Multi-Level Rotation Sampling Designs and the Variances of Extended Generalized Composite Estimators

YouSung Park, JaiWon Choi and KeeWhan Kim*

Abstract

We classify rotation sampling designs into two classes. The first class replaces sample units within the same rotation group while the second class replaces sample units between different rotation groups. The first class is specified by the three-way balanced design which is a multi-level version of previous balanced designs. We introduce an extended generalized composite estimator (EGCE) and derive its variance and mean squared error for each of the two classes of design, cooperating two types of correlations and three types of biases. Unbiased estimators are derived for difference between interview time biases, between recall time biases, and between rotation group biases. Using the variance and mean squared error, since any rotation design belongs to one of the two classes and the EGCE is a most general estimator for rotation design, we evaluate the efficiency of EGCE to simple weighted estimator and the effects of levels, design gaps, and rotation patterns on variance and mean squared error.

Keywords : The l-level rotation design; Three-way balancing; Extended Generalized Composite Estimator; Two types of correlations; Variance and Mean Squared Error; Compromise coefficients.

*YouSung Park is professor, Department of Statistics, Korea University, 5-1 Anam-Dong, Sungbuk-gu, KOREA(email:yspark@mail.korea.ac.kr); Jai Won Choi is mathematical statistician, NCHS, CDC, 6525 Belcrest RD, Hyattsville, MD, U.S.A.(email : jwc7@cdc.gov); and KeeWhan KIM is a researcher, Department of Statistics, Korea University, 5-1 Anam-Dong, Sungbuk-gu, KOREA(email:korpen@ahanet.co.kr) .

1 Introduction

The entire sample units or population are partitioned into a finite number of rotation groups in rotation design. One class of rotation designs replaces sample units within the same rotation group to retain all rotation groups in monthly sample. The other class consists monthly sample only with one rotation group and the rotation group is replaced with another rotation group as one month advances. Thus, replacement of sample units is occurred between different rotation groups in this second class.

Typical examples of the first class are U.S. Current Population Survey (CPS), the Canadian Labor Force Survey (CLFS) for one-level rotation design (i.e., $l = 1$). U.S. Consumer Expenditure Survey (CEX) for 3-level rotation design, and National Crime Victimization Survey (NCVS) for 6-level rotation design. Here, the respondent in a ℓ-level rotation design reports the information of current survey month and $\ell - 1$ previous months. We characterize this class as three-way balanced rotation design in which monthly sample is balanced in rotation group, interview time as well as recall time. Previous balanced designs are special cases of the three-way balanced design (Cantwell 1990, Park, Kim and Choi 2001).

We characterize the second class of ℓ-level rotation designs as ℓ/G designs where G is the number of rotation groups since each rotation group returns to sample every G months in this class of designs. Examples of such designs are the U.S. National Ambulatory Medical Care Survey (NAMCS) for one-level rotation design, the U.S. Monthly Retail Trade Survey (MRTS) and the U.S. Survey of Income and Program Participation (SIPP) for 2-level and 4-level rotation designs, respectively.

Because longer recalling is expected to provide less accurate information for the recalled month, recall time bias may exist in ℓ-level rotation design (Cantwell and Caldwell 1998). The entire sample units or population are partitioned into a finite number of rotation groups in the rotation design. Because different rotation groups may have different expectations of a characteristic of interest, rotation group bias can exist in rotation design. The same sample unit is interviewed more than one time in the rotation design. Because of response burden from repeated interviews, interview time bias may exist in rotation design (Bailar 1975). For each class of rotation designs, we present unbiased estimators for differences of recall time biases, rotation group bias, and interview time biases.

We modify the generalized composite estimator (GCE) for one-level rotation design (Breau and Ernst 1983) to a multi-level version and call it extended GCE (EGCE). This EGCE provides an interim estimator at any survey month before we have full information about month t and a final estimator for month t. We derive the variance and MSE of EGCE for two classes of designs, cooperating two types of correlations and three types of biases. Minimizing a weighted sum of MSEs of l-level EGCEs of general interest, we derive one set of the compromise coefficients to preserve the additivity of estimates.

By numerical examples, since any rotation design belongs to one of the two classes and the EGCE is a most general estimator for rotation design, we evaluate the efficiency of EGCE to simple weighted estimator and compare the two class of rotation designs. We also study the effects of levels, design gaps, and rotation patterns on variance and mean squared error for different values of two types of correlations.

The rest of this paper is divided into 4 sections. In section 2, we discuss the two general classes of l-level rotation designs and present a simple algorithm to construct them. In section 3, the EGCE and its variance and bias for various characteristics are presented for the two classes of l-level rotation designs. In section 4, we show the common coefficients for GCEs by minimizing their MSEs. Using the results of sections 3 and 4, we performs efficiency studies for two classes of designs in section 5. Finally, we conclude this paper in section 6.

2 Three-Way Balanced Design and l/G Rotation Design

We describe a general rotation system for a l-level rotation design. When a sample unit is selected from each rotation group, this unit returns to the sample for every lth month until its r_{11}th interview and is out of the sample for the next $r_{21}+l-1$ successive months. Then, the sample unit is again interviewed for every lth month until its $(r_{11}+r_{12})$th interview and is out of the sample for the next $r_{22}+l-1$ months. This procedure is repeated until this sample unit returns to the sample for its final $(\sum_{i=1}^{m} r_{1i})$th interview. We denote this rotation system as $\prod_{i=1}^{m}(r_{1i}(l)-r_{2i})$. When a l-level rotation design follows this rotation system with $m<\infty$ and the number of rotation groups is $\sum_{i=1}^{m} r_{1i}$, we call this design l-level $r_{11}-\cdots-r_{2,m-1}-r_{1m}$

rotation design. When $m = 1$, we call it the l-level r_{11} in-then-out design. In particular, when $l = 1$, the l-level $r_{11} - \cdots - r_{2,m-1} - r_{1m}$ rotation design is reduced to the one-level rotation design of Cantwell (1990) and the rotation design considered by Park, Kim and Choi (2001) for their two-way balanced design.

2.1 Three-Way Balanced Design

The l-level $r_{11} - \cdots - r_{2,m-1} - r_{1m}$ rotation design has $\sum_{i=1}^{m} r_{1i}$ rotation groups and interview times from 1 to $\sum_{i=1}^{m} r_{1i}$ and l recall times from 0 to $l-1$. Thus, we can define the three-way balanced design as follows.

Definition 2.1. The l-level $r_{11} - \cdots - r_{2,m-1} - r_{1m}$ design is balanced in three-ways if

(a) all $\sum_{i=1}^{m} r_{1i}$ rotation groups and interview times from 1 to $\sum_{i=1}^{m} r_{1i}$ are present in every monthly sample and

(b) (a) is true for any recall time from 0 to $l-1$.

The condition (a) for $l = 1$ is equivalent to the conditions that Park, Kim and Choi (2001) imposed for their two-way balanced design. Since all rotation groups should be included in every monthly sample in the three-way balanced design, all rotation groups have equal opportunity to be represented in monthly sample and the replaced sample unit comes from the same rotation group. Characteristics of the outgoing and incoming units are as similar as possible so that units within the same rotation group are homogeneous. The U.S. CES is a typical example of the three-way balanced design.

The following algorithm provides how to construct the three-way balanced design from the l-level $r_{11} - \cdots - r_{2,m-1} - r_{1m}$ rotation design as illustrated by Figure 1 of the 3-level 5-8-3 design.

Algorithm 1

Step 1: To create the column labels for the unit and group on the top two rows, arrange sample units by their affiliation indices (α, g) in the order of $(1,1)$, $(1,2), \cdots, (1, \sum_{i=1}^{m} r_{1i}), \cdots,$ $(\alpha^*, 1)$, $(\alpha^*, 2), \cdots, (\alpha^*, \sum_{i=1}^{m} r_{1i})$ where α^* is chosen to be large enough to last entire survey.

Step 2: Next fill in the third row of the first month t. According to the reverse order of the $\prod_{i=1}^{m}(r_{1i}(l) - r_{2i})$ rotation system, select every l-th index from Step 1 until the first r_{1m} indices are selected and leave the next $r_{2,m-1} + l - 1$ indices. Repeat this procedure until the last r_{11} indices are selected. The $\sum_{i=1}^{m} r_{1i}$ sample units with the selected indices are the sample of the initial month t. For example, the 3-level 5-8-3 design in Figure 1 has $r_{11} = 5, r_{21} = 8$ and $r_{12} = 3$. The first indices are $(1,1)$, $(1,4)$, and $(1,7)$.

Step 3: To fill in the remaining rows, shift the third row of the first month t one column to the right for each advancing month.

Figure 1: The three-way balanced 3-level 5-8-3 design.

α	1								2								3								4								5								6	
g	1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8	1	2
t	u_8	I	II	u_7	I	II	u_6	I	II									u_5	I	II	u_4	I	II	u_3	I	II	u_2	I	II	u_1	I	II										
$t+1$		u_8	I	II	u_7	I	II	u_6	I	II									u_5	I	II	u_4	I	II	u_3	I	II	u_2	I	II	u_1	I	II									
M $t+2$			u_8	I	II	u_7	I	II	u_6	I	II									u_5	I	II	u_4	I	II	u_3	I	II	u_2	I	II	u_1	I	II								
O $t+3$				u_8	I	II	u_7	I	II	u_6	I	II									u_5	I	II	u_4	I	II	u_3	I	II	u_2	I	II	u_1	I	II							
N $t+4$					u_8	I	II	u_7	I	II	u_6	I	II									u_5	I	II	u_4	I	II	u_3	I	II	u_2	I	II	u_1	I	II						
T $t+5$						u_8	I	II	u_7	I	II	u_6	I	II									u_5	I	II	u_4	I	II	u_3	I	II	u_2	I	II	u_1	I	II					
H $t+6$							u_8	I	II	u_7	I	II	u_6	I	II									u_5	I	II	u_4	I	II	u_3	I	II	u_2	I	II	u_1	I	II				
$t+7$								u_8	I	II	u_7	I	II	u_6	I	II									u_5	I	II	u_4	I	II	u_3	I	II	u_2	I	II	u_1	I	II			
$t+8$									u_8	I	II	u_7	I	II	u_6	I	II									u_5	I	II	u_4	I	II	u_3	I	II	u_2	I	II	u_1	I	II		
$t+9$										u_8	I	II	u_7	I	II	u_6	I	II									u_5	I	II	u_4	I	II	u_3	I	II	u_2	I	II	u_1	I	II	
$t+10$											u_8	I	II	u_7	I	II	u_6	I	II									u_5	I	II	u_4	I	II	u_3	I	II	u_2	I	II	u_1	I	II
$t+11$												u_8	I	II	u_7	I	II	u_6	I	II									u_5	I	II	u_4	I	II	u_3	I	II	u_2	I	II	u_1	I
$t+12$													u_8	I	II	u_7	I	II	u_6	I	II									u_5	I	II	u_4	I	II	u_3	I	II	u_2	I	II	u_1
$t+13$														u_8	I	II	u_7	I	II	u_6	I	II									u_5	I	II	u_4	I	II	u_3	I	II	u_2	I	II

Figure 1 illustrates how to construct the 3-level 5-8-3 design by Algorithm 1. The order in Step 1 is interpreted as the time order for a sample unit to be introduced in sample. For example, the sample unit indexed by $(4,6)$ which is denoted by u_1 at month t is introduced at month t, the unit indexed by $(4,7)$ is introduced at month $t+1$ and so on. By the rotation system of the 3-level 5-8-3 rotation design, a sample unit introduced at month t returns to the sample at months, $t+3$, $t+6$, $t+9$, $t+12$, $t+23$, $t+26$, and $t+29$. This implies that the notation u_i stands for the corresponding unit α in group g interviewed for the ith time (i.e, $i = 1, 2, \cdots, 8$). For example, because the sample unit indexed by $(5,1)$ is introduced at month $t+3$, it is denoted by u_1 at month $t+3$, by u_2 at month $t+6$, by u_3 at month $t+9$, and so on. The symbols "I" and "II" above the sample unit u_i means that the sample unit u_i provides the information for the two previous months. The recall time of u_i is 0 at the very survey month; the recall time of u_i represented by "I" is 1, one month prior to the survey month; and the recall time of u_i represented by "II" is 2, two months prior to the survey month. Therefore, Step 2 and Step 3 ensure that each monthly sample is balanced in interview times and recall times.

Note that if the selected $\sum_{i=1}^{m} r_{1i}$ sample units in Step 2 are from different rotation groups, all subsequence monthly samples contain all rotation groups because the rotation group indices are arranged in increasing order in Step 1. This means that a l-level $r_{11} - r_{21} - \cdots - r_{2,m-1} - r_{1m}$ design is three-way balanced when we have all $\sum_{i=1}^{m} r_{1i}$ rotation groups in Step 2. In particular, when we assume $l = 1$, $r_{11} = r_{12} = \cdots = r_{1m}$, and $r_{21} = \cdots = r_{2,m-1}$ in the three-way balanced design, the design is reduced to the two-way balanced design of Park, Kim and Choi (2001). Thus, the 2-way balanced designs (i.e., when $l = 1$) constructed by Algorithm 1 is more general than the previous two-way balanced design. Since the number of rotation groups used in monthly sample ($= \sum_{i=1}^{m} r_{1i}$) equals to the number of months that a sample unit is included in the sample, the three-way balanced design is also a multi-level extension of Cantwell's one-level rotation design.

2.2 l/G Rotation Designs

Suppose that a l-level rotation design has G rotation groups with the recall level $l \leq G$. At any survey month, only one rotation group is surveyed. Then the rotation group is out of the sample for the next $G-1$ consecutive months before it returns to the sample again. This procedure is repeated until the survey is completed. During the $G-1$ out-of-sample months, the remaining $G-1$ rotation groups sequentially enter to the sample and follow the same rotation pattern as described above. We call this l-level rotation design as l/G rotation design. The U.S. NAMCS and SIPP can be expressed by 1/5 and 4/4 rotation designs, respectively.

Unlike the three-way balanced design, incoming and outgoing sample units in l/G design are from different rotation group and the same sample unit is repeatedly used during the life of the survey. Thus, l/G design can not be three-way balanced. The one-level rotation design by Rao and Graham (1964) is an example of $1/G$ design. Moreover, one level-one group design or 1/1 design (i.e. $l = G = 1$) is equivalent to the usual fixed sampling in which the sample group selected at the initial month of survey is repeatedly used for the entire life of survey.

3 Extended Generalized Composite Estimators

Some sample units are used repeatedly for a pre-determined number of months according to their rotation pattern. We can obtain efficient estimator by using such repeated information of the same sample unit. The current composite estimator by Rao and Graham (1964) and the generalized composite estimators (GCE) by Breau and Ernst (1983) are good examples.

Denote $x_{t,i}^{(j)}$ be the measurement for month t obtained from the jth recall of the sample unit interviewed for the ith time at month $t+j$. In a l-level rotation design, since information for month t is obtained from $x_{t,i}^{(0)}$ through $x_{t,i}^{(l-1)}$ for $i = 1, \ldots, G$, we should wait $l-1$ months more for complete information of month t. Thus, we need preliminary estimators for month t at survey months from t to $t+l-2$ until we have a final estimator for month t which is obtained at month $t+l-1$. The followings are those $l-1$ preliminary estimators and one final estimator. For each recall time $j = 0, 1, \cdots, l-1$,

$$y_t^{(j)} = \sum_{i=1}^{G}\sum_{j'=0}^{j} a_{i,j'}^{(j)} x_{t,i}^{(j')} - \omega \sum_{i=1}^{G}\sum_{j'=0}^{j} b_{i,j'}^{(j)} x_{t-1,i}^{(j')} + \omega y_{t-1}^{(j)} \quad \text{for } j = 0, 1, \cdots, l-1 \tag{1}$$

where $0 \le \omega < 1$ and $\sum_{i=1}^{G}\sum_{j'=0}^{j} a_{i,j'}^{(j)} = \sum_{i=1}^{G}\sum_{j'=0}^{j} b_{i,j'}^{(j)} = 1$.

We call these l estimators as extended generalized composite estimators (EGCE). We can interpret the EGCE as follows: $y_t^{(0)}$ is obtained at the survey month t, and updated by $y_t^{(1)}$ one month later with the recall information from month $t+1$; the $y_t^{(1)}$ is again updated by $y_t^{(2)}$ two months later with the recall from month $t+2$ and so on until we have $y_t^{(l-1)}$. The $y_t^{(j)}$ in (1) uses only available information for month t, obtained from $j+1$ months from month t to month $t+j$. $y_t^{(j)}$ can be interpreted as an interim estimator of a characteristic for month t, which leads to the final estimator $y_t^{(l-1)}$ when all the information is available.

Define $\mathbf{a}_j = (a_{10}^{(j)}, \cdots, a_{G0}^{(j)}, \cdots, a_{1j}^{(j)}, \cdots, a_{Gj}^{(j)})'$ $\mathbf{b}_j = (b_{10}^{(j)}, \cdots, b_{G0}^{(j)}, \cdots, b_{1j}^{(j)}, \cdots, b_{Gj}^{(j)})'$, and $\mathcal{X}_t^{(j)} = (x_{t1}^{(0)}, \cdots, x_{tG}^{(0)}, \cdots, x_{t1}^{(j)}, \cdots, x_{tG}^{(j)})'$. Then, EGCE given in (1) can be written by a matrix form:

$$y_t^{(j)} = \mathbf{a}_j'\mathcal{X}_t^{(j)} - \omega\mathbf{b}_j'\mathcal{X}_{t-1}^{(j)} + \omega y_{t-1}^{(j)} \quad \text{for} \quad j = 0, 1, \cdots, l-1 \tag{2}$$

where $\mathbf{a}_j'\mathbf{1} = \mathbf{b}_j'\mathbf{1} = 1$ for all j.

When $l = 1$, EGCE is reduced to the one-level GCE of Breau and Ernst (1983) and hence the previous current and $A - K$ composite estimators are included in EGCE as special cases. When we let $\mathbf{a}_j = \mathbf{1}/jG$ where $\mathbf{1}$ is an unit vector and $\omega = 0$, $y_t^{(j)}$ defined in (2) is reduced to a simple estimator. It can be also shown that the 2-step composite estimator by Wolter (1979) is a special case of EGCE for $l = 2$ by choosing $\mathbf{a}_0 = \mathbf{b}_0 = \mathbf{1}/G$, $a_{i0}^{(1)} = b_{i0}^{(1)} = \beta$ where $0 \leq \beta \leq 1$, and $a_{i1}^{(1)} = b_{i1}^{(1)} = 1 - \beta$ for all $i = 1, \ldots, G$.

3.1 Bias and Variance of EGCE in Three-Way Balanced Design

It is common in a l-level rotation design that the expected random measurements from a sample unit may depend on its rotation group, interview time, and recall time. Let τ_g, η_i, and ξ_j be biases from the gth rotation group, the ith interview time, and the jth recall time. Because three-way balanced design has $G(= \sum_{i=1}^{m} r_{1i})$ rotation groups and interview time and l recall times, we assume that

$$E(x_{t,i}^{(j)}) = \mu_t + \tau_g + \eta_i + \xi_j, \quad \text{for} \quad g, i = 1, \ldots, G \quad \text{and} \quad j = 0, \ldots, l-1 \tag{3}$$

where μ_t is the monthly level to be estimated and the subscript g indicates the index for the rotation group containing the measurement $x_{t,i}^{(j)}$. This rotation group is uniquely determined in three-way balanced design since there is one-to-one correspondence between the rotation group and the sample unit producing $x_{t,i}^{(j)}$ by three-way balancing.

To establish such one-to-one correspondence, let $g_t(i)$ be the rotation group interviewed for the ith time at month t, in which the rotation group is interviewed through its sample unit. The three-way balancing ensures that there is the unique interview time m_k such that $g_t(m_k) = g_{t+k+1}(1)$ where $k = 0, 1, \ldots, G-1$. For example, in the 3-level 4-8-4 design in Figure 1, the rotation group indexed by 7 is introduced at month $t+1$, the rotation group indexed by 8 at month $t+2$, and the rotation groups indexed from 1 to 6 at respective months from $t+3$ to $t+8$. Since the rotation groups indexed by 7 and 8 are interviewed for 6th and

3rd times at month t, respectively, we have $m_0 = 6$ and $m_1 = 3$. Similarly, we have $m_2 = 8$, $m_3 = 5$, $m_4 = 2$, $m_5 = 7$, $m_6 = 4$, and $m_7 = 1$. This is also true for any t. Using this m_k, define L to be $G \times G$ matrix with the (i, j)th element

$$(L)_{i,j} = \begin{cases} 1 & \text{if } (i,j) \in \{(m_k, m_{k-1}); k = 1, 2, \cdots, G-1\} \text{ or } (i,j) = (m_0, m_{G-1}), \\ 0 & \text{otherwise.} \end{cases} \quad (4)$$

Then we have the following result.

Lemma 3.1. *For* $t_0 = 1, 2, \ldots$, *let* $L^{t_0} = L^{t_0 - 1}L$ *where* $L^0 = I$. *Then, by three-way balancing,* $(L^{t_0})_{i,j} = 1$ *implies that the two rotation groups interviewed for the ith time at month t and interviewed for the jth time at month* $t + t_0$ *are the same.*

By this lemma, $g_t(i) = g_{t+t_0}(j)$ if $(L^{t_0})_{i,j} = 1$ and $g_t(i) \neq g_{t+t_0}(j)$ if $(L^{t_0})_{i,j} = 0$. Thus, the matrix L^{t_0} is used to identify whether or not two sample units interviewed at two months t and $t + t_0$ from the same rotation group. As an example, consider 3-level 5-8-3 design in Figure 1. The matrix L for this 3-level 4-8-4 design is $(L)_{i_1,i_2} = 1$ if $(i_1, i_2) \in \{(1,4), (2,5), (3,6), (4,7), (5,8), (6,1), (7,2), (8,3)\}$ and $(L)_{i_1,i_2} = 0$ otherwise. Thus we have $(L)^3_{i_1,i_2} = 1$ if $(i_1, i_2) \in \{(1,2), (2,3), (3,4), (4,5), (5,6), (6,7), (7,8), (8,1)\}$ and $(L)_{i_1,i_2} = 0$ otherwise. That is, the two sample units interviewed for the 8th and the 1st times and the ith and the $i+1$ times for $i = 1, \ldots, 7$ at the respective months t and $t+3$ are from the same rotation group.

Let $\boldsymbol{\tau}_t = (\tau_{g_t(1)}, \tau_{g_t(2)}, \cdots, \tau_{g_t(G)})'$, $\boldsymbol{\eta} = (\eta_1, \eta_2, \cdots, \eta_G)'$, and $\boldsymbol{\xi}_j = (\xi_0, \xi_1, \cdots, \xi_j)'$ for $j = 0, \ldots, l-1$. Using L matrix, define $\mathcal{L}_{j,k} = (L'^{-k} : L'^{1-k} : \cdots : L'^{j-k})'$. Then, we have

Lemma 3.2.

$$\begin{aligned} E(y_t^{(j)}) &= \mu_t + \frac{1}{1-\omega^G}\mathbf{a}_j' \sum_{k=0}^{G-1} \omega^k \mathcal{L}_{j,k} \boldsymbol{\tau}_t + \frac{1}{1-\omega}\mathbf{a}_j' \mathbf{1}_j \otimes \boldsymbol{\eta} + \frac{1}{1-\omega}\mathbf{a}_j' \boldsymbol{\xi}_j \otimes \mathbf{1}_G \\ &\quad - \frac{1}{1-\omega^G}\mathbf{b}_j' \sum_{k=0}^{G-1} \omega^{k+1} \mathcal{L}_{j,k+1} \boldsymbol{\tau}_t - \frac{\omega}{1-\omega}\mathbf{b}_j' \mathbf{1}_j \otimes \boldsymbol{\eta} - \frac{\omega}{1-\omega}\mathbf{b}_j' \boldsymbol{\xi}_j \otimes \mathbf{1}_G \end{aligned}$$

where $\mathbf{1}_k$ *is the* $k \times 1$ *unit vector and* $\otimes$ *is Kronecker product.*

Since $\mathcal{L}_{j,k}$ is a permutation matrix and $\boldsymbol{\tau}_t$ is a permutation of $\boldsymbol{\tau}$, the bias of $y_t^{(j)}$ is invariant to survey month t and is a weighted average of rotation group biases, interview time biases, and recall time biases. In particular, when we assume that, for each $j = 0, 1, \ldots, l-1$, $a_{ik}^{(j)} = a_k^{(j)}$ regardless i for $k = 0, 1, \ldots, j$ (i.e., $\mathbf{a}_j = (a_0^{(j)}, a_1^{(j)}, \cdots, a_j^{(j)})' \otimes \mathbf{1}_G$) and similarly, $\mathbf{b}_j = (b_0^{(j)}, b_1^{(j)}, \cdots, b_j^{(j)})' \otimes \mathbf{1}_G$ such as Wolter's 2-step composite estimator and simple estimator, one can show that Lemma 3.2 is reduced to

$$E(y_t^{(j)}) = \mu_t + \frac{1}{G}\sum_{i=1}^{G}(\tau_i + \eta_i) + \frac{1}{1-\omega}G\sum_{k=0}^{j}(a_k^{(j)} - \omega b_k^{(j)})\xi_k. \tag{5}$$

In this case, $E(y_t^{(j)} - y_{t-t_0}^{(j)}) = \mu_t - \mu_{t-t_0}$ for $t_0 = 1, 2, \cdots$ which implies that $y_t^{(j)}$ is unbiased for changes such as monthly and yearly changes. The expected revision defined as $E(y_t^{(j+1)} - y_t^{(j)})$ for $j = 0, 1, \ldots, l-2$ depends only on recall time bias. This is not true in non three-way balanced design.

Remark 3.1. The ith rotation group bias τ_i may be different when it is represented by different sample units (Cantwell and Caldwell 1998). Namely, when we measure $x_{t,i}^{(j)}$ from the sample unit indexed by (α, g), $E(x_{t,i}^{(j)}) = \mu_t + \tau_g + \gamma_\alpha + \eta_i + \xi_j$ where γ_α is the effect of the αth panel. Then, the bias of EGCE given in Lemma 3.2 varies with survey month t because α varies with t. This means that γ_α and μ_t are confounded and they are not separately estimable. Thus, without a better alternative, it is practical that an average of panel biases included in each month t is taken to be part of the monthly effect μ_t so that μ_t in Lemma 3.2 reflects panel bias in a wide sense.

The repeated interviews of the same sample units are more likely correlated. We call this the first-order correlation or time correlation. Furthermore, since sample units in the same rotation group are usually close to each other characteristically and more likely correlated, we call this the second-order or spatial correlation. These two types of correlations are incorporated into our variance estimation. Previous works (Rao and Graham 1964, Huang and Ernst 1981, Cantwell 1990) ignored the second-order correlation for the calculation of variances, while Kumar and Lee (1983) and Park, Kim and Choi (2001) showed that the variance of the GCE is generally underestimated when the second-order correlation is ignored even for small second-order correlation.

The interview time and recall time of a sample unit may also affect the variance; hence we allow that the variance varies over the course of interviewing time and recalling level of recall time biases. In particular, when we assume that, for each $j = 0, 1, \ldots, l-1$, $a_{ik}^{(j)} = a_k^{(j)}$ regardless i for $k = 0, 1, \ldots, j$ (i.e., $\mathbf{a}_j = (a_0^{(j)}, a_1^{(j)}, \cdots, a_j^{(j)})' \otimes \mathbf{1}_G$) and similarly, $\mathbf{b}_j = (b_0^{(j)}, b_1^{(j)}, \cdots, b_j^{(j)})' \otimes \mathbf{1}_G$ such as Wolter's 2-step composite estimator and simple estimator, one can show that Lemma 3.2 is reduced to

$$E(y_t^{(j)}) = \mu_t + \frac{1}{G}\sum_{i=1}^{G}(\tau_i + \eta_i) + \frac{1}{1-\omega}G\sum_{k=0}^{j}(a_k^{(j)} - \omega b_k^{(j)})\xi_k. \tag{5}$$

In this case, $E(y_t^{(j)} - y_{t-t_0}^{(j)}) = \mu_t - \mu_{t-t_0}$ for $t_0 = 1, 2, \cdots$ which implies that $y_t^{(j)}$ is unbiased for changes such as monthly and yearly changes. The expected revision defined as $E(y_t^{(j+1)} - y_t^{(j)})$ for $j = 0, 1, \ldots, l-2$ depends only on recall time bias. This is not true in non three-way balanced design.

Remark 3.1. The ith rotation group bias τ_i may be different when it is represented by different sample units (Cantwell and Caldwell 1998). Namely, when we measure $x_{t,i}^{(j)}$ from the sample unit indexed by (α, g), $E(x_{t,i}^{(j)}) = \mu_t + \tau_g + \gamma_\alpha + \eta_i + \xi_j$ where γ_α is the effect of the αth panel. Then, the bias of EGCE given in Lemma 3.2 varies with survey month t because α varies with t. This means that γ_α and μ_t are confounded and they are not separately estimable. Thus, without a better alternative, it is practical that an average of panel biases included in each month t is taken to be part of the monthly effect μ_t so that μ_t in Lemma 3.2 reflects panel bias in a wide sense.

The repeated interviews of the same sample units are more likely correlated. We call this the first-order correlation or time correlation. Furthermore, since sample units in the same rotation group are usually close to each other characteristically and more likely correlated, we call this the second-order or spatial correlation. These two types of correlations are incorporated into our variance estimation. Previous works (Rao and Graham 1964, Huang and Ernst 1981, Cantwell 1990) ignored the second-order correlation for the calculation of variances, while Kumar and Lee (1983) and Park, Kim and Choi (2001) showed that the variance of the GCE is generally underestimated when the second-order correlation is ignored even for small second-order correlation.

The interview time and recall time of a sample unit may also affect the variance; hence we allow that the variance varies over the course of interviewing time and recalling level of sample unit. Thus, the variance and covariance of $x_{t,i}^{(j)}$ and $x_{t+t',i'}^{(j')}$ are summarized below.

$$Cov(x_{t,i}^{(j)}, x_{t+t',i'}^{(j')}) = \begin{cases} \sigma_{ij}^2 & \text{if } \; t' = 0, \; i = i' \; \text{ and } \; j = j', \\ \rho_{1t'}\sigma_{ij}\sigma_{i'j'} & \text{if both are from the same unit} \\ \rho_{2t'}\sigma_{ij}\sigma_{i'j'} & \text{if both are from the different units of the same group} \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

where $\rho_{1t'}$ is the first-order correlation and $\rho_{2t'}$ is the second-order correlation between months t and $t + t'$.

By the $\prod_{i=1}^{m}(r_{1i}(l) - r_{2i})$ rotation system, a sample unit introduced for the first time at month t is interviewed for the ith time at month $t + t_i$ where $i = 1, 2, \cdots, \sum_{i=1}^{m} r_{1i}$ and $t_i = (i-1)l + \sum_{j=1}^{m-1} r_{2j} I_{[i > \sum_{\xi=1}^{j} r_{1\xi}]}$. This implies that the sample unit interviewed for i_1th time at month t is the same unit as the sample unit interviewed for i_2th times at month $t + t'$ for $t' \geq 0$ only if $t_{i_2} = t_{i_1} + t'$. This relationship can be expressed by the $G \times G$ matrix of $L_1^{t'}$ in which the (i, j)th element is 1 if $t_i = t_j - t'$ for $j \geq i$ and is 0 otherwise.

The matrix $L^{t'}$ is used to identify two sample units interviewed at two months t and $t + t'$ from the same rotation group, and the matrix $L_1^{t'}$ is used to identify two measurements obtained at months t and $t+t'$ from the same sample unit. We define the matrix $L_2^{t'} = L^{t'} - L_1^{t'}$ to distinguish the two measurements from the two different sample units but from the same rotation group interviewed at months t and $t + t'$. If $(L_2^{t'})_{ij} = 1$, two sample units at month t and $t + t'$ with recall time 0 and respective interview times i and j are different but from the same rotation group.

Therefore $L_1^{t'}$ and $L_2^{t'}$ matrices completely identify the two sample units at the respective survey months t and $t + t'$ only by their interview times: if $(L_1^{t'})_{ij} = 1$ and $(L_2^{t'})_{ij} = 0$, the two sample units with the respective interview times i and j at months t and $t + t'$ are the same unit. On the other hand, if $(L_1^{t'})_{ij} = 0$ and $(L_2^{t'})_{ij} = 1$, they are different units but from the same rotation group.

Define $\mathbf{x}_t^{(j)} = (x_{t1}^{(j)}, x_{t2}^{(j)}, \cdots, x_{tG}^{(j)})'$ where $G = \sum_{i=1}^{m} r_{1i}$. Using the two identification matrices, L_1 and L_2 matrices, we show the following lemma.

Lemma 3.3. *Suppose that l-level rotation design is balanced in 3-ways. Then under the covariance structure given in (6), we have*

$$Cov(\mathbf{x}_t^{(j_1)}, \mathbf{x}_{t+t'}^{(j_2)}) = \rho_{1t'}\Lambda_{j_1} L_1^{|t'-j_1+j_2|}\Lambda_{j_2} + \rho_{2t'}\Lambda_{j_1} L_2^{|t'-j_1+j_2|}\Lambda_{j_2} \tag{7}$$

where $j_1, j_2 = 0, 1, \cdots, l-1$ *and* $\Lambda_{j_k} = diag(\sigma_{1j_k}, \ \sigma_{2j_k}, \ \cdots, \ \sigma_{Gj_k})$ *for* $k = 1, 2$.

From Lemma 3.3, for $j = 0, 1, \cdots, l-2$,

$$Cov(\mathcal{X}_t^{(j)}, \mathcal{X}_{t+t'}^{(j)}) = \begin{pmatrix} Q_{t',0,0} & Q_{t',0,1} & \cdots & Q_{t',0,j+1} \\ Q_{t',1,0} & Q_{t',1,1} & \cdots & Q_{t',1,j+1} \\ \vdots & \vdots & \vdots & \vdots \\ Q_{t',j+1,0} & Q_{t',j+1,1} & \cdots & Q_{t',j+1,j+1} \end{pmatrix}$$

where $Q_{t',j_1,j_2} = Cov(\mathbf{x}_t^{(j_1)}, \mathbf{x}_{t+t'}^{(j_2)})$ given in (7). Denote $Cov(\mathcal{X}_t^{(j)}, \mathcal{X}_{t+t'}^{(j)})$ be $V_{t',j}$ for $j = 0, 1, \cdots, l-2$ where $V_{t',l-1} = V_{t',l-2}$. Then we have the variance of $y_t^{(j)}$ as follows.

Theorem 3.4. *Under the same assumptions given in Lemma 3.3, the variance of EGCE for recall level* $j = 0, 1, \cdots, l-1$ *is*

$$\begin{aligned} (1-\omega^2)Var(y_t^{(j)}) &= \mathbf{a}_j'(V_{0,j} + 2\omega B_{1,0}(j))\mathbf{a}_j + \omega^2 \mathbf{b}_j'(V_{0,j} + 2\omega B_{1,0}(j))\mathbf{b}_j \\ &- 2\omega \mathbf{b}_j'(B_{1,0}(j) + B_{1,1}'(j))\mathbf{a}_j \end{aligned}$$

where $B_{n_1,n_2}(j) = \sum_{k=0}^{\infty} \omega^{k+n_2} V_{n_1+k-n_2,j}$ *for* $n_1 \geq n_2$.

For each $j = 0, 1, \cdots, l-1$ and the integers $t_0, t' > 0$, we are also interested in the variances of (a) the change $y_t^{(j)} - y_{t-t_0}^{(j)}$, (b) the aggregate $S_t^{t_0}(j) = \sum_{t'=0}^{t_0-1} y_{t-t'}^{(j)}$, and (c) the change of two aggregates $S_t^{t_0}(j) - S_{t-t_1}^{t_0}(j) = \sum_{i=0}^{t_0-1} y_{t-i}^{(j)} - \sum_{i=0}^{t_0-1} y_{t-t_1-i}^{(j)}$ for $t_1 \geq t_0$. For the variances of these three estimators, let $P_{1j}(t^*) = \frac{2(1-\omega^{t^*})}{1-\omega^2}(V_{0,j} + 2\omega B_{1,0}(j)) - 2\sum_{n=0}^{t^*-1} B_{t^*,n}(j)$, $P_{2j}(t^*) = -\frac{\omega(1-\omega^{t^*})}{1-\omega^2}(B_{1,0}(j) + B_{1,1}'(j)) + 2\sum_{n=0}^{t^*-1}\left(\omega B_{t^*+1,n}(j) + B_{t^*,n+1}'(j)\right)$ and $P_{3j}(t^*) = \omega^2 P_{1j}(t^*)$ for $t^* \geq 1$. Then using these definitions, we show the variances of the three estimators as follows.

Theorem 3.5. *Under the same conditions in Theorem 3.4,*

$$(a)\quad Var(y_t^{(j)} - y_{t-t_0}^{(j)}) = \mathbf{a}_j' P_{1j}(t_0)\mathbf{a}_j + \mathbf{b}_j' P_{2j}(t_0)\mathbf{a}_j + \mathbf{b}_j' P_{3j}(t_0)\mathbf{b}_j,$$

$$(b)\quad Var(S_t^{t_0}(j)) = \mathbf{a}_j'\big(t_0 Q_{1j} - \sum_{t^*=1}^{t_0-1}(t_0 - t^*)P_{1j}(t^*)\big)\mathbf{a}_j$$

$$+\mathbf{b}_j'\big(t_0 Q_{2j} - \sum_{t^*=1}^{t_0-1}(t_0 - t^*)P_{2j}(t^*)\big)\mathbf{a}_j + \mathbf{b}_j'\big(t_0\omega^2 Q_{1j} - \sum_{t^*=1}^{t_0-1}(t_0 - t^*)P_{3j}(t^*)\big)\mathbf{b}_j$$

and

$$(c)\quad Var(S_t^{t_0}(j) - S_{t-t_1}^{t_0}(j))$$

$$= \mathbf{a}_j'\big(\sum_{t^*=-t_0+1}^{t_0-1}(t_0 - |t^*|)P_{1j}(t_1 - t^*) - 2\sum_{t^*=1}^{t_0-1}(t_0 - t^*)P_{1j}(t^*)\big)\mathbf{a}_j$$

$$+\mathbf{b}_j'\big(\sum_{t^*=-t_0+1}^{t_0-1}(t_0 - |t^*|)P_{2j}(t_1 - t^*) - 2\sum_{t^*=1}^{t_0-1}(t_0 - t^*)P_{2j}(t^*)\big)\mathbf{a}_j$$

$$+\mathbf{b}_j'\big(\sum_{t^*=-t_0+1}^{t_0-1}(t_0 - |t^*|)P_{3j}(t_1 - t^*) - 2\sum_{t^*=1}^{t_0-1}(t_0 - t^*)P_{3j}(t^*)\big)\mathbf{b}_j.$$

Since the simple estimator is defined as $y_t^{(j)} = \mathbf{1}'\mathcal{X}_t^{(j)}/jG$, it can be shown that $P_{1j}(t^*) = 2(V_{0,j} - V_{t^*,j})$ and $P_{2j}(t^*) = P_{3j}(t^*) = 0$. The variances for the simple estimator are easily obtained by letting $\omega = 0$ and $\mathbf{a}_j = \mathbf{1}/jG$ in Theorem 3.4 and 3.5.

If the second-order correlation $\rho_{2t} = 0$ for all t, the L_2 matrix is no longer necessary in variance estimation for the l-level $r_{11} - r_{21} - \cdots - r_{2,m-1} - r_{1m}$ design because the matrix L_2 is related only to the second-order correlation. Thus, the following corollary can be stated for this situation.

Corollary 3.6. *Suppose that $\rho_{2t} = 0$ for all t is the l-level $r_{11} - r_{21} - \cdots - r_{2,m-1} - r_{1m}$ design. By letting $L_2^t = 0$ for all t in Lemma 3.3, Theorem 3.4, and 3.5 still hold for any l-level $r_{11} - r_{21} - \cdots - r_{2,m-1} - r_{1m}$ design without balancing on three-ways.*

Note that when $l = 1$ in the l-level $r_{11} - r_{21} - \cdots - r_{2,m-1} - r_{1m}$ design, the variances given in Corollary 3.6 are the same as those of Cantwell (1990).

3.2 Bias and Variances of l-level composite estimators in l/G design

Let $x_t^{(j)}$ be the measurement for month t by recalling the jth prior month from the interview month $t + j$ for $j = 0, 1, \ldots, l-1$. Because there is only one rotation group for each recall time j in l/G design, we define EGCE for l/G design as follows.

$$\begin{aligned} y_t^{(j)} &= \sum_{j'=0}^{j} a_{j'}^{(j)} x_t^{(j)} - \omega \sum_{j'=0}^{j} b_{j'}^{(j)} x_{t-1}^{(j')} + \omega y_{t-1}^{(j)} \\ &= \mathbf{a}^{*\prime}_j \mathcal{X}_{t,j} - \omega \mathbf{b}^{*\prime}_j \mathcal{X}_{t-1,j} + \omega y_{t-1}^{(j)} \quad \text{for } j = 0, 1, \cdots, l-1, \end{aligned} \tag{8}$$

where $\mathbf{a}_j^* = (a_0^{(j)}, a_1^{(j)}, \cdots, a_j^{(j)})'$, $\mathbf{b}_j^* = (b_0^{(j)}, b_1^{(j)}, \cdots, b_j^{(j)})'$, and $\mathcal{X}_{t,j} = (x_t^{(0)}, x_t^{(1)}, \cdots, x_t^{(j)})'$.

For bias of $y_t^{(j)}$, we assume in l/G rotation design that the first rotation group is interviewed at the initial month $t = 0$ and then the second rotation group is interviewed at month $t = 1$, and so on. Thus, the ith interview is performed from $t = (i-1)G$ to $t = iG - 1$ where $i = 1, 2, \cdots$. Let I_j be $(j+1) \times G$ with $(I_j)_{l,m} = 1$ if $l = m$ and $(I_j)_{l,m} = 0$ if $l \neq m$ and L_3 be $G \times G$ matrix with $(L)_{ij} = 1$ if $j = i+1$ for $i = 1, \cdots, G-1$ and if $j = 1$ for $i = G$ and $(L)_{ij} = 0$ otherwise. By the similar approach used for three-way balanced design, $E(y_t^{(j)})$ in l/G design is as follows.

Lemma 3.7. *For l/G design,*

$$\begin{aligned} E(y_t^{(j)}) &= \mu_t + \frac{1}{1-\omega^G} \mathbf{a}^{*\prime}_j \sum_{k=0}^{j} \omega^k I_j L_3^{t-k} \boldsymbol{\tau} - \frac{\omega}{1-\omega^G} \mathbf{b}^{*\prime}_j \sum_{k=0}^{j} \omega^k I_j L_3^{t-k-1} \boldsymbol{\tau} \\ &\quad + \frac{1}{1-\omega} (\mathbf{a}^{*\prime}_j - \omega \mathbf{b}^{*\prime}_j) \boldsymbol{\xi}_j. \end{aligned}$$

where $\boldsymbol{\tau}$ and $\boldsymbol{\xi}_j$ are given as before.

As for $I_j L_3^{t-k} \boldsymbol{\tau}$ in Lemma 3.7, $L_3^{t-k} \boldsymbol{\tau}$ is a permutation of $\boldsymbol{\tau}$ and I_j selects the first j elements from this permuted $\boldsymbol{\tau} = (\tau_1, \tau_2, \cdots, \tau_G)'$. Thus, unlike three-way balanced design, the bias of EGCE for l/G varies with survey month t because time dependence of $I_j L_3^{t-k} \boldsymbol{\tau}$. Since $L_3^t = L_3^{t+G}$, $E(y_t^{(j)}) = E(y_{t+G}^{(j)})$. This implies that l/G design has a cyclic bias with G months as one cycle period.

For the simple estimator, let $\mathbf{a}^*_j = \mathbf{1}_{j+1}/(j+1)$ and $\omega = 0$ where $\mathbf{1}_{j+1}$ is $(j+1) \times 1$ unit vector. Then, Lemma 3.7 is reduced to

$$E(y_t^{(j)}) = \mu_t + \frac{1}{j+1} \mathbf{1}'_{j+1} I_j L_3^t \boldsymbol{\tau} + \frac{1}{j+1} \sum_{k=0}^{j} \xi_k. \tag{9}$$

Thus, the bias of simple estimator in l/G design also vary with survey month t.

Remark 3.2. Interview time bias also exist in l/G design. Since interview time increases as time advances, interview time bias and monthly effect μ_t are confounded in l/G design like as panel bias and μ_t are confounded in three-way balanced design. Thus, an average of interview time biases is practically taken part of μ_t in l/G design.

Comparing to the simple estimator in the three-way balanced design whose expectation is $\mu_t + \frac{1}{G}\sum_{i=1}^{G}(\tau_i + \eta_i) + \frac{1}{j+1}\sum_{k=0}^{j}\xi_k$ from (5), the simple estimator for l/G design whose expectation is given in (9) is biased for characteristics such as monthly and yearly changes while that for three-way balanced design is unbiased. Since the expected revision of the simple estimator in l/G design is, from (9), $E(y_t^{(j+1)}) - E(y_t^{(j)}) = \frac{1}{j+2}\mathbf{1}'_{j+2}I_{j+1}L_3^t\boldsymbol{\tau} - \frac{1}{j+1}\mathbf{1}'_{j+1}I_jL_3^t\boldsymbol{\tau} + \frac{1}{j+2}\sum_{k=0}^{j+1}\xi_k - \frac{1}{j+1}\sum_{k=0}^{j}\xi_k$ while that in three-way balanced design is $\frac{1}{j+2}\sum_{k=0}^{j+1}\xi_k - \frac{1}{j+1}\sum_{k=0}^{j}\xi_k$ the simple estimator in l/G design needs the extra revision of $\frac{1}{j+2}\mathbf{1}'_{j+2}I_{j+1}L_3^t\boldsymbol{\tau} - \frac{1}{j+1}\mathbf{1}'_{j+1}I_jL_3^t\boldsymbol{\tau}$ than that in three-way balanced design. Moreover, if $\sum_{i=1}^{G}\tau_i = \sum_{i=1}^{G}\eta_i = \sum_{j=0}^{l-1}\xi_j = 0$ can be assumed, the simple estimator as an final estimator in l/G design is biased for μ_t but that in three-way balanced design is unbiased.

Since no second-order correlation exists in l/G design, if $x_t^{(j)}$ and $x_{t+t'}^{(j')}$ are the measurements from the same sample unit, the covariance structure in l/G design is

$$Cov\big(x_t^{(j)}, x_{t+t'}^{(j')}\big) = \rho_{1,t'}\sigma_j\sigma_{j'} \tag{10}$$

where $j, j' = 0, 1, \cdots, l-1$. Now, since each of G rotation groups returns to the sample for every Gth month in l/G rotation sampling design, the rotation pattern can be expressed as $g_t^{(j)} = g_{t+nG}^{(j)}$ for $n = 0, 1, \cdots$ and for each $j = 0, 1, \cdots, l-1$ where the $g_t^{(j)}$ is the rotation group with the recall level j at month t. Hence, the covariance of (10) can be expressed as

$$Cov\big(x_t^{(j)}, x_{t+t'}^{(j')}\big) = \rho_{1,t'}\sigma_j\sigma_{j'}\ d\big(mod_G(|t' + j' - j|), 0\big)$$

where $d\big(mod_G(|t' + j' - j|), 0\big) = 1$ only if $mod_G(|t' + j' - j|) = 0$. This expression yields $Cov(\mathcal{X}_{t,k}, \mathcal{X}_{t+t',k}) \equiv V^*_{t',k}$ with its (i,j)th element, for $k = 0, 1, \ldots, l-1$,

$$\rho_{1t'}\sigma_{i-1}\sigma_{j-1}d\big(mod_G(|t' + j - i|), 0\big)$$

where $i, j = 1, 2, \ldots, k+1$.

Since the EGCE for l/G design given by (8) is the same form as the EGCE for three-way balanced design given by (2), we can obtain the variances of monthly level, aggregate level, and changes of monthly and aggregate levels for the l/G design from Theorem 3.4 and 3.5 as summarized in the following corollary.

Corollary 3.8. *Suppose that a multi-level rotation design follows the l/G rotation system. When we replace $\mathbf{a}_j$, $\mathbf{b}_j$ and $V_{t',j}$ by $\mathbf{a}_j^*$, $\mathbf{b}_j^*$ and $V_{t',j}^*$, respectively and let $\rho_{2t} = 0$ for all t in Theorem 3.4 and 3.5, all results provided in these two theorems for the three-way balanced design hold for the l/G rotation design of l-level recall with G groups.*

Wolter (1979) derived the variance of his own GCE in 2/3 rotation design. His variance formula is obtainable by applying Corollary 3.8 when we let $a_0^{(0)} = 1$ and $b_0^{(0)} = 0$ for $j = 0$ and $(a_0^{(1)}, a_1^{(1)}) = (\omega_1, 1 - \omega_1)$ and $(b_0^{(1)}, b_1^{(1)}) = (0, 1)$ for $j = 1$ where $0 < \omega_1 < 1$.

4 Optimal coefficients of the l-level GCE

When we have many different estimators and these estimators need to be consistent, it may not be possible in practice to use different coefficients for every different estimators. Therefore, for this practical reason, we obtain one set of coefficients of the l-level GCE by minimizing a weighted sum of the MSEs of monthly level, aggregate level, and changes of monthly and aggregate levels.

We have the four types of l-level GCE : $y_t^{(j)}$, $y_t^{(j)} - y_{t-t_0}^{(j)}$, $S_t^{t_0}(j)$ and $S_t^{t_0}(j) - S_{t-t_1}^{t_0}(j)$ for $t_0 \geq 1$, $t_1 \geq t_0$ and each $j = 0, 1, \cdots, l-1$. Defining specific values of t_0 and t_1, we assume that there are H l-level GCEs of interest, and denote them by z_{tjh} for $h = 1, 2, \cdots, H$. Note that z_{tjh} and $z_{tjh'}$, $h \neq h'$ can be the same type of l-level GCE even when they are of different characteristics. For example, $y_t^{(j)}$ for Labor Force is the same notation as $y_t^{(j)}$ for Unemployed.

To have the consistency in total among estimates, we set the object function O_j, $j = 0, 1, \cdots, l-1$ for MSE: $O_j = \sum_{h=1}^{H} \delta_h MSE(z_{tjh}) - \lambda_1(\mathbf{1}'\mathbf{a}_j - 1) - \lambda_2(\mathbf{1}'\mathbf{b}_j - 1)$ where λs are the Lagrange multipliers and δs are the weights which represent the relative importance of the corresponding estimators. For example, all the H l-level GCEs are equally valuable, then the weight $\delta_h = 1/H$ for all h. By a suitable choice of z_{tjh}, we obtain one set of optimal coefficients by minimizing the object function O_j for each level $j = 0, 1, \cdots, l-1$. Then we can use it commonly for the estimators of different characteristics.

By Lemma 3.2, Theorems 3.4 and 3.5 for the 3-way balanced l-level $r_{11} - r_{21} - \cdots - r_{2,m-1} - r_{1m}$ design, $MSE(z_{tjh})$ can be expressed as

$$Var(z_{tjh}) = \mathbf{a}_j' C_{1jh}\mathbf{a}_j + \mathbf{b}_j' C_{2jh}\mathbf{a}_j + \mathbf{b}_j' C_{3jh}\mathbf{b}_j. \tag{11}$$

For example, when $z_{tjh} = y_t^{(j)}$, $\mathbf{C}_{kjh} = P_{kj} + B_{kj}^2$ for $k = 1, 2, 3$ where P_{kj} is given in Theorem 3.4 and B_{kj} is the bias from Lemma 3.2.

Using (11), we optimize the object function O_j and obtain the following optimal coefficients in the 3-way balanced $r_{11} - r_{21} - \cdots - r_{2,m-1} - r_{1m}$ design. Hereafter, we call these optimal coefficients as compromise coefficients to distinguish it from all other coefficients.

Lemma 4.1. *Suppose that a l-level rotation design is balanced in 3-ways. For given weights δ_h, $h = 1, 2, \cdots, H$, the compromise coefficients of $\mathbf{a}_j$ and $\mathbf{b}_j$, minimizing the weighted MSE $\sum_{h=1}^{H} \delta_h MSE(z_{tjh})$, are given by*

$$\begin{pmatrix} \widehat{\mathbf{a}}_j \\ \widehat{\mathbf{b}}_j \end{pmatrix} = \begin{pmatrix} C_{1j} + C_{1j}' & C_{2j} - s_{1j}^{-1}\mathbf{1}_j\mathbf{1}_j'(C_{1j} + C_{1j}')^{-1}C_{2j}' \\ C_{2j} - s_{3j}^{-1}\mathbf{1}_j\mathbf{1}_j'(C_{3j} + C_{3j}')^{-1}C_{2j}' & C_{3j} + C_{3j}' \end{pmatrix}^{-1} \begin{pmatrix} s_{1j}^{-1}\mathbf{1}_j \\ s_{3j}^{-1}\mathbf{1}_j \end{pmatrix}$$

where $j = 0, 1, \cdots, l-2$, $s_{1j} = \mathbf{1}_j'(C_{1j} + C_{1j}')^{-1}\mathbf{1}_j$, and $s_{3j} = \mathbf{1}_j'(C_{3j} + C_{3j}')^{-1}\mathbf{1}_j$ where $\mathbf{1}_j$ is an unit vector with an appropriate size depending on j.

When a multi-level rotation design follows the l/G rotation system, the corresponding compromise coefficients minimizing their sum of variances are easily obtained after appropriately defining $\mathbf{C}_{kj}$, for each of $k = 1, 2, 3$ and level $j = 0, 1, \cdots, l-1$. This can be done from Lemma 3.7 and Corollary 3.8 for the l/G designs.

5 Conclusion Remarks

All rotation designs studied before can be classified into one of three classes of multi-level rotation designs. Those are the life time balanced l-level $r_{11} - r_{21} - \cdots - r_{2,m-1} - r_{1m}$ design, three-way balanced l-level $r_{11} - r_{21} - \cdots - r_{2,m-1} - r_{1m}$ design, and l/G rotation design. We are mainly interested in the three-way balanced $r_{11} - r_{21} - \cdots - r_{2,m-1} - r_{1m}$ design because its basic framework and properties can be easily applied to the remaining two other classes of rotation designs as special cases.

This paper presents six major contributions to the multi-level rotation designs, and they can be summarized as follows. (1) We introduce the 3-way balanced rotation design and investigate its properties and rotation pattern. We provide the two conditions for the l-level $r_{11}-r_{21}-\cdots-r_{2,m-1}-r_{1m}$ design to be balanced in three-ways. (2) Using the basic framework of the three-way balanced design, we suggest a multi-level version of the Cantwell's one-level balanced design and establish the basic properties for the l-level $r_{11}-r_{21}-\cdots-r_{2,m-1}-r_{1m}$ design and the l/G rotation design. (3) We show that the l-level GCE is a general form of the previous one-level GCE and that all the previous estimators used in rotation sampling designs are special cases of our l-level GCE. (4) We derive the general variance formula of the l-level GCE from the three-way balanced l-level $r_{11}-r_{21}-\cdots-r_{2,m-1}-r_{1m}$ design in the presence of the first-order and second-order correlations. (5) When the second-order correlation is zero, our variance formula for the three-way balanced design are applicable to any l-level $r_{11}-r_{21}-\cdots-r_{2,m-1}-r_{1m}$ design regardless three-way balancing. After slight modification of the variance formula of the 3-way balanced design, we can obtain the variance of the l-level GCE of the l/G design as a special case. (6) To preserve the additivity of estimates in total, we obtain one set of common coefficients of the l-level GCE for each of the three types of rotation designs. These coefficients are optimal in the sense that they minimize a weighted sum of variances of all concerned characteristics of interest.

$-1/2Var(y_t^{(j)} - y_{t-t'}^{(j)})$, (??) is

References

Bailar, B. (1975). The effects of rotation group bias on estimates from panel survey, *Journal of the American Statistical Association*, **70**, 23-30.

Breau, P. and Ernst, L. (1983). Alternative estimators to the current composite estimators. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 397-492.

Cantwell, P.J. (1990). Variance formulae for composite estimators in rotation designs. *Survey Methodology*, **16**, 153-163.

Cantwell, P.J. and Caldwell, C.V. (1998). Examining the revisions in Monthly Retail and Wholesale Trade Surveys under a rotating panel design. *Journal of Official Statistics*, **14**, 47-59.

Huang, E.T. and Ernst, L. R. (1981). Comparison of an alternative estimator to the current composite estimator in CPS. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 303-308.

Kumar, S. and Lee, H. (1983). Evaluation of composite estimation for the Canadian Labor Force Survey. *Survey Methodology*, 9, 403-408.

Park, Y.S., Kim, K.W. and Choi, J. (2001). One-level rotation design balanced on time in monthly sample and in rotation group. *Journal of American Statistical Association*. 96, 1483-1496.

Rao, J.N.K. and Graham, J.E. (1964). Rotation designs for sampling on repeated occasions. *Journal of the American Statistical Association*, **59**, 492-509.

Wolter, K. (1979). Composite estimation in finite population. *Journal of the American Statistical Association*, **74**, 604-613.