A Study of the Composite Estimator for Change Rate

Dong-Yun Shin* Heungsun Park†

Abstract
Most panel surveys are performed repeatedly, so that the overlapped respondents could provide better estimates of the monthly/yearly change. But, as the repeated interviews with the same panel increase the burden, rotation sample design is used. Recent studies suggest AK-composite estimator for the level change, $Y_t - Y_{t-1}$, in rotated panel survey, which was first introduced by Kumar and Lee (1983). In this study, we introduce a composite estimator for estimating the change rate, $(Y_t - Y_{t-1})/Y_{t-1}$, not the change itself. We compare MSEs of composite estimators and those of simple ratio estimator with In-for-6 simulated data.

Key Words: panel data, overlapped samples, relative-error

1. Introduction

A panel survey is defined as a survey in which similar measurements are made on the same sample at different time points and it is often to have panel rotations in a certain way in order to decrease the burden of the respondents (Neuwenbroek, 1991). A rotation panel is referred as an alternative to solve this problem. In a rotation panel survey, the panel is supposed to be in a sample for a fixed period and come out of the sample, and a new panel is introduced instead. The overlapped portion is calculated as $p = 1 - 1/k$, where k is the staying period in the sample for a fixed panel. The rotation panel is a hybrid between a cross-sectional panel and a fixed panel. In general, in-for-6 and 4-8-4 rotation scheme are used.

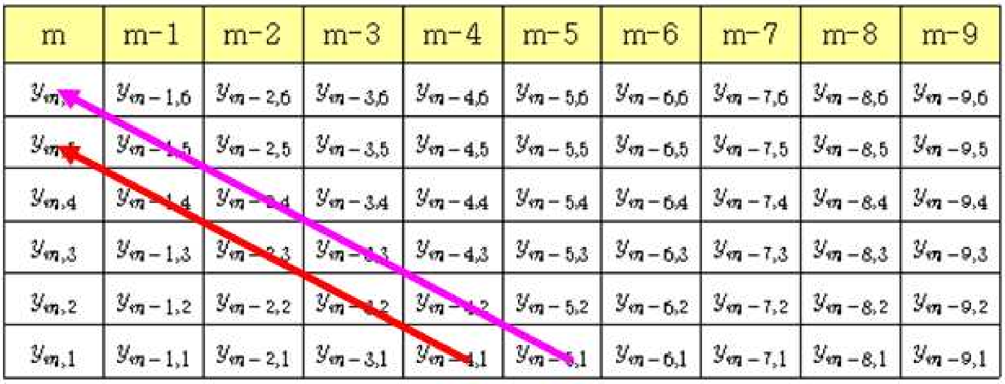
Besides Kummar and Lee (1983) introduced AK-composite estimator for in-for-6 rotation panels, several authors suggested other alternatives (Breau and Ernst, 1983; Gurney and Daly, 1965). In this article, we focuses on estimating the change rate based on in-for-6 rotation panel.

1.1 In-for-6 Rotation

Kummar and Lee(1983) suggested an estimating method for the change of Canadian Labor Force via composite estimators. As for in-for-6 rotation sampling, the units stays for 6 time points and leave out of the sample by being replaced by a new panel. The rotation scheme is illustrated at Figure 1.

From the bottom line in Figure 1, the interviews are repeated according to north-west direction, so that the panel remains in the sample up to 6 interviews and it will be replaced by new panel after that. Here, $y_{m,i}$ represents the simple ratio estimator for total sum at time t for the mth interviewed panel. Therefore, $y_{m-1,i-1}$, $y_{m-2,i-2}$, $\cdots$, $y_{m-i,1}$ are the repeated simple ratio estimates for the same panel.

*Dept. of Statistics, Hankuk University of Foreign Studies, Yongin, 449-791, South Korea
†Dept. of Statistics, Hankuk University of Foreign Studies, Yongin, 449-791, South Korea. email: hspark@hufs.ac.kr



m	m-1	m-2	m-3	m-4	m-5	m-6	m-7	m-8	m-9
$y_{m,6}$	$y_{m-1,6}$	$y_{m-2,6}$	$y_{m-3,6}$	$y_{m-4,6}$	$y_{m-5,6}$	$y_{m-6,6}$	$y_{m-7,6}$	$y_{m-8,6}$	$y_{m-9,6}$
$y_{m,5}$	$y_{m-1,5}$	$y_{m-2,5}$	$y_{m-3,5}$	$y_{m-4,5}$	$y_{m-5,5}$	$y_{m-6,5}$	$y_{m-7,5}$	$y_{m-8,5}$	$y_{m-9,5}$
$y_{m,4}$	$y_{m-1,4}$	$y_{m-2,4}$	$y_{m-3,4}$	$y_{m-4,4}$	$y_{m-5,4}$	$y_{m-6,4}$	$y_{m-7,4}$	$y_{m-8,4}$	$y_{m-9,4}$
$y_{m,3}$	$y_{m-1,3}$	$y_{m-2,3}$	$y_{m-3,3}$	$y_{m-4,3}$	$y_{m-5,3}$	$y_{m-6,3}$	$y_{m-7,3}$	$y_{m-8,3}$	$y_{m-9,3}$
$y_{m,2}$	$y_{m-1,2}$	$y_{m-2,2}$	$y_{m-3,2}$	$y_{m-4,2}$	$y_{m-5,2}$	$y_{m-6,2}$	$y_{m-7,2}$	$y_{m-8,2}$	$y_{m-9,2}$
$y_{m,1}$	$y_{m-1,1}$	$y_{m-2,1}$	$y_{m-3,1}$	$y_{m-4,1}$	$y_{m-5,1}$	$y_{m-6,1}$	$y_{m-7,1}$	$y_{m-8,1}$	$y_{m-9,1}$

Figure 1: In-for-6 Rotation Pattern

2. Composite Estimators

Since AK-composite estimator was introduced by Gurney and Daly (1965), Kumar and Lee (1983) developed it for estimating the change of levels in Canadian Labor Force Survey. The AK composite estimator for the mth month is defined as

$$y'_m = (1-K+A)y_{m,1}/6 + (1-K-A/5)\sum_{j=2}^{6} y_{m,j}/6 + K(y'_{m-1} + d_{m,m-1}), \quad (1)$$

where $y_{m,1}$ is the simple ratio estimator for the new panel at the m month, $d_{m,m-1}$ is the average of the difference of simple ratio estimates between m month and $m-1$ month. In this, 'A' puts more weights on the newly come panel group whereas 'K' puts more weights on the change from the previous month. If $A = 0$ implies K-composite estimator and $A = K = 0$ implies just the ordinary simple ratio estimator.

In survey sampling, people are interested not only in estimating the total sum such as the number of unemployment, participants in national health insurance, or aged persons above 65, but also in estimating the change of total sum such as the increased number of unemployment people or the change amount of household income since last year.

Kumar and Lee (1983) calculated the variance function of AK-composite estimator for change as follows:

$$\begin{aligned} V(y'_m - y'_{m-1}) &= \sigma^2 \left[A^2/30 - (1-\rho_1)KA/15 + (1-K)^2/6 \right. \\ &\quad \left. +(1-\rho_1)K(K+5)/15\right]/K - (1-K)^2 V(y'_m)/K \end{aligned} \quad (2)$$

where

$$\begin{aligned} V(y'_m) &= \left[V(y_m) + K^2 V(d_{m,m-1}) + 2KCov(y_m, d_{m,m-1}) \right. \\ &\quad \left. +2KCov(y_m, y'_{m-1}) + 2K^2 Cov(d_{m,m-1}, y'_{m-1})\right]/(1-K^2) \end{aligned} \quad (3)$$

In this, Kumar and Lee (1983) found the optimal K and A for estimating the change for the Canadian Labor Force Survey as $K = 0.9$ if A-composite is used and as $A = 0.1$ and $K = 0.9$ if AK-composite is used.

3. AK Composite Estimator for Change Rate

The change amount between the previous time point and the present time point is denoted as $Y_t - Y_{t-1}$, however sometimes the primary interest may lie on the change rate (the percentage form of change amount). For example, the survey is performed for estimating the increased household income percentage during the last one year. This concept mainly stems from the relative-error or percentile error studies(Park and Stefanski, 1998)

In this study, we suggest how to calculate the variance function of the AK-composite estimator for the relevant issue and the corresponding optimal A, K coefficients.

First of all, the variance function of the simple ratio estimator is obtained as follows:

$$\begin{aligned} V\left(\frac{\hat{Y}_2 - \hat{Y}_1}{\hat{Y}_1}\right) &= V\left(\frac{\hat{Y}_2}{\hat{Y}_1} - 1\right) = V\left(\frac{\hat{Y}_2}{\hat{Y}_1}\right) \\ &= V\left(\frac{N\bar{y}_2}{N\bar{y}_1}\right) = V\left(\frac{\bar{y}_2}{\bar{y}_1}\right) \\ &\approx \frac{1}{\mu_1^2}\left(\frac{S_2^2}{n_1}\right) + \frac{\mu_2^2}{\mu_1^4}\left(\frac{S_1^2}{n_1}\right) - 2\frac{\mu_2}{\mu_1^3}\left(\frac{n_{1r}}{n_1^2}\rho S_1 S_2\right) \end{aligned} \tag{4}$$

which is based on the first order Taylor's expansion.

Similarly, the variance function of AK-composite estimator for the change rate can be obtained as

$$\begin{aligned} V\left(\frac{y'_m - y'_{m-1}}{y'_{m-1}}\right) &= V\left(\frac{y'_m}{y'_{m-1}}\right) \\ &\approx \frac{1}{\mu_{m-1}^2}V\left(y'_m\right) + \frac{\mu_m^2}{\mu_{m-1}^4}V\left(y'_{m-1}\right) - 2\frac{\mu_m}{\mu_{m-1}^3}Cov\left(y'_m, y'_{m-1}\right) \\ &= \frac{1}{\mu_{m-1}^2}V\left(y'_m\right) + \frac{\mu_m^2}{\mu_{m-1}^4}V\left(y'_m\right) - 2\frac{\mu_m}{\mu_{m-1}^3}\frac{1}{2K}\Big[(1-K^2)V(y'_m) \\ &\quad -\left\{\frac{(1-K)^2}{6} + \frac{A^2}{30}\right\}\sigma^2 - 2K(1-K)(1-\rho_1)\frac{\sigma^2}{6} \\ &\quad -A(1-\rho_1)\frac{\sigma^2}{30} - K^2\frac{2}{5}(1-\rho_1)\sigma^2\Big]. \end{aligned} \tag{5}$$

Likewise, the MSE's for the average of simple ratio estimators and AK-composite estimator are

$$\begin{aligned} MSE\left(\frac{\bar{y}_m - \bar{y}_{m-1}}{\bar{y}_{m-1}}\right) &= V\left(\frac{\bar{y}_m - \bar{y}_{m-1}}{\bar{y}_{m-1}}\right) + \left\{E\left(\frac{\bar{y}_m - \bar{y}_{m-1}}{\bar{y}_{m-1}}\right) - \left(\frac{Y_m - Y_{m-1}}{Y_{m-1}}\right)\right\}^2 \\ &\approx \frac{1}{E(\bar{y}_m)^2}V(\bar{y}_m) + \frac{E(\bar{y}_m)^2}{E(\bar{y}_{m-1})^4}V(\bar{y}_{m-1}) \\ &\quad -2\frac{E(\bar{y}_m)}{E(\bar{y}_{m-1})^3}Cov(\bar{y}_m, \bar{y}_{m-1}) \end{aligned} \tag{6}$$

and

$$MSE\left(\frac{y'_m - y'_{m-1}}{y'_{m-1}}\right) = V\left(\frac{y'_m - y'_{m-1}}{y'_{m-1}}\right) + \left\{E\left(\frac{y'_m - y'_{m-1}}{y'_{m-1}}\right) - \left(\frac{Y_m - Y_{m-1}}{Y_{m-1}}\right)\right\}^2$$

m	m-1	m-2	m-3	m-4	m-5	m-6	m-7	m-8	m-9
25.34	25.1	26.03	21.77	26.38	20.77	17.44	23.06	19.51	24.09
21.08	26.75	24	26.07	22.31	25.19	19.84	19.63	21.3	23.51
21.04	22.27	24.63	24.34	25.43	23.61	19.89	19.69	18.89	24.51
20.16	21.47	17.1	25.31	20.99	20.78	22.03	21.61	18.48	20.19
26.85	20.39	21.12	15.82	25.43	21.03	20.67	17.24	21.38	19.51
17.87	25.16	20.88	21.3	17.02	20.88	20.28	20.37	15.89	19.97

Figure 2: Simple Ratio Estimates for each month.

which means

$$MSE\left(\frac{y'_m - y'_{m-1}}{y'_{m-1}}\right) \approx \frac{1}{\mu^2_{m-1}}V(y'_m) + \frac{\mu^2_m}{\mu^4_{m-1}}V(y'_{m-1}) \\ -2\frac{\mu_m}{\mu^3_{m-1}}Cov(y'_m, y'_{m-1}) + \left(\frac{\mu_m}{\mu_{m-1}} - \frac{Y_m}{Y_{m-1}}\right)^2. \quad (7)$$

4. Simulated Data

The performance of estimators are validated by comparing their MSEs, and for the In-for-6 rotation panel data we would like to show how to decide the optimal A, K coefficients and compare it with MSE of the average of simple ratio estimators.

In this study, we generate 6 simple ratio estimators for a fixed month and each ratio estimator are made positively correlated within AR(1) process, which implies given time point t,

$$y_{m-9+t,j} = (1-\rho)\mu + \rho y_{m-9+(t-1),j-1} + \epsilon_t \quad (8)$$

where

$$\begin{aligned} y_{m,1} &\sim N(20,5) \\ \rho &= 0.8 \\ \mu &= 20 \\ \epsilon_t &\sim N(0, 5(1-0.64)). \end{aligned}$$

The generated estimates are shown at Figure 2. And, Table 1 illustrates MSEs for each A-K combinations.

Therefore, the optimal K-composite estimator is obtained by $K = 0.9$ and the optimal AK-composite estimator is obtained by $A = 0.2$ and $K = 0.7$ for this data.

5. Conclusion

Rotation panel is necessary to reduce the interviewee's burden and eliminate the group bias effect caused by the successive interviews. Kumar and Lee (1983) introduced AK-composite estimator for the change of monthly labor force. In this study,

Table 1: MSE for AK-composite estimates of change rate for m month

	K								
A	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.0	0.0035	0.0041	0.0047	0.0034	0.0035	0.0030	0.0024	0.0029	0.0021
0.1	0.0032	0.0031	0.0040	0.0034	0.0034	0.0030	0.0029	0.0030	0.0023
0.2	0.0042	0.0032	0.0037	0.0039	0.0036	0.0035	0.0020	0.0027	0.0026
0.3	0.0050	0.0037	0.0030	0.0033	0.0031	0.0029	0.0028	0.0043	0.0020
0.4	0.0053	0.0074	0.0051	0.0050	0.0025	0.0029	0.0033	0.0026	0.0029
0.5	0.0078	0.0049	0.0030	0.0030	0.0035	0.0030	0.0032	0.0037	0.0023
0.6	0.0061	0.0044	0.0031	0.0029	0.0029	0.0022	0.0039	0.0042	0.0026
0.7	0.0069	0.0044	0.0041	0.0038	0.0030	0.0032	0.0029	0.0048	0.0026
0.8	0.0072	0.0071	0.0030	0.0051	0.0033	0.0049	0.0026	0.0034	0.0041
0.9	0.0080	0.0053	0.0032	0.0044	0.0047	0.0029	0.0028	0.0047	0.0028

we suggest AK-composite estimator for the change rate and its variance function with its corresponding optimal coefficients. The scope of this study is limited because only In-for-6 rotation pattern is considered with simulated data. However, it wouldn't be difficult to use this procedure for the real data and expand to other rotation pattern.

REFERENCES

Breau, P., and Ernst, L. R. (1983), "Alternative Estimators to the Current Composite Estimator", *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 397–402.

Gurney, M. and Daly, J. F. (1965), "A Multivariate Approach to Estimation in Periodic Sample Surveys", *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 237–257.

Kumar, S., and Lee, H. (1983), "Evaluation of Composite Estimation for the Canadian Labor Force Survey," *Survey Methodology*, 9, 403–408.

Nieuwenbroek, N. J.(1991), "Precision of Net Change in a Rotating Panel Survey," *The Statistician*, 40, 195-201.

Park, H. and Stefanski, L. A. (1998), "Relative Error Prediction" *Statistics & Probability Letters*, 40, 227–236.