Journal of the Korean Statistical Society (2003), **32**: *3, pp 245–259*

THREE-WAY BALANCED MULTI-LEVEL ROTATION SAMPLING DESIGNS†

Y. S. PARK[1], K. W. KIM[1] AND N. Y. KIM[1]

ABSTRACT

The 2-way balanced one-level rotation design has been discussed (Park *et al.*, 2001), where the 2-way balancing is done on interview time in monthly sample and rotation group. We extend it to 3-way balanced multi-level design to obtain more information of the same sample unit for one or more previous months. The 3-way balancing is accomplished not only on interview time in monthly sample and rotation group but also on recall time as well. The 3-way balancing eliminates or reduces any bias arising from unbalanced interview time, rotation group and recall time, and all rotation groups are equally represented in the monthly sample. We present the rule and rotation algorithm which guarantee the 3-way balancing. In particular, we specify the necessary and sufficient condition for the 3-way balanced multi-level rotation design.

AMS 2000 subject classifications. Primary 62D05.
Keywords. 3-way balancing, multi-level rotation design, necessary and sufficient condition.

1. INTRODUCTION

Rotation sampling designs may be classified into two categories by the frequency that a selected sample unit appears in the sample. The first type uses the same sample unit for the entire life of the survey. The typical examples of such design are the U. S. Monthly Retail Trade (MRTS), the U. S. Monthly Wholesale Trade Surveys (these two surveys recently changed their designs from rotating panels to a single fixed panel) and the Survey of Income and Program Participation (SIPP). The second type uses the sample unit only for a fixed number of times, and we call it semi-rotation sampling design. The U. S. Current

Received January 2003; accepted April 2003.
†This research was supported by a Korea University Grant.
[1]Department of Statistics, Korea University, Seoul 136-701, Korea (e-mail : yspark@korea.ac.kr)

Population Survey (CPS), the Canadian Labor Force Survey (CLFS) and U. S. Consumer Expenditure Survey (CEX) belong to this category.

In both types of designs the entire sample or population is partitioned into a finite number of rotation groups and each group includes a number of sample units. In the first type, we systematically select some of the rotation groups and then use all sample units in each selected group for the monthly sample. However, in the semi-rotation design all rotation groups are included in the sample, but the sample unit in a certain group is replaced by a new unit from the same group following a specific rotation plan. Thus, the respondents in the semi-rotation plan have less response burden than those in the first type.

The first type have been studied extensively for one-level and multi-level rotation designs (Rao and Graham, 1964; Wolter, 1979; Cantwell, 1990; Fuller, 1990; Cantwell and Caldwell, 1998). Here, respondents provide information only for the current survey month in an one-level rotation design. Unlike one-level design, in a multi-level rotation design respondents provide information not only for the current survey month but also for the previous survey months. In the second type, although many researchers studied one-level semi-rotation design (Woodruff, 1963; Bailer, 1975; Jones, 1980; Cantwell, 1990; Fuller *et al.*, 1992; Lent *et al.*, 1999; Mansur and Shoemaker, 1999; Park *et al.*, 2001), the multi-level aspect of this design has not been investigated before. This is the focus of out paper.

In multi-level rotation design, there are three types of possible biases in monthly sample (Bailar, 1975; Cantwell and Caldwell, 1998). These biases arise from different interview times of monthly sample (*i.e.*, rotation group bias), different rotation groups in monthly sample (*i.e.*, panel imbalance), and different recall times in monthly sample (*i.e.*, recall time bias). Because the data for a given month t are collected during the current month t as well as the $l-1$ proceeding months in l-level rotation design, we do not have complete information for month t until the month $t+l-1$. Until then we release only the preliminary estimate of month t for characteristics such as monthly level and monthly change. A month later we obtain a second-preliminary estimate, adding one month recall from month $t+1$ to the data of the previous month t. This is repeated until the final estimate for month t includes all $l-1$ recalls from month $t+1$ to month $t+l-1$. The difference between the two successive estimates for month t is called the revision to the estimate. It is desirable that this revision is as small as possible to have a stable estimate. Cantwell and Caldwell (1998) indicated that this revision was accounted mainly by the panel imbalance.

As we want to reduce the three types of biases from monthly sample and revision as much as possible, we propose a class of multi-level semi-rotation designs which is a general version of the previous one-level semi-rotation designs (Park *et al.*, 2001; Cantwell, 1990). Such multi-level semi-rotation designs are balanced on interview times in monthly sample, rotation group as well as recall time. By this 3-way balancing, all rotation groups have equal opportunity to be represented in the sample for every recall time, and some sample units are replaced by similar ones from the same rotation groups. Thus the 3-way balanced design eliminates bias arising from taking certain rotation groups more often than others.

Our main point in this paper is to present the necessary and sufficient condition for the 3-way balanced multi-level rotation design to eliminate or reduce the design biases. The remainder of this paper is divided into 5 sections. In Section 2, we extend the previous 2-way balanced one-level rotation design (Park *et al.*, 2001) to the 3-way balanced multi-level rotation design. We discuss the properties of the 3-way balanced design as well as its rotation pattern. The necessary and sufficient condition for a multi-level rotation design to be balanced in 3-ways is also provided in Section 3. In Section 4, we present an algorithm to construct 3-way balanced multi-level rotation designs. A conclusion remark is provided in Section 5. All theoretical proofs are provided in Appendix.

2. THREE-WAY BALANCED MULTI-LEVEL ROTATION SAMPLING DESIGNS

To understand the basic concept of the 3-way balanced design, we illustrate the multi-level rotation design with the "3-level" 4-8-4 design (Table 2.1) which is a multi-level version of the 4-8-4 design. The "3-level" means that each sample unit reports the information of the interview month as well as of the two previous months. In the 3-level 4-8-4 rotation design, a sample unit is in the sample every third month for a total of 4 times, gives no information for the next 8 months and finally returns to the sample every third month for another 4 times. The notation (α, g) in Table 2.1 is the index for the α^{th} sample unit in the g^{th} rotation group, and the u_i indicates that the corresponding unit α is interviewed for the i^{th} time in any given month $(i = 1, 2, \ldots, 8)$. In spite of the two different times of observation, u_6 in month t and u_7 in month $t+3$ are the same sample unit from the same group 7 since both are indexed by $(1, 7)$; but u_6 in month t and u_6 in month $t+1$ are different units since they are indexed by $(1, 7)$ and $(1, 8)$. The symbols "I" and "II" above the sample unit u_i means that the same sample unit

u_i provides the information of the 2 previous months. The recall time of the unit u_i is 0 at the very survey month, "I" right above u_i means one month recall time from the survey month and "II" means 2 months recall time from the survey month. In the first picture of Table 2.1, all 8 rotation groups are included in the sample for any survey month and each group is represented three times by 3 different sample units, each with respective recall times of 0, 1 and 2.

TABLE 2.1 *Three-way balanced multi-level 4-8-4 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	I	u_5	I	II									u_4	I	II	u_3	I	II	u_2	I	II

(i) recall time=0

month	*Rotation Groups*							
	1	2	3	4	5	6	7	8
t	u_8	u_5	u_2	u_7	u_4	u_1	u_6	u_3
$t+1$	u_3	u_8	u_5	u_2	u_7	u_4	u_1	u_6
$t+2$	u_6	u_3	u_8	u_5	u_2	u_7	u_4	u_1
$t+3$	u_1	u_6	u_3	u_8	u_5	u_2	u_7	u_4
$t+4$	u_4	u_1	u_6	u_3	u_8	u_5	u_2	u_7
$t+5$	u_7	u_4	u_1	u_6	u_3	u_8	u_5	u_2
$t+6$	u_2	u_7	u_4	u_1	u_6	u_3	u_8	u_5
$t+7$	u_5	u_2	u_7	u_4	u_1	u_6	u_3	u_8
$t+8$	u_8	u_5	u_2	u_7	u_4	u_1	u_6	u_3
$t+9$	u_3	u_8	u_5	u_2	u_7	u_4	u_1	u_6
$t+10$	u_6	u_3	u_8	u_5	u_2	u_7	u_4	u_1
$t+11$	u_1	u_6	u_3	u_8	u_5	u_2	u_7	u_4
$t+12$	u_4	u_1	u_6	u_3	u_8	u_5	u_2	u_7
$t+13$	u_7	u_4	u_1	u_6	u_3	u_8	u_5	u_2

(ii) recall time=1

month	*Rotation Groups*							
	1	2	3	4	5	6	7	8
t	u_3	u_8	u_5	u_2	u_7	u_4	u_1	u_6
$t+1$	u_6	u_3	u_8	u_5	u_2	u_7	u_4	u_1
$t+2$	u_1	u_6	u_3	u_8	u_5	u_2	u_7	u_4
$t+3$	u_4	u_1	u_6	u_3	u_8	u_5	u_2	u_7
$t+4$	u_7	u_4	u_1	u_6	u_3	u_8	u_5	u_2
$t+5$	u_2	u_7	u_4	u_1	u_6	u_3	u_8	u_5
$t+6$	u_5	u_2	u_7	u_4	u_1	u_6	u_3	u_8
$t+7$	u_8	u_5	u_2	u_7	u_4	u_1	u_6	u_3
$t+8$	u_3	u_8	u_5	u_2	u_7	u_4	u_1	u_6
$t+9$	u_6	u_3	u_8	u_5	u_2	u_7	u_4	u_1
$t+10$	u_1	u_6	u_3	u_8	u_5	u_2	u_7	u_4
$t+11$	u_4	u_1	u_6	u_3	u_8	u_5	u_2	u_7
$t+12$	u_7	u_4	u_1	u_6	u_3	u_8	u_5	u_2
$t+13$	u_2	u_7	u_4	u_1	u_6	u_3	u_8	u_5

(iii) recall time=2

month	*Rotation Groups*							
	1	2	3	4	5	6	7	8
t	u_6	u_3	u_8	u_5	u_2	u_7	u_4	u_1
$t+1$	u_1	u_6	u_3	u_8	u_5	u_2	u_7	u_4
$t+2$	u_4	u_1	u_6	u_3	u_8	u_5	u_2	u_7
$t+3$	u_7	u_4	u_1	u_6	u_3	u_8	u_5	u_2
$t+4$	u_2	u_7	u_4	u_1	u_6	u_3	u_8	u_5
$t+5$	u_5	u_2	u_7	u_4	u_1	u_6	u_3	u_8
$t+6$	u_8	u_5	u_2	u_7	u_4	u_1	u_6	u_3
$t+7$	u_3	u_8	u_5	u_2	u_7	u_4	u_1	u_6
$t+8$	u_6	u_3	u_8	u_5	u_2	u_7	u_4	u_1
$t+9$	u_1	u_6	u_3	u_8	u_5	u_2	u_7	u_4
$t+10$	u_4	u_1	u_6	u_3	u_8	u_5	u_2	u_7
$t+11$	u_7	u_4	u_1	u_6	u_3	u_8	u_5	u_2
$t+12$	u_2	u_7	u_4	u_1	u_6	u_3	u_8	u_5
$t+13$	u_5	u_2	u_7	u_4	u_1	u_6	u_3	u_8

Tables 2.1 (i), (ii) and (iii) show the 2-way balancing on interview times at the respective recall times of 0, 1 and 2. These three pictures are obtained from Table 2.1 when ignoring α. We observe that the picture of each recall time is balanced in 2-ways, horizontally and perpendicularly, where the perpendicular balancing is done for any span of 8 months as in one-level 4-8-4 design. Table 2.1 is a typical example of a multi-level rotation design which is balanced in 3-ways by interview time (horizontally in each picture), by rotation group (perpendicularly in each picture) and by recall time (the same pattern in all three pictures (i), (ii) and (iii)). This 3-way balancing ensures that all 8 rotation groups are included in

the monthly sample at any survey month and the rotation pattern of a sample unit depends only on its interview time regardless of its rotation group and recall time. This enables us to obtain the variance and MSE of estimators used in the 3-way balanced design.

The 4-8-4 design can be written as $4^2 - 8^1$. Similarly, a general form of the "one-level" rotation design is expressed as $r_1^m - r_2^{m-1}$ where a selected sample unit is interviewed for the first r_1 months, out of the sample for the next r_2 months (equivalently, no information for those months) and back to the sample for another r_1 months. This procedure is repeated m times until the sample unit is surveyed for mr_1 time.

In a l-level rotation design, each sample unit reports the information for the current month as well as for the $l-1$ previous months. When the sample unit returns to the sample for every l^{th} month, it again provides l months information. One-level $r_1^m - r_2^{m-1}$ rotation system is generalized to "multi-level" rotation system: once a sample unit is selected from each rotation group, the sample unit returns to the sample for every l^{th} month until its r_1^{th} interview. Then no information is obtained from the sample unit for the next r_2 months; this procedure is repeated until this sample unit returns to the sample for its final mr_1^{th} interview. We denote this rotation system as $r_1^m(l) - r_2^{m-1}$ for the l-level design. Note that when $l = 1$, $r_1^m(l) - r_2^{m-1}$ rotation system is reduced back to the $r_1^m - r_2^{m-1}$ rotation system. In Table 2.1, we have $r_1 = 4$, $r_2 = 8$, $m = 2$ and $l = 3$. That is, the 3-level 4-8-4 design is written as $4^2(3) - 8^1$ design. We call a l-level rotation design to be the $r_1^m(l) - r_2^{m-1}$ design if the design follows the $r_1^m(l) - r_2^{m-1}$ system and if it has mr_1 rotation groups.

Each sample unit reports all the information simultaneously for the l months in the l-level rotation design. In other words, we have the information for the months $t - l_1$, $l_1 = 0, 1, \ldots, l-1$ where t is the month of a sample unit being surveyed. Since the information for the $t-1$ month, ..., $t-l+1$ month are also obtained from the t month unit, we define that this sample unit has the recall time l_1 at month $t - l_1$. Using this definition of recall time, we formally define the 3-way balanced multi-level rotation design as follows.

DEFINITION 2.1. *A l-level $r_1^m(l) - r_2^{m-1}$ design is balanced in 3-ways if the following three properties are satisfied:*

(a) *For each survey month, all mr_1 rotation groups are present in the sample. Each rotation group is represented by its l sample units, one with recall* *time 0, another with recall time 1, ..., and the last with recall time $l-1$ (see Table 2.1).*

(b) *For each survey month and recall time l_1, the sample is balanced in such a way that one of the mr_1 sample units is interviewed for the 1^{st} time, one for the 2^{nd} time, ..., and one for the mr_1^{th} time (see Table 2.1 (i), (ii) and (iii)).*

(c) *For every span of mr_1 months and for each recall time, each of the mr_1 rotation groups contributes its mr_1 sample units in which one sample unit is interviewed for the first time, ..., one for the mr_1^{th} time.*

We need (a) for the design to be balanced on recall time, (b) to be horizontally balanced on interview time, and (c) to be perpendicularly balanced on interview time.

The U. S. CEX uses $5^1(3)-0^0$ design which satisfies all properties of Definition 2.1. Since the U. S. MRTS and SIPP survey only one of their rotation groups in each survey month and use a sample unit repeatedly for the life of the survey, they do not satisfy all three properties of Definition 2.1. As seen in Table 2.1, the 3-level 4-8-4 rotation designs satisfy all the properties.

3. Necessary and Sufficient Condition for Three-Way Balancing

In this section, we investigate the relationship among r_1, r_2, m and l, the general rotation pattern, and the necessary and sufficient condition for the 3-way balanced multi-level rotation design. There would be an unlimited number of multi-level designs if no restriction is applied to the numbers, r_1, r_2, m and l. Although many multi-level designs lack the property of the 3-way balancing, some of them may be useful for various practical reasons (*e.g.*, Rao and Graham, 1964; Cantwell, 1990). The main point of this paper is not to eliminate such useful designs but to develop 3-way balanced semi multi-level rotation designs. We provide the necessary and sufficient condition to have the 3-way balanced designs and show how to construct such designs. By 3-way balancing, all rotation groups and units in the group are represented equally in the sample. We impose certain rules on the numbers, r_1, r_2, m and l for a $r_1^m(l)-r_2^{m-1}$ design to achieve 3-way balancing.

THEOREM 3.1. *Suppose that the $r_1^m(l)-r_2^{m-1}$ design is balanced in 3-ways.*

For each given i, $i = 0, 1, \ldots, r_1 - 1$, there is a unique integer m_i^*, $1 \le m_i^* \le mr_1$ satisfying

$$\text{mod}_{mr_1}\left\{m_i^* + i + (m_i^* - 1)(l-1) + \left[\frac{m_i^* - 1}{r_1}\right]r_2\right\} = 0, \tag{3.1}$$

where $[\cdot]$ is the integer operator.

The relationship among the numbers, r_1, r_2, m and l is formulated by the existence of the m_i^* satisfying (3.1) for all $i = 0, 1, 2, 3$. In $4^2(3) - 8^1$ design in Table 2.1, $m_0^* = 6$ for $i = 0$, $m_1^* = 3$ for $i = 1$, $m_2^* = 8$ for $i = 2$ and $m_3^* = 5$ for $i = 3$. However, $4^2(2) - 8^1$ design for $l = 2$ does not have m_i^* for all $i = 0, 1, 2, 3$ satisfying (3.1). Similarly, for $2^2(3) - 2^1$, $m_0^* = 2$ but m_1^* is not obtainable. Thus, $4^2(2) - 8^1$ and $2^2(3) - 2^1$ designs are not 3-way balanced by Theorem 3.1. In 2-way balanced design, r_2/r_1 should be nonnegatively integer-valued as shown in Park *et al.* (2001). Since the 3-way balanced design is a general version of the 2-way balanced design, hereafter we assume that

$$r_2 = k_1 r_1 \quad \text{for} \quad k_1 = 0, 1, \ldots \tag{3.2}$$

When $l = 1$ under (3.2), in particular, the necessary condition provided by Park *et al.* (2001) for the 2-way balanced one-level rotation design is a special case of Theorem 3.1 as shown below.

COROLLARY 3.1. *When $l = 1$ in (3.1), $m_i^* = m^* r_1 - i$, $0 \le i \le r_1 - 1$ where the integer m^* $(1 \le m^* \le m)$ satisfies $\text{mod}_m\{m^*(k_1 + 1) - k_1\} = 0$ for $k_1 = r_2/r_1$ where k_1 is a nonnegative integer.*

In $4^2 - 8^1$ design, the $m^* = 2$. Hence, we have $m_0^* = 8$, $m_1^* = 7$, $m_2^* = 6$ and $m_3^* = 5$ for $l = 1$.

With m_i^* given in Theorem 3.1, we define indices I_{ij} for the interview time to identify the same group in the sample months. Let $I_{ij} = \text{mod}_{mr_1}(m_i^* + jr_1)$ for $j = 0, 1, \ldots, m-1$, $i = 0, 1, \ldots, r_1 - 1$ where we replace $I_{ij} = mr_1$ if $I_{ij} = 0$. Then, it is shown in Appendix that the set $\{I_{ij}\}$ is equivalent to the set $\{1, 2, \ldots, mr_1\}$. Moreover, this index I_{ij} plays an important role to identify which two sample units come from the same rotation group and to show the perpendicular balancing on interview time for every span of mr_1 months (see Appendix for proof).

Since $m = 2$, $r_1 = 4$, $l = 3$ and the equation (3.1) give $m_0^* = 6$, $m_1^* = 3$, $m_2^* = 8$ and $m_3^* = 5$ in $4^2(3) - 8^1$ design, we have $I_{00} = 6$, $I_{01} = 2$, $I_{10} = 3$, $I_{11} = 7$, $I_{20} = 8$, $I_{21} = 4$, $I_{30} = 5$ and $I_{31} = 1$. Note from Table 2.1 (i) that the sample unit being interviewed for the $(I_{ij})^{th}$ time at month t and the sample unit interviewed for $(I_{i-1,j})^{th}$ time at month $t+1$ are from the same rotation group where $i = 1,2,3$ and $j = 0,1$. For $i = 0$ and $j = 0,1$, the sample unit with the $(I_{0j})^{th}$ interview time (6^{th} and 2^{nd}) at month t and the sample unit with the $(jr_1+1)^{th}$ interview time (1^{st} and 5^{th}) at month $t+1$ are from the same rotation group, 7 and 3, respectively. This means that in $4^2(3)-8^1$ design once we know interview times of two sample units during any two successive months, we can identify whether or not they come from the same rotation group through the indices I_{ij}'s. This is because $\{jr_1+1, j=0,1\} = \{1,5\} = \{I_{3j^*}, j^* = \text{mod}_2(1+j)\}$.

To generalize this observation in the $4^2(3)-8^1$ design, let $g_t^{l_1}(n)$ be the rotation group represented by the sample unit which has the recall time l_1, $0 \le l_1 \le l-1$ and has been interviewed for the n^{th} time ($n = 1,2,\ldots,mr_1$) in month t. Since I_{ij} can be used as the index for the interview time n, we present the following general rotation pattern for $r_1^m(l) - r_2^{m-1}$ design: for each recall time $l_1 = 0,1,\ldots,l-1$ and each $j = 0,1,\ldots,m-1$,

$$\begin{aligned} g_t^{l_1}(I_{0j}) &= g_{t+1}^{l_1}(jr_1+1), \\ g_t^{l_1}(I_{ij}) &= g_{t+1}^{l_1}(I_{i-1,j}) \quad \text{for} \quad 1 \le i \le r_1 - 1. \end{aligned} \tag{3.3}$$

One can easily check that the $4^2(3)-8^1$ design follows the rotation pattern (3.3). For example, when $l_1 = 2$, $g_t^2(I_{00} = 6) = g_{t+1}^2(I_{31} = 1)$, $g_t^2(I_{11} = 7) = g_{t+1}^2(I_{01} = 2)$ and $g_t^2(I_{31} = 1) = g_{t+1}^2(I_{21} = 4)$. When a sample unit at month t is interviewed for the $(I_{ij})^{th}$ time with recall time l_1, and another unit at month $t+1$ is interviewed for the $(I_{i-1,j})^{th}$ time with the same recall time l_1, these two units are from the same rotation group. As shown in Lemma A.1 given in Appendix, jr_1+1 can be written as I_{r_1-1,j^*} where $j^* = \text{mod}_m\{(mr_1 - m^*_{r_1-1}+1)/r_1 + j\}$. Thus, the rotation group containing the sample unit which is interviewed for $(I_{0j})^{th}$ time is the same rotation group containing the sample unit which is interviewed for the $(I_{r_1-1,j^*})^{th}$ time.

By Corollary 3.1, $I_{ij} = \text{mod}_{mr_1}(m^*r_1 - i + jr_1)$ when $l = 1$ in (3.3). Therefore, since $jr_1 + 1 = \text{mod}_{mr_1}\{mr_1 + jr_1 + 1\}$ for $j = 0,1,\ldots,m-1$, we have

$$g_t^0(i) = \begin{cases} g_{t+1}^0((m-m^*)r_1 + i + 1), & \text{for } i = r_1, 2r_1, \ldots, (m^*-1)r_1, \\ g_{t+1}^0(i - m^*r_1 + 1), & \text{for } i = m^*r_1, (m^*+1)r_1, \ldots, mr_1, \\ g_{t+1}^0(i+1), & \text{for } i = 1,2,\ldots,mr_1 \text{ except } r_1, r_2, \ldots, mr_1. \end{cases}$$

This rotation pattern, when $l = 1$, is the 2-way balanced one-level rotation pattern given in Park *et al.* (2001) and is a special case of (3.3).

More importantly, the general rotation pattern given in (3.3) is the necessary and sufficient condition for a multi-level rotation design to be balanced in 3-ways.

THEOREM 3.2. *Suppose that all $g_t^{l_1}(i)$ for $i = 1, 2, \ldots, mr_1$ and $l_1 = 0, 1, \ldots$ $l-1$ exist at an initial month t. Then the rotation pattern described in (3.3) is the necessary and sufficient condition for $r_1^m(l) - r_2^{m-1}$ design to be balanced in 3-ways where $r_2 = k_1 r_1$ for $k_1 = 0, 1, \ldots$*

The existence of unique integers m_i^*, for each $i = 0, 1, \ldots, r_1 - 1$, satisfying (3.1) given in Theorem 3.1 is the necessary condition for the 3-way balancing. Thus, by Theorem 3.2, the $r_1^m(l) - r_2^{m-1}$ design with such m_i^*'s belongs to the class of 3-way balanced multi-level designs when the design is constructed following the rotation pattern (3.3).

An algorithm is provided in the following section for the construction of the 3-way balanced design.

4. AN ALGORITHM FOR CONSTRUCTING THREE-WAY BALANCED $r_1^m(l) - r_2^{m-1}$ DESIGNS

With some appropriate numbers, r_1 of in-sample months, r_2 of out-sample months, m of repeating time and l of recall levels, we can construct the $r_1^m(l) - r_2^{m-1}$ design which is balanced in 3-ways. Under the existence of m_i^* satisfying $\text{mod}_{mr_1}\{m_i^* + i + (m_i^* - 1)(l-1) + [(m_i^* - 1)/r_1]r_2\} = 0$ for $i = 0, 1, \ldots, r_1 - 1$, the following algorithm provides the 3-way balanced $r_1^m(l) - r_2^{m-1}$ design with the rotation pattern (3.3).

The algorithm allocates only sample units with the recall time 0 at each survey month and this allocation automatically determines those sample units with more than 0 recall time (*i.e.*, $l_1 = 0, 1, \ldots, l-1$).

ALGORITHM. Using Table 4.1 with $m = 2$, $r_1 = r_2 = 3$ and $l = 4$,

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), $\ldots$, $(1, mr_1), \ldots, (\alpha^*, g^*)$, $(\alpha^*, g^* + 1), \ldots, (\alpha^*, mr_1)$ where $mr_1 = 6$, $\alpha^* > k_1 l$ and $k_1 = r_2/r_1$; $k_1 l = 4$ and $k_1 = 1$ in Table 4.1.

Step 2. Next fill in the first row of month t. According to the $r_1^m(l) - r_2^{m-1}$ rotation system, select every l^{th} index from Step 1 until r_1 indices are selected; then leave the next r_2 indices and repeat this procedure until the m^{th} r_1 successive indices are selected. The mr_1 sample units with the selected indices are the initial sample of the initial month t. For the $3^2(4) - 3^1$ design in Table 4.1, the first $r_1 = 3$ indices are $(1,1)$, $(1,5)$ and $(2,3)$, and the second 3 indices are $(3,4)$, $(4,2)$ and $(4,6)$ after skipping $r_2 = 3$ successive indices.

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

TABLE 4.1 *Three-way balanced* $3^2(4) - 3^1$ *design*

α	1						2						3						4						5				
g	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5
t	u_6	I	II	III	u_5	I	II	III	u_4	I	II	III				u_3	I	II	III	u_2	I	II	III	u_1	I	II	III		
m $t+1$		u_6	I	II	III	u_5	I	II	III	u_4	I	II	III				u_3	I	II	III	u_2	I	II	III	u_1	I	II	III	
o $t+2$			u_6	I	II	III	u_5	I	II	III	u_4	I	II	III				u_3	I	II	III	u_2	I	II	III	u_1	I	II	III
n $t+3$				u_6	I	II	III	u_5	I	II	III	u_4	I	II	II				u_3	I	II	III	u_2	I	II	III	u_1	I	II
t $t+4$					u_6	I	II	III	u_5	I	II	III	u_4	I	II	III				u_3	I	II	III	u_2	I	II	III	u_1	I
h $t+5$						u_6	I	II	III	u_5	I	II	III	u_4		II	III				u_3	I	II	III	u_2	I	II	III	u_1
$t+6$							u_6	I	II	III	u_5	I	II	III	u_4	I	II	III				u_3	I	II	III	u_2	I	II	III

Table 4.1 shows how to construct the 3-way balanced $3^2(4) - 3^1$ design. In this table, u_i indicates the sample unit with the recall time 0 which is interviewed for the i^{th} time at each given month. By Step 1, we construct the first two rows for the affiliation indices for the sample unit (α, g). Step 2 creates the initial row for the sample for survey month t. The sample unit $(4,6)$ is interviewed for the first time; the sample unit $(4,2)$ is interviewed for the second time; $\cdots$; and the sample unit $(1,1)$ is interviewed for the 6^{th} time. Step 3 generates the sample for month $t + t_0$, $t_0 \geq 1$ by shifting t_0 steps to the right from the initial survey month t. All 6 rotation groups are represented by their respective sample units in each month.

Once the allocation of the sample units with the recall time 0 at each survey month is determined, then the sample units report their information for the previous $l - 1$ months. Therefore, the allocation of the sample units with the recall time more than 0 is straightforward. For example, consider the months t, $t+1$ and $t+2$. At month t, all sample units indexed by $(\alpha, g) = (1,2)$, $(1,6)$, $(2,4)$, $(3,5)$, $(4,3)$ and $(5,1)$ have the same recall time $l_1 = 1$ since these sample units have the recall time $l_1 = 0$ at the survey month $t+1$. Similarly, the sample units indexed by (1,3), (2,1), (2,5), (3,6), (4,4) and (5,2) have recall time $l_1 = 2$ at month t.

In this $3^2(4) - 3^1$ design, since $r_1 = 3$, $r_2 = 3$, $m = 2$ and $l = 4$, we have $m_0^* = 6, m_1^* = 2$ and $m_2^* = 4$, and hence $I_{00} = 6$, $I_{01} = 3$, $I_{10} = 2$, $I_{11} = 5$, $I_{20} = 4$ and $I_{21} = 1$. Therefore, this design follows the rotation pattern (3.3): the group $1 = g_t^0(I_{00}) = g_{t+1}^0(I_{21})$ and $4 = g_t^0(I_{01}) = g_{t+1}^0(I_{20})$; the group $2 = g_t^0(I_{10}) = g_{t+1}^0(I_{00})$ and $5 = g_t^0(I_{11}) = g_{t+1}^0(I_{01})$; the group $3 = g_t^0(I_{20}) = g_{t+1}^0(I_{10})$ and $6 = g_t^0(I_{21}) = g_{t+1}^0(I_{11})$. Hence, the $3^2(4) - 3^1$ design is balanced in 3-ways by Theorem 3.2.

Similarly, for given r_1, r_2, m and l in $r_1^m(l) - r_2^{m-1}$ design, one can check the existence of m_i^* satisfying $\text{mod}_{mr_1}\{m_i^* + i + (m_i^* - 1)(l-1) + [(m_i^* - 1)/r_1]r_2\} = 0$ for $i = 0, 1, \ldots, r_1 - 1$. Whenever such m_i^*'s exists for all i, the algorithm given above always provides the 3-way balanced $r_1^m(l) - r_2^{m-1}$ design.

5. CONCLUSIONS

The multi-level design balanced in 3-ways is potentially usual tool to obtain the current information in relation to the previous information. This is one way to eliminate biases occurring from unbalanced interview times, rotation groups and recall times in the monthly sample.

We have demonstrated that the semi multi-level rotation design can be created to be balanced in 3-ways: balanced on recall time and on interview time both in monthly sample and rotation group. All the units in the group have equal opportunity to be represented in the sample. We have shown the algorithm to create the 3-way balanced multi-level rotation sample plan and also gives the necessary and sufficient condition for 3-way balanced design. The 2-way balanced designs such as the CPS, the Canadian and Australian Labor Force Surveys are considered as special cases of the 3-way balanced design.

We provide a method how to construct a 3-way balanced design and show the conditions for the balancing. However, this paper excludes the properties of estimation from the 3-way balanced design. The only reason of this exclusion is the length of the paper, and the remaining sections can be presented as a second part.

APPENDIX : PROOF OF THEOREMS 3.1, 3.2 AND COROLLARY 3.1

PROOF OF THEOREM 3.1. By definition of the 3-way balancing, there is one rotation group interviewed for the first time at any survey month. Thus there exists m_i^* such that

$$g_t^0(m_i^*) = g_{t+i+1}^0(1) \tag{A.1}$$

where $g_t^0(i)$ denotes the rotation group containing the sample unit with recall time 0 which is interviewed for the i^{th} time in survey month t.

We hereafter denote $g_t^0(i) = g_t(i)$ for simplicity of notation. Assume that there exist $m_{i_1}^*$ and $m_{i_2}^*$ such that $m_{i_1}^* \neq m_{i_2}^*$ but $g_t(m_{i_1}^*) = g_{t+i}(1)$ and $g_t(m_{i_2}^*) = g_{t+i}(1)$ for some i, $1 \leq i \leq r_1$ or such that $m_{i_1}^* = m_{i_2}^*$ but $g_t(m_{i_1}^*) = g_{t+i_1}(1) = g_{t+i_2}(1)$ for some $1 \leq i_1 \neq i_2 \leq r_1$. Then these violate (b) and (c) of Definition 2.1. Thus the rotation group satisfying (A.1) is unique.

The $r_1^m(l) - r_2^{m-1}$ system defined in Section 2 implies

$$\begin{aligned} g_t(m_i^*) &= g_t(k_{1i}r_1 + k_{2i}) = g_{t+l}(k_{1i}r_1 + k_{2i} + 1) \\ &= g_{t+2l}(k_{1i}r_1 + k_{2i} + 2) = \cdots = g_{t+(r_1-k_{2i})l}\big((k_{1i}+1)r_1\big). \end{aligned} \tag{A.2}$$

where $m_i^* = k_{1i}r_1 + k_{2i}$ with $0 \leq k_{1i} \leq m-1$ and $1 \leq k_{2i} \leq r_1$ without loss of generality. That is, this particular rotation group is to be interviewed for its $(k_{1i}+1)r_1$ time at survey month $t + (r_1 - k_{2i})l$. By the definition of $r_1^m(l) - r_2^{m-1}$ system and (A.1), we also have

$$\begin{aligned} g_t(m_i^*) &= g_{t+i+1}(1) = g_{t+i+1+l}(2) = \cdots = g_{t+i+1+(r_1-1)l}(r_1) \\ &= g_{t+i+1+(r_1-1)l+r_2+l}(r_1+1) = \cdots = g_{t+i+1+(r_1-1)l+r_2+r_1l}(2r_1). \end{aligned} \tag{A.3}$$

Repeat this procedure until we arrive at

$$\begin{aligned} &g_{t+i+1+2r_1l+2r_2}(2r_1+1) = \cdots = g_{t+i+1+2r_1l+2r_2+(r_1-1)l}(3r_1) \\ &= \cdots = g_{t+i+1+k_{1i}r_1l+k_{1i}r_2+(r_1-1)l}\big((k_{1i}+1)r_1\big). \end{aligned} \tag{A.4}$$

Equation (A.4) means that the rotation group $g_t(m_i^*)$ will be in sample at month $t + i + 1 + (k_{1i}+1)r_1l + k_{1i}r_2 - l$ for its $(k_{1i}+1)r_1^{th}$ interview. Therefore, by (A.2) and (A.4), the rotation group $g_t(m_i^*)$ is to be interviewed for the same time, $(k_{1i}+1)r_1$, in both months $t + (r_1 - k_{2i})l$ and $t + (k_{1i}+1)r_1l + k_{1i}r_2 - l + i + 1$. Hence, $\{t + (k_{1i}+1)r_1l + k_{1i}r_2 - l + i + 1\} - \{t + (r_1 - k_{2i})l\}$ should be a multiple of mr_1 since the same interview time happens only for every mr_1 months by (c) of Definition 2.1. This is precisely

$$\text{mod}_{mr_1}\{(k_{1i}r_1 + k_{2i})l + k_{1i}r_2 - l + i + 1\} = 0. \tag{A.5}$$

Finally, since $m_i^* = k_{1i}r_1 + k_{2i}$ and $[(m_i^* - 1)/r_1] = [(k_{1i}r_1 + k_{2i} - 1)/r_1] = k_{1i}$, (A.5) produces (3.1). □

PROOF OF COROLLARY 3.1. Equation (3.1) with $l = 1$ is,

$$\text{mod}_{mr_1}\{m_i^* + i + [(m_i^* - 1)/r_1]r_2\} = 0. \tag{A.6}$$

Since r_2 is a multiple of r_1, (A.6) implies that $m_i^* + i$ should be a multiple of r_1. Thus, we may let $m_i^* = m^*r_1 - i$ where $1 \leq m^* \leq m$. Then (A.6) can be written as $\text{mod}_{mr_1}\{m^*r_1 + [(m^*r_1 - i - 1)/r_1]r_2\} = 0$. Since $(m^* - 1)r_1 \leq m^*r_1 - i - 1 \leq m^*r_1 - 1$, we have $[(m^*r_1 - i - 1)/r_1] = m^* - 1$. Thus (A.6) is $\text{mod}_{mr_1}\{m^*r_1 + (m^* - 1)k_1r_1\} = \text{mod}_m\{m^* + (m^* - 1)k_1\} = 0$. □

The following three lemmas are necessary for proving Theorem 3.2.

LEMMA A.1. *Under (3.2),* $\{\{I_{0j}\}, \{I_{1j}\}, \ldots, \{I_{r_1-1,j}\}\}$ *is a partition of the set of* $\{1, 2, \ldots, mr_1\}$. *Specially,* $\{I_{r_1-1,j}\} = \{jr_1 + 1, j = 0, 1, \ldots, m - 1\}$.

LEMMA A.2. *Suppose that the* $r_1^m(l) - r_2^{m-1}$ *design is balanced in 3-ways with* $r_2 = k_1r_1$ *for* $k_1 = 0, 1, \ldots$ *When we have two sample units interviewed for* $(I_{ij})^{th}$ *time in month* t *and for* $jr_1 + 1$ *in month* $t + i + 1$ *for* $i = 0, 1, \ldots, r_1 - 1$ *and* $j = 0, 1, \ldots, m - 1$. *Then these two sample units are from the same rotation group.*

LEMMA A.3. *For each fixed* j, $j = 0, 1, \ldots, m - 1$, *let* $j_k^* = k(mr_1 - m_{r_1-1}^* + 1)/r_1 + j$. *Then,*

$$I_{0j} \neq I_{ij_1^*} \neq I_{ij_2^*} \neq \cdots \neq I_{ij_m^*} \quad \textit{for} \quad i = 0, 1, \ldots, r_1 - 1 \quad \textit{except} \quad I_{0j} = I_{0j_m^*}.$$

PROOF OF THEOREM 3.2. Suppose that the $r_1^m(l) - r_2^{m-1}$ design follows the rotation pattern given in (3.3). Then by Lemma A.1, the rotation pattern of (3.3) implies that all I_{ij}'s exist at month $t + 1$. Hence (a) and (b) given in Definition 2.1 are satisfied for month t+1 by the assumption and Lemma A.1. Furthermore, since this is true for any two successive months once the previous month satisfies (a) and (b) of Definition 2.1, we have these (a) and (b) for every survey month.

Since the rotation pattern given in (3.3) is the same for all recall times l_1, it is enough to consider the rotation pattern with the recall time 0 to show (c) of Definition 2.1. We pick an arbitrary rotation group $g_t(I_{0j})$ at month t where $g_t(\cdot) \equiv g_t^0(\cdot)$. Since $m_{r_1-1}^* = \tau r_1 + 1$ for some $\tau = 0, 1, \ldots, m - 1$ from Lemma A.1, it is easy to see that

$$jr_1 + 1 = I_{r_1-1,(mr_1-m_{r_1-1}^*+1)/r_1+j} \quad \text{for} \quad j = 0, 1, \ldots, m - 1. \tag{A.7}$$

Recursively applying the rotation pattern given in (3.3), we have, during the first r_1 months from t to $t + r_1 - 1$,

$$g_t(I_{0j}) = g_{t+1}(I_{r_1-1,j_1^*}) = g_{t+2}(I_{r_1-2,j_1^*}) = \cdots = g_{t+r_1-1}(I_{1j_1^*}), \qquad \text{(A.8)}$$

where we used (A.7) for the first equality with $j_1^* = (mr_1 - m_{r_1-1}^* + 1)/r_1 + j$. Similarly, by again using the rotation pattern, for $1 \le k \le m-1$, (A.8) is expressed until the month $t + mr_1 - 1$ by

$$\begin{aligned} g_{t+kr_1-1}(I_{1j_k^*}) &= g_{t+kr_1}(I_{0j_k^*}) = g_{t+kr_1+1}(I_{r_1-1,j_{k+1}^*}) = \cdots \\ &= g_{t+(k+1)r_1-1}(I_{1j_{k+1}^*}), \\ g_{t+mr_1-1}(I_{1j_m^*}) &= g_{t+mr_1}(I_{0j_m^*}), \end{aligned} \qquad \text{(A.9)}$$

where (A.7) is used for the second equality with $j_k^* = k(mr_1 - m_{r_1-1}^* + 1)/r_1 + j$.

By Lemmas A.1 and A.3, observe that rotation group $g_t(I_{0j})$ contains all interview times from 1 to mr_1 through its mr_1 sample units during mr_1 months from t to $t + mr_1 - 1$. Moreover, the rotation group is restarted for its I_{0j}^{th} interview at month $t + mr_1$ since $I_{0j} = I_{0j_m^*}$ in (A.9) which was shown in the proof of Lemma A.3. Therefore, the rotation group $g_t(I_{0j})$ satisfies property (c) given in Definition 2.1.

Also, we have $g_{t+1}(I_{r_1-1,j_1^*}) = g_{t+mr_1}(I_{0j}) = g_{t+mr_1+1}(I_{r_1-1,j_1^*})$ from (A.8) where we used $I_{0j} = I_{0j_m^*}$ in the second equality. Thus, by Lemma A.3, the rotation group $g_{t+1}(I_{r_1-1,j_1^*})$ satisfies property (c) given in Definition 2.1 for every span of mr_1 months starting from month $t + 1$. Similarly, the $mr_1 - 2$ rotation groups defined in (A.8) and (A.9) from $g_{t-2}(I_{r_1-2,j_1^*})$ to $g_{t+mr_1-1}(I_{1j_m^*})$ satisfy property (c) for every span of mr_1 months starting from their respective months from $t+2$ to $t + mr_1 - 1$. This shows property (c) given in Definition 2.1 since the rotation pattern described in (3.3) depends only on the interview times at each month.

Now, suppose that a $r_1^m(l) - r_2^{m-1}$ design is balanced in 3-ways. By Lemma A.2, then we have $g_{t_1}(I_{0\xi}) = g_{t_1+1}(\xi r_1 + 1)$ when $i = 0$. This is the general rotation pattern given in (3.3) for I_{0j}, $j = 0, 1, \ldots, m-1$.
Lemma A.2 also produces $g_{t_1+1}(I_{i\xi}) = g_{t_1+2+i}(\xi r_1 + 1)$ and $g_{t_1}(I_{i+1,\xi}) = g_{t_1+2+i}$ $(\xi r_1 + 1)$ for I_{ij}, $1 \le i \le r_1 - 1$. Hence we arrive at $g_{t_1}(I_{i+1,\xi}) = g_{t_1+1}(I_{i\xi})$.

Finally, since $g_t^{l_1}(I_{ij}) = g_{t+l_1}(I_{ij})$ for all l_1 and I_{ij} by the nature of the $r_1^m(l) - r_2^{m-1}$ rotation system, the rotation pattern of $g_t^{l_1}(\cdot)$ for all l_1 are the same as that of $g_t(\cdot) = g_t^0(\cdot)$. This completes the proof. □

REFERENCES

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

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.

FULLER, W. A. (1990). "Analysis of repeated survey", *Survey Methodology*, **16**, 167–180.

FULLER, W. A., ADAM, A. AND YANSANEH, I. S. (1992). "Estimators for longitudinal surveys with application to the U. S. Current Population Survey", *Proceedings of Statistics Canada Symposium 92*, 309–324.

JONES, R. G. (1980). "Best linear unbiased estimators for repeated surveys", *Journal of the Royal Statistical Society*, **B42**, 221–226.

LENT, J., MILLER, S. M., CANTWELL, P. J. AND DUFF, M. (1999). "Effects of composite weights on some estimates from the Current Population Survey", *Journal of Official Statistics*, **15**, 431–448.

MANSUR, K. A. AND SHOEMAKER, H. H. (1999). "The impact of changes in the Current Population Survey on time-in-sample bias and correlation between rotation groups", *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 180–185.

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 the 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.

WOODRUFF, R. S. (1963). "The use of rotating samples in the census bureau's monthly surveys", *Journal of the American Statistical Association*, **58**, 454–457.