RATIO ESTIMATORS IN POST-STRATIFICATION

Sevil BACANLI*
Eda EVİN AKSU**

In this study, some ratio-type estimators are taken into consideration in literature and their properties are studied in post-stratification. Mean square error (MSE) of all ratio estimators in post-stratification is obtained and compared with the MSE of classical estimators in stratified random sampling. Within the frame work of the data from 2000 General Population size Census of Turkey which was carried out by the Turkish Statistical Institute (TURKSTAT), average employment has been estimated, and the population is taken as auxiliary variable by NUTS-1 Level. An application is carried out to show the superiority of the suggested ratio-estimators in post-stratification under the guidance of Turkey 2000 Population Census data.

1. INTRODUCTION

A ratio estimate of the population mean $\bar{Y}$ can be made in stratified random sampling in two ways. One way is to make a separate ratio estimate of the total of each stratum and weighting these totals. The other one is combined ratio estimate which is derived from a single combined ratio. From the sample data, we compute sample mean of the variates in stratified random sampling ($\bar{y}_s$) method are computed as such

$$\bar{y}_s = \sum_{=1}^{\ell} W_h \bar{y}_h \qquad \bar{x}_s = \sum_{h=1}^{\ell} W_h \bar{x}_h \tag{1}$$

where ℓ is the number of stratum, $W_h = N_h/N$ is stratum weight, N is the number of units in population, N_h is the number of units in stratum h, $\bar{y}_h$ is the sample mean of variate of interest in stratum h and $\bar{x}_h$ is the sample mean of auxiliary variate in stratum h. The separate ratio estimate ($\bar{y}_{sr}$) and the MSE of this estimator are given by

$$\bar{y}_{sr} = \sum_{h=1}^{\ell} W_h \frac{\bar{y}_h}{\bar{x}_h} \bar{X}_h \ , \tag{2}$$

$$MSE(\bar{y}_{sr}) = \sum_{h=1}^{l} W_h^2 \left(\frac{1-f_h}{n_h}\right)\left(S_{yh}^2 + S_{xh}^2 R_h^2 - 2R_h S_{xyh}\right)$$

$$= \sum_{h=1}^{l} W_h^2 \left(\frac{1-f_h}{n_h}\right) \bar{Y}_h^2 \left(C_{hy}^2 + C_{hx}^2 - \rho_{hxy} C_{hy} C_{hx}\right) \tag{3}$$

respectively. Here $\bar{y}_h$ is the sample mean of the study variable in stratum h; $\bar{x}_h$ is the sample mean of the auxiliary variable in stratum h; $f_h = n_h/N_h$; n_h is the sample size of the h. stratum and N_h is the

---

* Associated Professor, Hacettepe University Department of Statistics
** Expert, TURKSTAT, Department of Demography Statistics

population size of the h. stratum. In the equation (3), $S_{yh}^2 = \frac{\sum_{i=1}^{N_h}\left(y_{h_i} - \bar{Y}_h\right)^2}{N_h - 1}$ and $S_{xh}^2 = \frac{\sum_{i=1}^{N_h}\left(x_{h_i} - \bar{X}_h\right)^2}{N_h - 1}$ are population variances of auxiliary and study variables in stratum h, $S_{xyh}^2 = \frac{\sum_{i=1}^{N_h}\left(y_{h_i} - \bar{Y}_h\right)\left(x_{h_i} - X_h\right)}{N_h - 1}$ is the population covariance between the auxiliary and study variables, $R_h = \bar{Y}_h / \bar{X}_h$ is the population ratio of h. stratum ; $\rho_{hxy} = \frac{S_{hxy}}{S_{hx} S_{hy}}$ is the coefficient of correlation between x and y; $C_{hy}^2 = \frac{S_{hy}^2}{\bar{Y}_h^2}$ and $C_{hx}^2 = \frac{S_{hx}^2}{\bar{X}_h^2}$ are the coefficient of variation of y and x in stratum h. It should not be forgotten in here, it is assumed that the population mean $\bar{X}$ of the auxiliary variable x is known (Cochran, 1977; Singh, 2003).

When the population coefficient of variation C_x and kurtosis $\beta_2(x)$ of the auxiliary variable, are known, Sisodia and Dwivedi (1981) ($\bar{y}_{SD}$), Upadhyaya and Singh (1999) suggest ratio-type estimators ($\bar{y}_{SK}$, $\bar{y}_{US1}$, $\bar{y}_{US2}$) for $\bar{Y}$ in simple random sampling as

$$\bar{y}_{SD} = \bar{y}\frac{\bar{X} + C_x}{\bar{x} + C_x} = \frac{\bar{y}}{\bar{x}_{SD}}\bar{X}_{SD} \tag{4}$$

$$\bar{y}_{SK} = \bar{y}\frac{\bar{X} + \beta_2(x)}{\bar{x} + \beta_2(x)} = \frac{\bar{y}}{\bar{x}_{SK}}\bar{X}_{SK} \tag{5}$$

$$\bar{y}_{US1} = \bar{y}\frac{\bar{X}\,\beta_2(x) + C_x}{\bar{x}\beta_2(x) + C_x} = \frac{\bar{y}}{\bar{x}_{US1}}\bar{X}_{US1} \tag{6}$$

$$\bar{y}_{US2} = \bar{y}\frac{\bar{X}\,C_x + \beta_2(x)}{\bar{x}C_x + \beta_2(x)} = \frac{\bar{y}}{\bar{x}_{US2}}\bar{X}_{US2} \quad . \tag{7}$$

The MSE equation of these estimators are given by

$$MSE(\bar{y}_{SD}) = \frac{(1-f)}{n}\bar{Y}^2\left[C_y^2 + C_x^2\alpha^2 - 2\alpha\rho_{xy}C_yC_x\right] \tag{8}$$

$$MSE(\bar{y}_{SK}) = \frac{(1-f)}{n}\bar{Y}^2\left[C_y^2 + C_x^2\delta^2 - 2\delta\rho_{xy}C_yC_x\right] \tag{9}$$

$$MSE(\bar{y}_{US1}) = \frac{(1-f)}{n}\bar{Y}^2\left[C_y^2 + C_x^2\pi^2 - 2\pi\rho_{xy}C_yC_x\right] \tag{10}$$

$$MSE(\bar{y}_{US2}) = \frac{(1-f)}{n}\bar{Y}^2\left[C_y^2 + C_x^2\theta^2 - 2\theta\rho_{xy}C_yC_x\right] \tag{11}$$

where $\alpha = \bar{X}/(\bar{X} + C_x)$, $\delta = \bar{X}/(\bar{X} + \beta_2(x))$, $\pi = \bar{X}\beta_2(x)/(\bar{X}\beta_2(x) + C_x)$,

$\theta = \bar{X}C_x/(\bar{X}C_x + \beta_2(x))$.

These estimators are developed by assuming that the sample is selected from the population with equal probability under simple random sampling. Kadilar and Cingi (2003) analyze these estimators for combined ratio estimator in the stratified random sampling and Rueda et al. (2006), Bacanli and Kadilar (2008) examine them under the light of various sampling designs. However, none of these studies examined those estimators within the frame of the post-stratified sampling. Therefore, this paper aims to investigate these ratio-type estimators, by using post-stratified sampling.

Post-stratification is a method of estimation that is very popular among survey practitioners. Holt and Smith (1979) describe the post-stratification as follows: Firstly, the sample is selected and then this selected sample is divided into groups such as age, gender, region, occupation and other factors. The reason behind the usage of post-stratification after sampling is that the information for the classification of the sampling units cannot be achieved prior data collection or is very high priced to use when creating sampling strata (Cervantes and Brick, 2009).

Two important problems can occur within the process of application post-stratification. The first of these problems is the empty strata. The estimations regarding the population cannot be calculated in the case of empty strata. This problem can be solved by combining the strata. However, this is a difficult and time-consuming process in the surveys. The other second problem in stratification is that the strata size in population cannot be known before (Bethlehem and Keller, 1987).

A number of articles have been written about the usages and benefits of post-stratification. For instance, Zhang (2000) examined a calibration estimator in post-stratification. Liu (2002) applied a three-stage sampling procedure to the estimation of mean in post-stratification. Kim et al. (2007) suggested findings from their research which is specific to cell collapsing in post-stratification. Kim and Wang (2009) proposed a simple second-order linearization variance estimator for the post-stratified estimator of the population total in two-stage sampling. Martinez et al. (2011) proposed a post-stratified calibration estimator for estimating quantiles.

In this study, firstly ratio estimators are considered within the frame work of separate ratio estimate *(*$\bar{y}_{srSD}$, $\bar{y}_{srSK}$, $\bar{y}_{srUS1}$,$\bar{y}_{srUS2}$*)* in the stratified random sampling. These estimators are given below:

$$\bar{y}_{srSD} = \sum_{h=1}^{l} \bar{y}_h W_h \frac{(\bar{X}_h + C_{hx})}{(\bar{x}_h + C_{hx})} \quad (12)$$

$$\bar{y}_{srSK} = \sum_{h=1}^{l} \bar{y}_h W_h \frac{(\bar{X}_h + \beta_{2h}(x))}{(\bar{x}_h + \beta_{2h}(x))} \quad (13)$$

$$\bar{y}_{srUS1} = \sum_{h=1}^{l} \bar{y}_h W_h \frac{(\bar{X}_h \beta_{2h}(x) + C_{hx})}{(\bar{x}_h \beta_{2h}(x) + C_{hx})} \quad (14)$$

$$\bar{y}_{srUS2} = \sum_{h=1}^{l} \bar{y}_h W_h \frac{(\bar{X}_h C_{hx} + \beta_{2h}(x))}{(\bar{x}_h C_{hx} + \beta_{2h}(x))} \quad (15)$$

MSE equations of these estimators are given by:

$$MSE(\bar{y}_{srSD}) = \sum_{h=1}^{l} W_h^2 \frac{(1-f_h)}{n_h} \bar{Y}_h^2 \left[C_{hy}^2 + C_{hx}^2 \alpha_h^2 - 2\alpha_h \rho_{hxy} C_{hy} C_{hx}\right] \quad (16)$$

$$MSE(\bar{y}_{srSK}) = \sum_{h=1}^{l} W_h^2 \frac{(1-f_h)}{n_h} \bar{Y}_h^2 \left[C_{hy}^2 + C_{hx}^2 \delta_h^2 - 2\delta_h \rho_{hxy} C_{hy} C_{hx}\right] \quad (17)$$

$$MSE(\bar{y}_{srUS1}) = \sum_{h=1}^{l} W_h^2 \frac{(1-f_h)}{n_h} \bar{Y}_h^2 \left[C_{hy}^2 + C_{hx}^2 \pi_h^2 - 2\pi_h \rho_{hxy} C_{hy} C_{hx}\right] \quad (18)$$

$$MSE(\bar{y}_{srUS2}) = \sum_{h=1}^{l} W_h^2 \frac{(1-f_h)}{n_h} \bar{Y}_h^2 \left[C_{hy}^2 + C_{hx}^2 \theta_h^2 - 2\theta_h \rho_{hxy} C_{hy} C_{hx}\right] \quad (19)$$

Where $\alpha_h = \bar{X}_h/(\bar{X}_h + C_{hx})$, $\delta_h = \bar{X}_h/(\bar{X}_h + \beta_{2h}(x))$,

$\pi_h = \bar{X}_h \beta_{2h}(x)/(\bar{X}_h \beta_{2h}(x) + C_{hx})$ *and* $\theta_h = \bar{X}_h C_{hx}/(\bar{X}_h C_{hx} + \beta_{2h}(x))$.

2. THE SUGGESTED ESTIMATORS

Post-stratification is often used in sample surveys, when the identification of stratum cannot be achieved in advance. In a post-stratified sampling scheme a sample of n units is first selected from the population of N units by using simple random sampling. The population is stratified into L strata on the basis of some known auxiliary information. In post stratified sampling, the values of N_h, where $h=1,2,\ldots,L$ and $\boldsymbol{N = \sum_{h=1}^{L} N_h}$ may or may not be known for each sample unit which is selected with the chosen design. Then post-stratified or placed in the h^{th} stratum based on the auxiliary information associated with each sampled unit such as $\boldsymbol{n = \sum_{h=1}^{L} n_h}$. Thus the difference between stratified and post-stratified sampling schemes is that in stratified sampling the sub-sample size n_h is a fixed or predefined number in stratified sampling, whereas it is random variable in post-stratified sampling (Singh, 2003).

For the post-stratified sampling, the ratio estimator ($\bar{y}_{pr}$) can be written as

$$\bar{y}_{pr} = \sum_{h=1}^{L} W_h \frac{\bar{y}_h}{\bar{x}_h} \bar{X}_h, \quad (20)$$

This estimation is called as the separate ratio estimation, when the strata are identified before the sampling process in stratified sampling. The stratum totals must be derived from the frame or from a reliable external source (Särndal et al., 2003).

In post-stratification, estimator n_h are random variables. If n_h were fixed, post-stratified ratio estimator would function as separate ratio estimation in the stratified sampling under proportional allocation. If n_h were fixed, the MSE of the separate ratio estimator ($\bar{y}_{sr}$) is

$$MSE(\bar{y}_{sr}) = \sum_{h=1}^{l} W_h^2 \left(\frac{1-f_h}{n_h}\right) \bar{Y}_h^2 \left(C_{hy}^2 + C_{hx}^2 - \rho_{hxy} C_{hy} C_{hx}\right) \quad . \quad (21)$$

It can be defined that $H_{sr} = \bar{Y}_h^2 \left(C_{hy}^2 + C_{hx}^2 - \rho_{hxy} C_{hy} C_{hx}\right)$. Then, the equation will be

$$MSE(\bar{y}_{sr}) = \sum_{h=1}^{L} W_h^2 \left(\frac{N_h - n_h}{N_h n_h}\right) H_{pr}$$

$$= \sum_{h=1}^{L} W_h^2 H_{pr} \frac{1}{n_h} - \frac{1}{N} \sum_{h=1}^{L} W_h H_{pr} \quad . \quad (22)$$

In this situation, a general expression for $MSE(\bar{y}_{pr})$ can be approximated by replacing $1/n_h$ with its expected value.

It is difficult to find the expected value of the reciprocal of a random variable; a good approximation can be given as (Hansen et al., 1953),

$$E\left(\frac{1}{n_h}\right) \cong \frac{1}{nW_h}\left\{1+\frac{(1-W_h)}{nW_h}\right\}$$

$$\cong \frac{1}{nW_h}+\frac{(1-W_h)}{n^2W_h^2} \quad . \tag{23}$$

By replacing this with $1/n_h$ in equation for MSE of the separate ratio estimator, MSE of the post-stratified ratio estimator would be

$$MSE(\bar{y}_{pr}) = \sum_{h=1}^{L} W_h^2 H_{pr}\, E(\frac{1}{n_h}) - \frac{1}{N}\sum_{h=1}^{L} W_h\, H_{pr}$$

$$= \sum_{h=1}^{L} W_h^2\, H_{pr}\left(\frac{1}{nW_h}+\frac{(1-W_h)}{n^2W_h^2}\right) - \frac{1}{N}\sum_{h=1}^{L} W_h H_{pr}$$

$$= \frac{1}{n}\sum_{h=1}^{L} W_h H_{pr} + \frac{1}{n^2}\sum_{h=1}^{L} H_{pr}(1-W_h) - \frac{1}{N}\sum_{h=1}^{L} W_h H_{pr}$$

$$= \frac{1-f}{n}\sum_{h=1}^{L} W_h H_{pr} + \frac{1}{n^2}\sum_{h=1}^{L} H_{pr}(1-W_h) \ . \tag{24}$$

Ratio estimators in post-stratification are same as the separate ratio estimators in stratified sampling. But MSE equations differ in these methods.

The estimators given in the Section-1 are combined with post-stratified ratio estimator given in (20), following estimators $(\bar{y}_{pSD},\ \bar{y}_{pSK}, \bar{y}_{pUS1},\ \bar{y}_{pUS2})$ are proposed as such:

$$\bar{y}_{pSD} = \sum_{h=1}^{L} \bar{y}_h W_h \frac{(\bar{X}_h + C_{hx})}{(\bar{x}_h + C_{hx})} \tag{25}$$

$$\bar{y}_{pSK} = \sum_{h=1}^{L} \bar{y}_h W_h \frac{(\bar{X}_h + \beta_{2h}(x))}{(\bar{x}_h + \beta_{2h}(x))} \tag{26}$$

$$\bar{y}_{pUS1} = \sum_{h=1}^{L} \bar{y}_h W_h \frac{(\bar{X}_h \beta_{2h}(x) + C_{hx})}{(\bar{x}_h \beta_{2h}(x) + C_{hx})} \tag{27}$$

$$\bar{y}_{pUS2} = \sum_{h=1}^{L} \bar{y}_h W_h \frac{(\bar{X}_h C_{hx} + \beta_{2h}(x))}{(\bar{x}_h C_{hx} + \beta_{2h}(x))} \quad . \tag{28}$$

By using (24), the MSE of the proposed estimators can be given as

$$MSE(\bar{y}_{pSD}) = \frac{1-f}{n}\sum_{h=1}^{L} W_h H_{pSD} + \frac{1}{n^2}\sum_{h=1}^{L}(1-W_h)\, H_{pSD} \tag{29}$$

where $H_{pSD} = \bar{Y}_h^2\left(C_{yh}^2 + C_{xh}^2\alpha_h^2 - 2\rho_{xyh}C_xC_y\alpha_h\right)$, $\alpha_h = \frac{\bar{X}_h}{\bar{X}_h + C_{hx}}$;

$$MSE\left(\bar{y}_{pSK}\right) = \frac{1-f}{n}\sum_{h=1}^{L} W_h H_{pSK} + \frac{1}{n^2}\sum_{h=1}^{L}(1-W_h)\, H_{pSK} \quad (30)$$

where $H_{pSK} = \bar{Y}_h^2\left(C_{yh}^2 + C_{xh}^2\delta_h^2 - 2\rho_{xyh}C_xC_y\delta_h\right)$, $\delta_h = \frac{\bar{X}_h}{\bar{X}_h + \beta_{2h}(x)}$;

$$MSE\left(\bar{y}_{pUS1}\right) = \frac{1-f}{n}\sum_{h=1}^{L} W_h H_{pUS1} + \frac{1}{n^2}\sum_{h=1}^{L}(1-W_h)\, H_{pUS1} \quad (31)$$

where $H_{pUS1} = \bar{Y}_h^2\left(C_{yh}^2 + C_{xh}^2\pi_h^2 - 2\rho_{xyh}C_xC_y\pi_h\right)$, $\pi_h = \frac{\bar{X}_h\beta_{2h}(x)}{\bar{X}_h\beta_{2h}(x) + C_{hx}}$;

$$MSE\left(\bar{y}_{pUS2}\right) = \frac{1-f}{n}\sum_{h=1}^{L} W_h H_{pUS1} + \frac{1}{n^2}\sum_{h=1}^{L}(1-W_h)\, H_{pUS1} \quad (32)$$

where $H_{pUS2} = \bar{Y}_h^2\left(C_{yh}^2 + C_{xh}^2\theta_h^2 - 2\rho_{xyh}C_xC_y\theta_h\right)$, $\theta_h = \frac{\bar{X}_hC_{hx}}{\bar{X}_hC_{hx} + \beta_{2h}(x)}$.

3. APPLICATION AND MAIN RESULTS

As the numerical example, the data from 2000 Population and Housing Census of Turkey which was carried out by the Turkish Statistical Institute is used, average employment is estimated as: y: employment (study variable), x: population (auxiliary variable) for each of the 12 regions within the Nomenclature of Territorial Units for Statistics (NUTS-1) level. For this data set, each regions are considered as a population *(TR_1, TR_2, TR_3 , TR_4 , TR_5 , TR_6 , TR_7 , TR_8 , TR_9 , TR_A , TR_B , TR_C).*

Turkey is divided into 12 regions on the basis of the Nomenclature of Territorial Units for Statistics level. Regions and cities in the each region are given in the Table 1.

In the study, population in employment defined as the persons who take place in an economic activity at least one hour on the reference date either as a regular or casual employee or as unpaid family worker for an income either in kind (good) or in cash (money) and who is 12 years of age or over (TURKSTAT, 2003). The persons with unknown employment status are not covered in the study.

Consequently, persons who are at the age of 12 or older in the population of cities and villages are taken into consideration in accordance with the definition of employment. City population can be defined as the population of municipal areas of the province and district centers, while village population can be defined as the population of sub-districts and villages.

Administrative units are taken as a sampling unit, whereas the distinction between city and village area is taken as a stratification variable. In order to estimate the average employment of the regions in NUTS-1 level, it is determined that the margin of error d=200 and risk α= 0,05.

In stratified random sampling (SS), population is stratified into strata and then samples are selected from each stratum by using Neyman allocation. On the other hand, in post-stratified sampling (PS), firstly, a sample of n units is selected with using simple random sampling without replacement and then, the sample size is stratified in to strata. The sample size and the statistics of regions are given in Table-2 for each sampling scheme. However, since the sample size which is calculated for TR-1 region is equal to the population size, the MSE of the ratio estimators for this region cannot be calculated, so this region cannot be included to the application.

Table 1. Cities and regions in NUTS-1 Level

NUTS-1	Cities	NUTS-1	Cities
TR1 (İstanbul)	İstanbul	TR7 (Central-Anatolia)	Kırıkkale, Aksaray, Niğde, Nevşehir, Kırşehir, Kayseri, Sivas, Yozgat
TR2 (Western-Marmara)	Tekirdağ, Edirne, Kırklareli, Balıkesir, Çanakkale	TR8 (Western-Blacksea)	Zonguldak, Karabük, Bartın, Kastamonu, Çankırı, Sinop, Samsun, Çorum, Tokat, Amasya
TR3 (Aegean)	İzmir, Aydın, Denizli, Muğla, Manisa, Afyon, Kütahya, Uşak	TR9 (Eastern Blacksea)	Trabzon, Ordu, Giresun, Rize, Artvin, Gümüşhane
TR4 (Eastern-Marmara)	Bursa, Eskişehir, Bilecik, Kocaeli, Sakarya, Düzce, Bolu, Yalova	TRA (Northeast-Anatolia)	Erzurum, Erzincan, Bayburt, Ağrı, Kars, Ardahan
TR5 (Western-Anatolia)	Ankara, Konya, Karaman	TRB (Eastern-Anatolia)	Malatya, Elazığ, Bingöl, Tunceli, Van, Muş, Bitlis, Hakkari
TR6 (Mediterranean)	Antalya, Isparta, Burdur, Adana, Mersin, Hatay, Kahramanmaraş, Osmaniye	TRC (Southeastern-Anatolia)	Gaziantep, Adıyaman, Kilis, Şanlıurfa, Diyarbakır, Mardin, Batman, Siirt, Şırnak

Table 2. Data Statistics and sample size for post-stratified (PS) and stratified random sampling (SS)

Regions	Strata	N_h	PS n_h	SS n_h	$\bar{Y}$	$\bar{X}$	S_y^2	S_x^2	ρ	$\beta_2(x)$	Cx
TR1	City	32	32	32	91.524,03	283.924,97	3.463.599.046,55	34.847.228.468,87	0,991	0,52	0,43
	Village	214	214	214	2.535,66	4.360,45	52.889.125,46	167.619.986,90	0,999	75,63	8,82
	Total	246	246	246	14.111,38	40.726,57	1.383.997.448,52	13.435.248.085,12	0,994	12,10	8,10
TR2	City	57	9	9	9.109,63	23.183,98	165.173.697,1	1.576.655.334,00	0,989	9,3	1,71
	Village	2.272	460	347	377,51	474,72	333.710,85	843.395,24	0,987	97,61	1,93
	Total	2.329	469	356	591,22	1.030,5	6.120.042,14	57.017.115,37	0,986	395,72	7,33
TR3	City	129	50	44	12.382,96	34.382,61	809.395.273,1	5.854.780.666,00	0,996	33,46	2,23
	Village	4.306	1.622	1.479	542,17	637,31	641.415,5	967.330,25	0,996	49,09	1,54
	Total	4.435	1.672	1.523	886,58	1.618,86	27.948.620,36	202.120.344,30	0,993	1083,86	8,78
TR4	City	81	39	36	13.656,14	38.095,16	945.480.731,00	7.087.960.254,00	0,997	14,33	2,21
	Village	3.017	1.463	1.351	407,91	496,01	1.489.859,89	2.566.225,39	0,997	871,94	3,23
	Total	3.098	1.502	1.387	754,29	1.479,07	30.344.502,96	221.598.965,70	0,991	624,15	10,1
TR5	City	61	48	49	22.928,31	64.472,36	2.604.605.525,00	16.830.189.341	0,993	8,77	2,01
	Village	1.880	1.556	1.510	477,62	584,38	646.594,50	1.190.710,77	0,982	76,86	1,87
	Total	1.941	1.604	1.559	1.183,18	2.592,19	96.580.325,11	646.311.738,10	0,993	367,76	9,81
TR6	City	91	33	39	12.857,87	43.208,53	916.596.096,00	9.170.787.191,00	0,993	19,09	2,22
	Village	3.304	1.516	1.401	650,39	806,07	1.501.612,65	2.784.540,05	0,989	132,28	2,07
	Total	3.395	1.549	1.440	977,6	1.942,63	29.655.576,27	292.810.181,20	0,985	751,21	8,81
TR7	City	84	13	9	5.853,71	21.384,37	152.255.918,80	1.697.373.955,00	0,995	12,52	1,93
	Village	3.341	388	341	337,23	407,74	322.150,00	513.569,23	0,995	34,4	1,76
	Total	3.425	401	350	472,52	922,21	4.733.290,02	52.176.529,61	0,982	587,44	7,83
TR8	City	105	3	4	5.384,7	18.120,11	122.826.551,00	1.201.758.891,00	0,993	36,18	1,91
	Village	5.654	259	218	274,55	337,41	173.531,09	397.270,08	0,969	744,38	1,87
	Total	5.759	262	222	367,72	661,63	2.856.807,15	27.762.058,70	0,985	1657,62	7,96
TR9	City	79	6	5	3.972,34	15.201,11	56.995.016,89	549.432.895,20	0,992	26,4	1,54
	Village	2.637	203	175	379,45	467,89	363.117,01	588.347,24	0,993	26,98	1,64
	Total	2.716	209	180	483,95	896,43	2.354.669,06	22.488.481,83	0,96	695,84	5,29
TRA	City	57	2	4	4.707,84	16.606,54	116.160.955,20	1.452.467.676,00	0,991	37,08	2,29
	Village	3.125	220	192	221,25	264,12	90.654,64	140.329,42	0,996	43,37	1,42
	Total	3.182	222	196	301,62	556,86	2.488.226,11	30.407.781,65	0,985	1.869,18	9,9
TRB	City	70	8	7	5.212,67	20.458,21	141.478.876,30	2.151.065.028,00	0,996	21,07	2,27
	Village	3.011	332	295	308,63	369,24	282.567,67	435.420,57	0,997	62,81	1,79
	Total	3.081	340	302	420,05	825,66	3.979.804,59	57.578.555,27	0,98	957,78	9,19
TRC	City	77	8	11	8.724,32	35.756,57	369.359.947,4	4.817.432.866,00	0,982	12,58	1,94
	Village	4.021	678	592	320,38	383,95	431.708,92	765.952,83	0,997	1.617,67	2,28
	Total	4.098	686	603	478,29	1.048,59	8.577.706,63	113.189.623,7	0,974	766,58	10,2

Within the frame work of the above results (Table 2), the MSE and the efficiency for all the estimators in post-stratification and stratified random sampling, are calculated for every 11 regions. The efficiency of each estimator in post-stratification with respect to the sample mean of a stratified random sampling is defined as follows:

$$e(\bar{y}) = \frac{MSE(\bar{y}_{sr.})}{MSE(\bar{y}_{p.})}$$

where $MSE(\bar{y}_{sr.})$ is the mean square error for each estimator which is suggested in stratified random sampling for separate ratio estimate while $MSE(\bar{y}_{p.})$ is the mean square error for each estimator which is suggested in post-stratified sampling.

Therefore, the efficiency of ratio estimators which are suggested in post-stratification with respect to the stratified random sampling, a comparison is carried out for 11 different data sets.

In Table 3 and Table 4, the MSE and efficiency for estimators given in Section 2 are presented. Based on these results, it is noticed that the suggested estimators in post-stratification have the highest efficiency, i.e., they have smaller MSE than separate ratio estimate.

Table 3. $\boldsymbol{MSE(\bar{y}_{sr.})}$ **and** $\boldsymbol{MSE(\bar{y}_{p.})}$ **of estimators** $\boldsymbol{\bar{y}_r, \bar{y}_{SD}, \bar{y}_{SK}, \bar{y}_{US1}, \bar{y}_{US2}}$ **according to data of regions**

$MSE(\bar{y}_{sr.})$ / Regions	Classical-Ratio-R	SD	SK	US1	US2
TR-2	742.32	739.77	**683.78**	742.26	700.44
TR-3	86.74	86.72	86.85	86.74	**86.36**
TR-4	65.57	**64.95**	273.71	65.57	113.08
TR-5	224.80	224.78	**223.74**	224.80	223.86
TR-6	179.00	178.69	173.58	179.00	**172.65**
TR-7	161.46	161.24	161.14	161.46	**159.93**
TR-8	263.98	**261.51**	467.79	263.98	334.96
TR-9	467.03	**466.80**	469.36	467.02	467.23
TR-A	164.19	**163.96**	167.27	164.19	164.77
TR-B	81.84	**81.53**	90.69	81.83	82.60
TR-C	480.73	**479.95**	841.09	480.74	693.46
$MSE(\bar{y}_{p.})$ / Regions	**Classical-Ratio-R**	**SD**	**SK**	**US1**	**US2**
TR-2	602.24	600.37	**560.17**	602.19	572.17
TR-3	76.53	76.51	76.63	76.53	**76.20**
TR-4	58.21	**57.68**	237.57	58.21	99.15
TR-5	213.81	213.80	**212.92**	213.81	213.01
TR-6	165.80	165.54	161.07	165.80	**160.24**
TR-7	161.08	160.91	160.92	161.08	**159.82**
TR-8	250.39	**248.33**	422.28	250.39	310.33
TR-9	441.27	**441.10**	443.71	441.26	441.74
TR-A	207.52	**207.31**	210.04	207.52	207.95
TR-B	82.92	**82.66**	90.68	82.92	83.60
TR-C	434.40	**433.74**	743.88	434.41	617.09

The best estimators are highlighted in bold.

Table 4. Efficiencies of estimators in post-stratification ($\bar{y}_{p.}$) with respect to separate ratio estimators ($\bar{y}_{sr.}$)

$e(\bar{y})$ / Regions	Classical-Ratio-R	SD	SK	US1	US2
TR-2	**1,2326**	**1,2322**	**1,2207**	**1,2326**	**1,2242**
TR-3	**1,1334**	**1,1334**	**1,1334**	**1,1334**	**1,1333**
TR-4	**1,1264**	**1,1260**	**1,1521**	**1,1264**	**1,1405**
TR-5	**1,0514**	**1,0514**	**1,0508**	**1,0514**	**1,0509**
TR-6	**1,0796**	**1,0794**	**1,0777**	**1,0796**	**1,0774**
TR-7	**1,0024**	**1,0021**	**1,0014**	**1,0024**	**1,0007**
TR-8	**1,0543**	**1,0531**	**1,1078**	**1,0543**	**1,0794**
TR-9	**1,0584**	**1,0583**	**1,0578**	**1,0584**	**1,0577**
TR-A	0,7912	0,7909	0,7964	0,7912	0,7924
TR-B	0,9870	0,9863	**1,0001**	0,9869	0,9880
TR-C	**1,1067**	**1,1065**	**1,1307**	**1,1067**	**1,1238**

The best estimators are highlighted in bold.

The striking feature of the Table 3 is that the proposed estimators in post-stratification are uniformly most efficient for all the 9 data sets except 2 of them.

In addition, it should be pointed out that the order of MSE values of the estimators from smaller to bigger is similar in both post-stratification and separate ratio estimate.

Ratio estimators ($\bar{y}_r$, $\bar{y}_{SD}$, $\bar{y}_{SK}$, $\bar{y}_{US1}$, $\bar{y}_{US2}$) which are examined in this study are studied according to simple and stratified random sampling in literature before whereas in this study, these ratio estimators are suggested for post-stratified sampling. In addition to that, the MSE of these estimators are calculated for 11 different data sets and their efficiencies are calculated by comparing them with stratified random sampling. As a consequence, it is seen that ratio estimators are more effective in post-stratification.

It is under consideration that in subsequent studies, the usage of regression estimators in post-stratification will be examined and they will be compared with ratio estimators.

REFERENCES

Aksu, E.E., (2009). **Post-stratification and an application**. TURKSTAT Expertness thesis.
Bacanli, S., Kadilar, C. (2008). "Ratio estimators with unequal probability designs", **Pakistan Journal of statistics**, 24(3), 167-172.

Bethlehem, J. G. and Keller, W.J. (1987). "Linear Weighting of Sample Survey Data", **Journal Of Official Statistics** 3(2), 141-153.

Cervantes, I.F. and Brick, J.F. (2009). "Efficacy of Poststratification in Complex Sample Design", **Proceedings of the Survey Research Methods Section, Alexandria, VA: American Statistical Association**. 4642-4655.

Cochran, W.G. (1977). **Sampling Techniques**. New York; Wiley.

Hansen, M.H., Hurwitz, W.N., Madow W.G. (1953). **Sample Survey Methods and Theory Volume II-Theory**. Canada; John Wiley&Sons, Incorporation.

Holt, D., Smith, T.M.F.(1979). "Post-Stratification" , **Journal of the Royal Statistical Society A**, 142, 33- 46.

Kadilar, C., Cingi, H. (2003). "Ratio estimators in stratified random sampling", **Biometrical Journal**. 45(2003)2, 218-225.

Kim, J.J., Li, J., and Valliant, R. (2007). "Cell Collapsing in Poststratification", **Survey Methodology,** 33, 139-150.

Rueda, M. M., Gonzalez, S. and Arcos, A. (2006). "A general class of estimators with auxiliary information based on available units", **Applied Mathematics and computation**. 175, 131-148.
Särndal, C.E., Swensson, B., Wretman, J. (2003). **Model Assisted Survey Sampling**. New York; Springer.

Singh, S. (2003). **Advanced sampling theory with applications**. Kluwer, Netherlands

Sisodia, B. V. S. and Dwivedi, V. K.. (1981). "A Modified ratio estimator using coefficient of variation of auxiliary variable", **Journal of Indian Society Agricultural Statistics**. 33, 13–18.

Upadhyaya, L. N. and Singh, H. P., (1999). "Use of transformed auxiliary variable in estimating the finite population mean", **Biometrical Journal** 41, 5, 627–636.

Zhang, Li-Chun (2000). " Post-Stratification and Calibration- A Synthesis", **The American Statistician**, 54, 178-184.

ÖZET

SONRADAN TABAKALAMADA ORANSAL TAHMİN EDİCİLER

Bu çalışmada, literatürde verilen oransal tahmin edicilerin sonradan tabakalama yönteminde kullanımı incelenmiştir. Hata kareler ortalaması (HKO) elde edilmiş ve tabakalı rasgele örneklemede verilen oransal tahmin edicilerin HKO ile karşılaştırılmıştır. Türkiye İstatistik Kurumu tarafından gerçekleştirilen 2000 Türkiye Genel Nüfus Sayımı sonuçlarından idari birim bazında istihdam ve nüfus değerleri kullanılarak İstatistiki Bölge Birimleri Sınıflaması Düzey-1 bazında ortalama istihdam tahmin edilmiştir. Uygulama sonuçlarına göre sonradan tabakalama için önerilen oransal tahmin edicilerin daha iyi sonuç verdiği gösterilmiştir.