)\right]^{-1}\mathrm{cov}_1\left(\hat{\mathbf{T}}_\mathbf{z},\hat{T}_y\right)\right)+\left(\mathrm{cov}_2\left(\hat{\mathbf{T}}_\mathbf{x},T_y\right)-\mathrm{cov}_2\left(\hat{\mathbf{T}}_\mathbf{x},\hat{\mathbf{T}}_\mathbf{v}\right)\left[v_2\left(\hat{\mathbf{T}}_\mathbf{v}\right)\right]^{-1}\mathrm{cov}_2\left(\hat{\mathbf{T}}_\mathbf{v},\hat{T}_y\right)\right)\right]$ $\beta_\mathbf{z}=\left[v_1\left(\hat{\mathbf{T}}_\mathbf{z}\right)\right]^{-1}\left[\mathrm{cov}_1\left(\hat{\mathbf{T}}_\mathbf{z},\hat{T}_y\right)-\mathrm{cov}_1\left(\hat{\mathbf{T}}_\mathbf{z},\hat{\mathbf{T}}_\mathbf{x}\right)\beta_\mathbf{x}\right]$, $\beta_\mathbf{v}=\left[v_2\left(\hat{\mathbf{T}}_\mathbf{v}\right)\right]^{-1}\left[\mathrm{cov}_2\left(\hat{\mathbf{T}}_\mathbf{v},\hat{T}_y\right)-\mathrm{cov}_2\left(\hat{\mathbf{T}}_\mathbf{v},\hat{\mathbf{T}}_\mathbf{x}\right)\beta_\mathbf{x}\right]$

Table 5: Results from the Empirical Example

Estimator	E&S case	Estimator	Population 12(0): $y_k = x_{1k} + x_{2k} + e_k$ SimVar	Est Var	Est VE	Est EV	Population 1(0): $y_k = \sqrt{2}x_{1k} + e_k$ SimVar	Est Var	Est VE	Est EV	Population 2(0): $y_k = \sqrt{2}x_{2k} + e_k$ SimVar	Est Var	Est VE	Est EV
(0) No auxiliary info	C4		10.01	9.41	2.35	7.06	9.09	8.79	2.20	6.59	9.63	9.25	2.31	6.94
(1a) All (x_{1k})	A2	opt	8.55	8.47	2.12	6.35	2.43	2.43	0.61	1.82	15.19	15.03	3.76	11.27
		cal	9.00	8.92	2.23	6.69	2.43	2.43	0.61	1.82	16.45	16.30	4.07	12.22
(1b) Overall (x_{1k})	B2	opt	8.55	8.47	2.12	6.35	2.43	2.43	0.61	1.82	15.19	15.03	3.76	11.27
		cal	9.00	8.92	2.23	6.69	2.43	2.43	0.61	1.82	16.45	16.30	4.07	12.22
(1c) First (x_{1k})	A4	opt	9.61	9.18	2.12	7.06	7.43	7.20	0.61	6.59	11.04	10.70	3.76	6.94
		cal	9.72	9.29	2.23	7.06	7.73	7.20	0.61	6.59	11.36	11.02	4.08	6.94
(1d) Second (x_{1k}, x_{2k})	C1	opt	4.42	4.22	2.35	1.86	4.19	4.10	2.20	1.90	4.30	4.16	2.31	1.85
		cal	4.38	4.17	2.35	1.82	4.11	4.02	2.20	1.82	4.30	4.19	2.31	1.85
(1d) Second (x_{1k})	C2	opt	8.96	8.71	2.35	6.35	4.09	4.02	2.20	1.82	13.76	14.58	2.31	11.27
		cal	9.29	9.04	2.35	6.69	4.09	4.02	2.20	1.82	14.69	14.54	2.31	12.22
(1d) Second (x_{2k})	C3	opt	8.91	8.68	2.35	6.33	12.60	12.60	2.20	10.29	4.29	4.17	2.31	1.86
		cal	9.32	9.09	2.35	6.73	13.33	13.33	2.20	11.04	4.29	4.17	2.31	1.86
(2a) All (x_{1k}) & First (x_{2k})	n/a	opt	7.03	6.97	0.62	6.35	2.46	2.46	0.63	1.82	11.96	11.89	0.62	11.27
(2b) Overall (x_{1k}) & First (x_{2k})	n/a	opt	7.52	7.44	1.41	6.03	2.44	2.43	0.61	1.82	12.96	12.82	2.33	10.48
(2c) All (x_{1k}) & Second (x_{2k})	A1	opt	4.03	3.98	2.12	1.86	2.53	2.53	0.61	1.90	5.70	5.60	3.76	1.85
		cal	4.09	4.05	2.23	1.82	2.44	2.43	0.61	1.82	6.01	5.93	4.08	1.85
(2d) Overall (x_{1k}) & Second (x_{2k})	B1	opt	4.37	4.31	1.70	2.61	2.47	2.47	0.61	1.85	6.28	6.14	2.71	3.43
		cal	3.43	3.36	1.54	1.82	2.45	2.45	0.61	1.82	4.30	4.16	2.31	1.85
(2e) First (x_{1k}) & Second (x_{2k})	A3	opt	8.53	8.45	2.12	6.33	10.93	10.94	0.60	10.29	5.69	5.61	3.76	1.86
		cal	8.78	8.71	1.97	6.73	11.70	11.74	0.65	11.04	6.01	5.94	4.08	1.86