

# On the Use of Auxiliary Variables in Agricultural Surveys Design

Federica Piersimoni

Istat, Servizio Agricoltura, Italy, piersimo@istat.it

Roberto Benedetti

Università degli Studi “G.d’Annunzio” di Chieti-Pescara, Italy, benedett@unich.it

Giuseppe Espa

Università degli Studi di Trento, Italy, giuseppe.espa@economia.unitn.it

## Abstract

Auxiliary variables, both univariate and multivariate, must be efficiently used to obtain accurate estimates. They are useful *ex ante*, that is when the sample has to be drawn, but also *ex post*, as the weight calibration method.

The classical issue on efficient sample design through a stratification based on auxiliary information will be reviewed in comparison with sampling units selection methods that make an appropriate use of auxiliary variables. We will focus our attention on three approaches: the *model-based* approach, the  $\pi$ ps approach and the ranked set sampling method.

While the  $\pi$ ps is a well known sample design method, the *model-based* approach of sample surveys assumes that a *superpopulation* model is specified.

Ranked set sampling (further on *RSS*) has been introduced by McIntyre (1952). Since the publication of this seminal work, the literature proposed numerous *RSS* extensions for both parametric and non-parametric estimates.

In its original formulation, *RSS* starts with the selection of a simple random sample without reinsertion of  $n$  population units. Then the mean must be estimated from those units. The  $n$  are ranked in increasing order with respect to an auxiliary variable  $x$ , that is, without the effective calculation of the interest variable  $y$ .

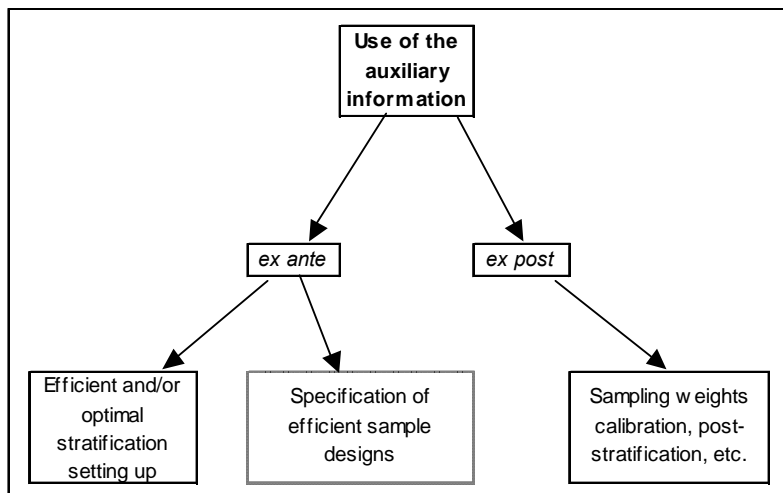
Which method is the best depends on the application at hand. To show some evidence, we compare the methods on real data on the slaughtering sector, for which the sample strata and the variables estimates are calculated, following the suggestions that are contained in Dorfman and Valliant (2000). Finally a comparison between *ex-ante* and *ex-post* use of the auxiliary information is performed. We will draw, for each selection method applied, 2000 samples based on the censuses lists in order to build the sample spaces and on the basis of them we will estimate which distribution is the “nearest” to the known true value in terms of MSE.

## 1. Introduction

Several National Institutes of Statistics (NIS), as the Italian one (ISTAT), usually make large efforts to project surveys that generally are based on an efficient use of all the available auxiliary information, either univariate or multivariate, in order to obtain more precise and reliable estimates. Such an utilization mainly consists of actions performed after the sample selection, as shown in not dashed boxes in the scheme of fig. 1. In fact, the most common context for the production of sample estimates consists in a *standard* design, that is generally a stratification of the list and a simple random selection of the units within the determined strata. Only after

the data collection and editing phase, the auxiliary information are used. It is in this phase that NIS make the greatest effort in the use, but also in the developing of very complex estimators that could lead to efficiency improvements. A lot of sample surveys done by ISTAT referring to the primary sector follow the above practice too. Among these, a special example is, also for the course of this paper, the red meat slaughtering monthly survey; it foresees a stratified sampling, with a stratification by kind of slaughter-houses and geographical division, for a total of 5 strata, two of which with geographical references. Strata are the following: stratum 1 (always totally observed), consisting of private with European Economic Community (EEC) stamp slaughter-houses in the geographical division 1 or 2; stratum 2: consisting of private with EEC stamp slaughter-houses in the geographical division 3, 4 or 5; stratum 3: private with low capacity slaughter-houses (apart from geographical division); stratum 4: private in derogation, public with EEC stamp and public in derogation (apart from geographical division) slaughter-houses; stratum 5: public with low capacity slaughter-houses. Two dimensional criteria that assign to stratum 1 those enterprises with more than 10.000 sheep and goats or more than 50.000 pigs slaughterings act in the stratification too. On the average the sample is of about 460 units for a population of about 2.200 units.

After the survey has been performed, the initial direct estimates are corrected through the use of calibration estimators (Deville and Särndal, 1992 - an outline will be given in the next section). The external consistency constraints are assigned referring to the total survey of two years before the current figures, because, at the estimates production time, the data of previous year survey are seldom available. Calibration acts in practice on each single category (for example calves) and not on the species (for example cattle) imposing, for each stratum, the respect of the two years before the current survey census total; in theory it would be possible to constrain with respect to all the 24 categories object of the survey; in practice this never happens because of non responses and the weighting generally is reduced to operate on the most important 4 or 5 variables.



**Figure 1:** use of the auxiliary information in sample surveys.

The purpose of this paper is to show how, combining the *ex ante* use of the auxiliary information with that one *ex post*, further improvements in terms of efficiency of the estimates can be gathered. The *ex ante* use of the auxiliary information to which we

are referring to, concerns the choice of sampling designs in which the auxiliary variables guide the selection of the sample units from the list of interest (cf. the dashed box in the scheme in fig. 1). A series of controlled experiments was conducted through simulations that will be described in section 2. Such experiments were built in order to test the efficiency of some sample selection criteria applied to the ISTAT slaughtering monthly survey (Espa *et al.*, 2001). Paragraph 3. is devoted to explain the main results of the simulation done and some final remarks. It could be asserted that the test carried out responds at the same time to another need: to supply proposals for the re-planning of the actual monthly sample of the ISTAT slaughtering survey.

## 2. The simulation performed: list and selection criteria adopted

The sampling frame for the experiments done was made up by  $N = 2.211$  units (enterprises) for which 12 variables were available. These were the same 4 variables (total number of slaughtered cattle, of pigs, of sheep and goats, and of equines), observed in all the population every year and in particular used with reference to the years 1999, 2000 and 2001. Such a frame was used to:

- i) plan the database necessary to obtain estimates with the use of *ex post* auxiliary information (auxiliary variables are those totally observed in 1999 and 2000);
- ii) verify (through the calculation of *RMSE*) the accuracy of the estimates obtained for 2001 (object year of the survey) by comparisons with known census figures of 2001. This is a way justified by the periodicity with which slaughtering figures are available in reality: beside census survey repeated every year, ISTAT carries out a monthly sample survey. Therefore our experiments ties to simulate just this monthly survey designed, as already said, using different sampling selection criteria; the simulated analysis then offers the possibility to carry out significant comparisons with the census data (through estimates empirical sampling distributions) to suggest a more efficient design for the monthly survey.

For every simulation, 2.000 samples with size  $n = 200$  were selected. Furthermore for each of the scenarios about which any simulation is performed, weight vector was calibrated, where possible, on the same variables as observed in 1999 and 2000. This in order to verify if the temporal delay of the auxiliary information gives rise to possible efficiency decreases of the estimates as well. The variables estimated through sampling means at 2001 are, in every single replication of the experiment, the totals of the four variables previously considered (cf. tab. 2).

For each of the selection criteria used, the estimates have been produced through the classical expansion estimator known in specialized literature as the Horvitz–Thompson estimator and through the calibration estimator of Deville and Särndal (1992). It is, in synthesis, an estimator that acts a calibration of the weights resulting from the original sample design that respect the external consistency conditions. In other words it is imposed that sampling distributions weighted for the auxiliary variables be in conformity with the same variables distributions known *a priori* through census data. An exception is the balanced sampling (or model–based approach to the inference in sampling survey, whose bases can be also found in Royall (1970)) that is constrained by definition to select a sample  $s$  satisfying the

condition  $\sum_s \pi_k X_k = \sum_U X_k$ .

Some needed clarifications about each of the selection criteria used are, in the same order of the first column of tab. 2, the following (more extended for those criteria less usually adopted in practice):

#### *Simple random sampling (SRS)*

For this sampling design, nothing has to be specified except the fact that the direct estimate doesn't use auxiliary information by definition and therefore it's not necessary to distinguish the two reference years.

#### *Stratified sampling (ST)*

On the contrary, in the case of stratified sampling, the auxiliary information acts, as already said, *ex ante* the samples selection in the setting up of the strata. Therefore in tab. 1 there are the results concerning direct estimates with basis 1999 and 2000 just to distinguish the reference year of the auxiliary variables on the basis of which strata have been set up. The strata themselves, in number of five for each experiment, have been set up in a way, so to say, based on the following considerations. Taking into account for each of the four variables object of estimate, the four census thresholds (Hidioglou, 1986), we dispose of two strata for each variable (one to be totally observed and the other one to be sampled). In total, therefore,  $2^4 = 16$  strata. By integrating strata similar in internal composition (namely integrating by addition of the strata codes), we obtain the five final strata:

- stratum 1, made up by firms always totally observed for all four variables;
- stratum 2, made up by firms totally observed three times (i.e. in relation to three out of the four variables);
- stratum 3, with units totally observed two times;
- stratum 4, with units totally observed only once;
- stratum 5, made up by units always sampled.

To conclude, it is specified that the allocation of the 200 units to the five strata has been made in accordance with the multivariate allocation model of Bethel (1989), taking into account that we have just simulated a multivariate survey.

#### *Ranked Set Sampling (RSS)*

The third selection criterion adopted is linked to the logic of Ranked Set Sampling, or RSS. The introduction of such sampling criterion is due to McIntyre (1952), who designed it with the aim of estimating the mean of agricultural variables and studied the advantages in comparison to the use of simple random sampling. Starting with this fundamental contribution, the specialized literature obtained then other contributions proposing numerous variants of the RSS both for parametric and non-parametric estimates (cf., among others, Li *et al.*, 1999, Bai and Chen, 2003 and Rosén, 1997) and interesting applications concerning sampling problems in many fields, taking into account that the most common utilizations concern agriculture and ecological studies (Patil *et al.*, 1994a).

In its original formulation, here utilized except for some modifications that will be mentioned later on, the RSS procedure calls for the simple random selection, without reinsertion, of  $n$  units of the reference population. The  $n$  sample units are ranked in a not-decreasing order, by mean of a criterion that doesn't call for an effective measure, in each unit, of the variable object of interest  $y$ . We could produce the ranking on the basis of the values assumed by the units in relation to an auxiliary  $x$

known at population  $U$  level. In this case it is evident that the “quality” of the auxiliary has to be linked not to the linear correlation between  $x$  and  $y$  but mainly to the correlation between the  $x$  and the  $y$  ranks (i.e. a weaker link). Later on we will suppose that the ranking has been made in respect to the auxiliary variable  $x$ , known at population  $U$  level.

Continuing with the  $RSS$ , with reference to the first chosen sample, we proceed to the quantification, that is to the measure of the variable object of interest  $y$ , of only the first unit of the ranked group. Then we draw a second sample in the same way of the first one, and the sample observation is applied only on the second unit of the ranking. This procedure is repeated up to the completion of the  $RSS$  cycle requiring  $n$  replications. A whole cycle can be repeated also  $m$  times.

The  $RSS$  increases the precision of the estimate of an average in respect of simple random sampling of same size  $n$ . This remains valid also if there are errors in the ranking; however the relative precision decreases with the increase of errors in the ranking. At the limit  $RSS$  and  $SRS$  are equivalent if the ranking is made up in a completely random manner. Therefore the trade-off to be carefully evaluated in the applications is that one between the precision increase and computational cost increase for ranking. Anyway an  $RSS$  efficiency increase (in comparison with the simple random sampling) cannot be obtained increasing the cycles number  $m$ , but increasing the sample size  $n$  which, on the other side, in practice has to be very reduced (Al-Saleh and Al-Omari, 2002).

An  $RSS$  selection requires therefore a basis of  $n^2$  units ( $n$  samples, each of  $n$  units), but it imposes to make the sample observation only on the  $n$  units fixed according to the principles specified above. For  $RSS$  comprehension, can be useful to refer to table 1 below:

$y_{(1)1}$	$y_{(1)2}$	...	$y_{(1)i}$	...	$y_{(1)n}$
$y_{(2)1}$	$y_{(2)2}$	...	$y_{(2)i}$	...	$y_{(2)n}$
$\vdots$	$\vdots$		$\vdots$		$\vdots$
$y_{(i)1}$	$y_{(i)2}$	...	$y_{(i)i}$	...	$y_{(i)n}$
$\vdots$	$\vdots$		$\vdots$		$\vdots$
$y_{(n)1}$	$y_{(n)2}$	...	$y_{(n)i}$	...	$y_{(n)n}$

**Tab. 1.** Database of  $n^2$  units:  $n$  ranked sets, each composed by  $n$  units. The observations corresponding to the chosen units are in evidence.

The final sample  $\{y_{(1)1}, \dots, y_{(i)i}, \dots, y_{(n)n}\}$ , known as rank ordered sample is that one utilized for inference tasks.

To be noted that, also if in the various phases of the cycle the selection is without reinsertion, ex post the sample can result with reinsertion. In fact it can happen that a certain unit ( $i$ ) coming from the  $i$ -th phase of the  $RSS$  cycle, be the ( $k$ )-th of the phase  $k$  with  $k \neq i$ .

A part the technical aspects that here are not relevant, the key parameter to implement a  $RSS$  selection is the variable utilized for ranking. The way utilized to treat a multivariate problem, with a selection criterion that is, by nature, based on only one variable, was to utilize as ranking variable the vector of the first order

inclusion probabilities, generated for implementing the selection with probabilities proportional to the dimension of the units ( $\pi PS$ ).

This solution could possibly refer to that of Patil *et al.* (1994b), who propose to base the samples selection on the ranking in relation of only an auxiliary variable, individualized as the most relevant, with the other variables called to perform as concomitant. This way to proceed is applicable with a certain reason if the primary variable is strongly and positively correlated with the concomitant (the rankings based on the single variables should be enough similar). It is of dubious applicability in the case such correlation is not important or quite negative. We will come back on this point in section 3., with particular reference to the simulation performed. To conclude we point out that four different algorithms of samples selection representing effectively extensions to the case of two or more variables, are described and treated by Ridout (2003).

#### *Probability proportional to size ( $\pi PS$ )*

Referring to this last criterion, we had initially four vectors of first order inclusion probabilities, one for each auxiliary at time  $t$ :  $\pi_{k,i}^t = \frac{nx_{k,i}}{\sum_U x_{k,i}}$ ,  $k=1,2,\dots, N$ ,  $i=1,4$  and

$t=1999$  and  $2000$ . In order to consider all the four auxiliaries in the sample selection process, we have then decided to change from the four vectors of first order inclusion probabilities to the vector of the averages of such probabilities. This choice assures, among others, the minimization of the maximum of the four variances of the auxiliary variables (the result is easily demonstrated by observing that the probability medium vector makes equal the four variances in question). To conclude, we only specify that all these units  $k$  for which  $\bar{\pi}_k > 1$ , caused by the presence of values  $x_k$  excessively high, are surely included in the sample. Therefore the solution here adopted to build a multivariate survey (that requires the measure of more than a variable in each sample) is not a true multivariate solution.

#### *Balanced samplings*

As already said, the model-based approach of sample surveys assumes that a model  $\xi$  is specified. Model  $\xi$  is called the *superpopulation* model and represents the distribution of the random vector  $\mathbf{Y}=(Y_1,\dots,Y_k,\dots,Y_N)$  where  $Y_k$  is a random variable of the  $k$  vector entry. The real population  $\mathbf{y}=(y_1,\dots,y_k,\dots,y_N)$  is a realization of  $\mathbf{Y}$ .

Without losing generality, let the total  $t = \sum_{k=1}^N y_k$  to be estimated. Then sample  $s$  is to be drawn, whose data are  $\{y_k : k \in s\}$ . A  $t$  estimator depends on those data, in such a way that  $\hat{t} = \hat{t}(\{y_k : k \in s\})$ .

If  $\xi$  is fixed and for a given  $s$ , the distribution of  $(\hat{t} - t)$  can be determined and  $MSE_{\xi}(\hat{t})$  can be estimated. A prediction interval for  $t$  is calculated, to be not confused with a confidence interval, because the approach is different.

The *model-based* inference assumes that a long series of vector  $\mathbf{Y}$  realizations are available, for fixed drawn sample  $s$ . The inference relies on the specified model and is constrained by  $s$ , the sample composed of the effective realization of the variable. Sample design  $p(s)$  and inclusion probability do not play any role in the

inference. This approach has some methodological advantages, see among others Royall, 1970, Royall, 1992, Royall and Herson, 1973a e 1973b, where the *model-based* approach is introduced and then formalized within a robustness approach. Conversely, some authors claim that finite populations cannot be replied, since there is only one finite population so that the so called *design-based* (or *randomization inference*) is more appropriate, see (Särndal *et al.*, 1992, pp. 533–535). The model-based approach consists in finding an estimator in such a way that, for given  $s$ , the mean square error determined by the model  $MSE_{\xi}(\hat{t}) = E_{\xi}[(\hat{t} - t)^2 | s]$  under the constraint that the estimator must be unbiased, that is:  $E_{\xi}[(\hat{t} - t) | s] = 0$ , is minimum. In practice, it may be the case that the sample is drawn in such a way that the minimum of the objective function  $MSE_{\xi}(\hat{t})$  is reached by the selected units, but in this case the sample is not random anymore (Royall, 1970). Detractors of the Model-based approach claim that such a sample can be risky (Särndal *et al.*, 1992, p. 517).

If the model has not been well specified, common estimators can be biased. Royall and Eberhardt (1975) studied the problem in the case of the ratio estimator proposing estimators more robust to the errors of model specification. As far the sample selection is concerned, Royall and Eberhardt (1975) suggest to balance the sample with respect to the auxiliary variable  $x$ . In other words, they suggest to choose a sample such that  $\bar{x}_U = \bar{x}_s$ . The sample balance prevents to control and to limit the effects of the bias that occurs if model is false.

To conclude, with reference to the last two selection criteria used (balanced sampling and  $\pi PS$  balanced sampling), it is only necessary to specify that the balance constraint is valid for all the four variables, subject to estimation. The difference between the two selection criteria simply relates to the respect of the balance constraint, that in the second case is imposed *ex post* to selected samples with probabilities proportional to the dimension.

### 3. Conclusions

In table 2 the main results of the simulation performed and in figure 2 the simulated sampling distributions are shown. The principal indications about the auxiliary information use coming out from the figures can be synthetized in the following considerations. Above all and with equal conditions, it is better to impose the balance constraints in design phase, than in *ex post*. This result, in reality obvious thinking by similarity to the advantages of the stratification in respect of post-stratification, is evident by comparing the RMSE of the simple random sampling with calibration and that one relating to the balanced sampling.

Secondly, the best performances among the adopted criteria are clearly associated to the balanced  $\pi PS$  selections and  $\pi PS$  with calibration. The two situations differ by almost nothing and in reality are not improvable because very near of a minimum point. From an applicative point of view our opinion is that only a joint use of complex estimators together with efficient sampling designs may reduce considerably the variability of the estimates. The drawback is that balanced  $\pi PS$  and  $\pi PS$  with calibration selection criteria are on the one hand more efficient but on the other hand less robust of the others when outliers are present (cf. figure 2. In which the bimodal

distributions corresponding to the criteria at issue can be observed). A not expected result concerns the bad performance of *RSS* method. Such a result can be ascribed to the forced univariate use of the auxiliary information for the ranking setting up when linear independence is present.

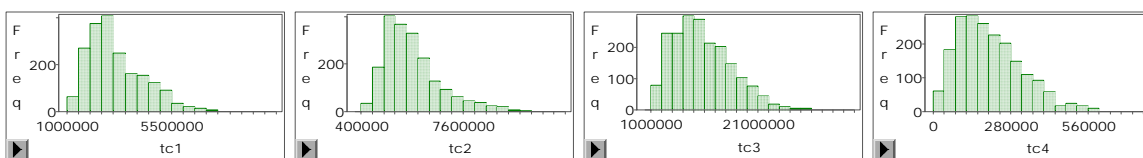
**Tab. 2: Synthesis prospect of the main results of the simulation performed.**

Selection criterion	Estimator	Aux. var. year	RMSE (as % of the estimate)			
			CATTLE	SHEEP AND GOATS	PIGS	EQUINES
SRS	Direct	-----	40,00%	51,25%	51,65%	57,70%
	Calibration	2000	14,12%	25,31%	27,43%	26,45%
		1999	20,31%	31,66%	26,73%	35,12%
Stratified	Direct	2000	28,04%	32,76%	34,43%	33,84%
		1999	26,47%	32,17%	33,92%	30,36%
	Calibration	2000	4,71%	9,04%	13,62%	9,77%
		1999	12,95%	12,43%	14,53%	12,87%
Ranked	Direct	2000	39,33%	47,91%	47,33%	51,23%
		1999	38,82%	47,82%	46,12%	52,43%
	Calibration	2000	13,52%	21,80%	22,36%	25,68%
		1999	18,18%	31,77%	22,80%	35,18%
$\pi$ PS	Direct	2000	6,17%	6,38%	15,61%	3,74%
		1999	7,28%	10,18%	17,20%	6,79%
	Calibration	2000	4,52%	5,04%	14,87%	2,57%
		1999	6,04%	9,28%	16,67%	6,48%
Balanced	Direct	2000	6,24%	13,48%	17,26%	15,05%
		1999	23,57%	20,05%	17,46%	19,43%
Bal/ $\pi$ PS	Direct	2000	5,55%	5,08%	14,60%	2,50%
		1999	6,37%	9,45%	16,72%	6,96%

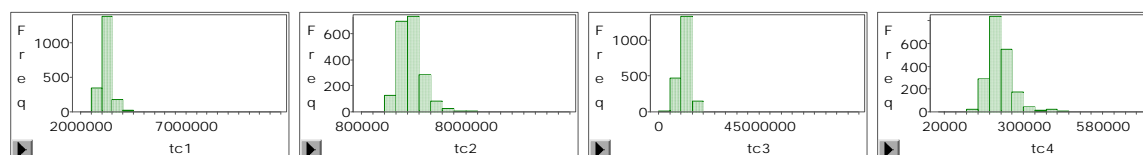
**Figure 2:** simulated sampling distributions according to the different sampling designs, where

- tc1 = total number cattle slaughtering in 2001;
- tc2 = total number sheep and goats slaughtering in 2001;
- tc3 = total number pigs slaughtering in 2001;
- tc4 = total number equines slaughtering in 2001:

SRS, direct estimates

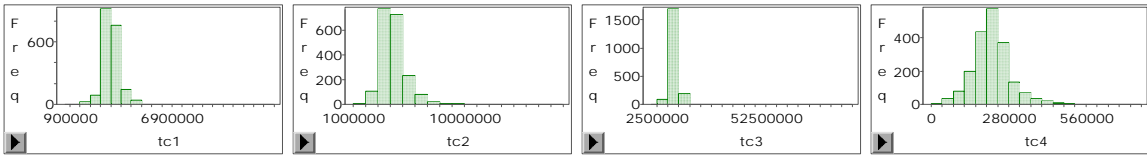


SRS, calibration, based on auxiliary variables of 2000

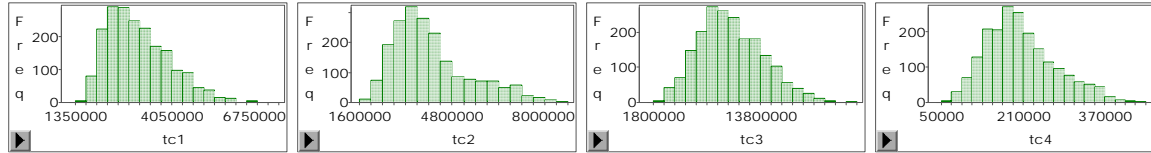




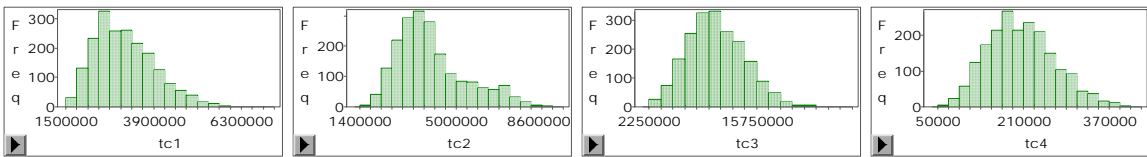
SRS, calibration, based on auxiliary variables of 1999



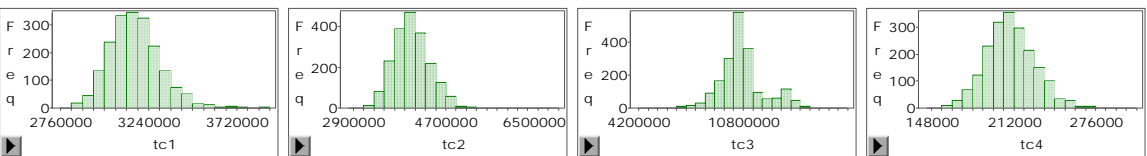
Stratified, direct estimates, based on auxiliary variables of 2000



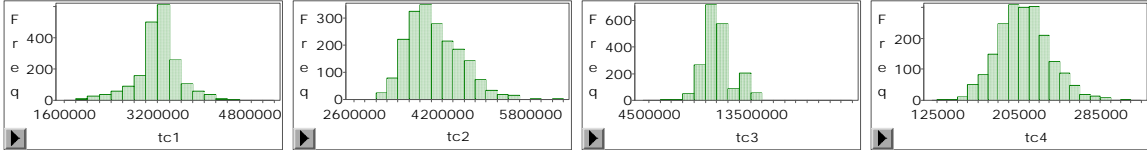
Stratified, direct estimates, based on auxiliary variables of 1999



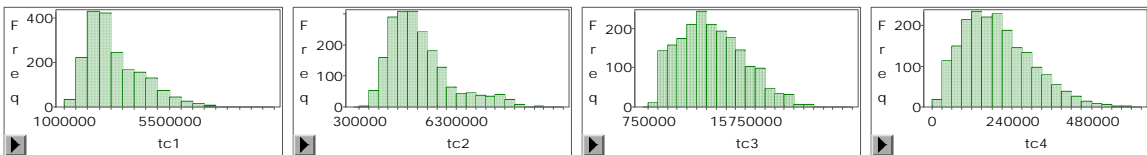
Stratified, calibration, based on auxiliary variables of 2000



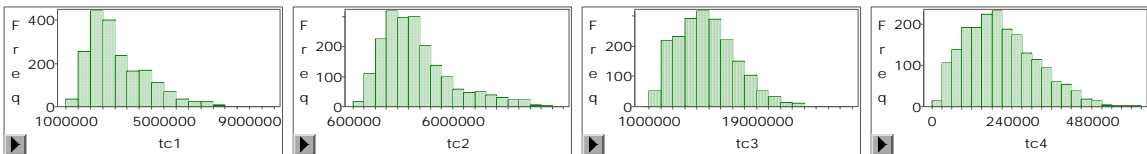
Stratified, calibration, based on auxiliary variables of 1999



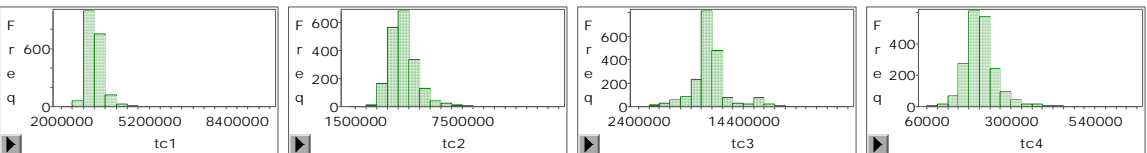
Ranked set, direct estimates, based on auxiliary variables of 2000



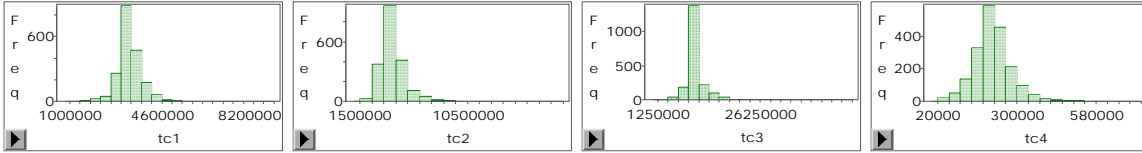
Ranked set, direct estimates, based on auxiliary variables of 1999



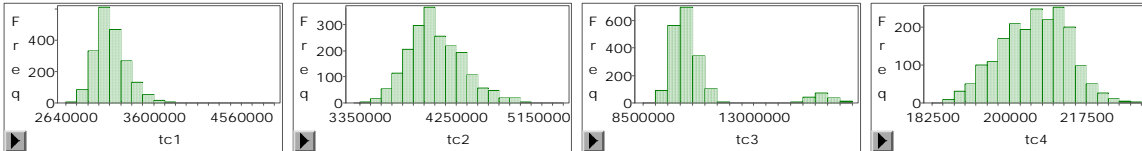
Ranked set, calibration, based on auxiliary variables of 2000



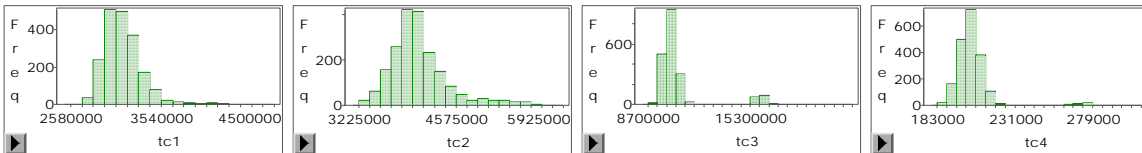
Ranked set, calibration, based on auxiliary variables of 1999



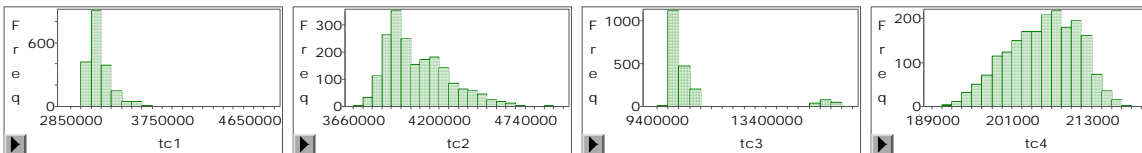
$\pi$ ps, direct estimates, based on auxiliary variables of 2000



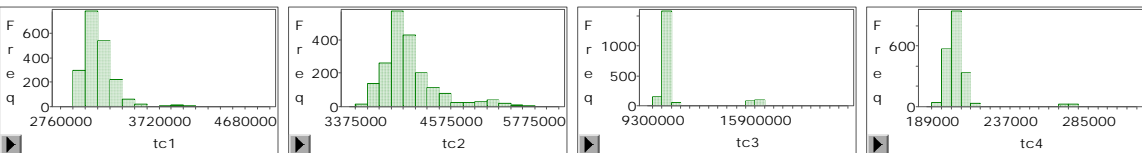
$\pi$ ps, direct estimates, based on auxiliary variables of 1999



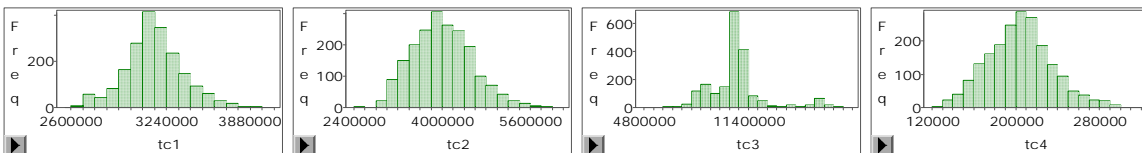
$\pi$ ps, calibration, based on auxiliary variables of 2000



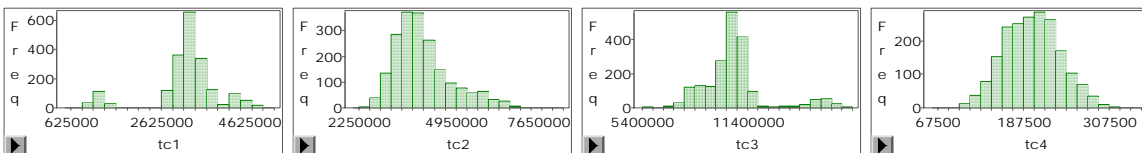
$\pi$ ps, calibration, based on auxiliary variables of 1999



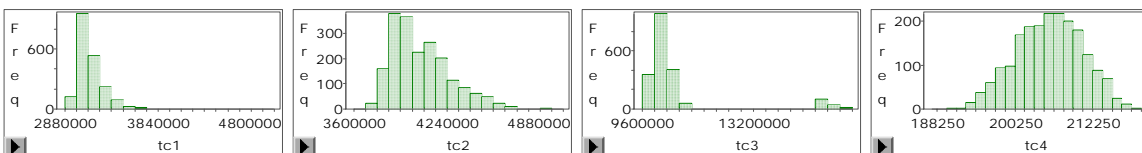
Balanced, direct estimates, based on auxiliary variables of 2000



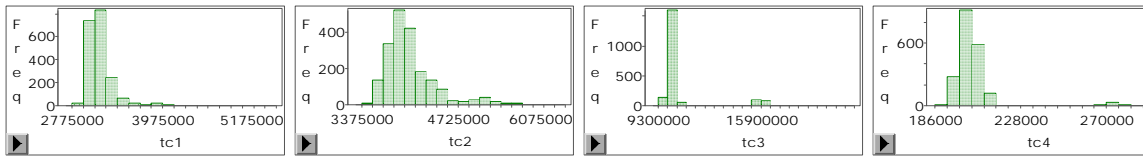
Balanced, direct estimates, based on auxiliary variables of 1999



Balanced,  $\pi$ ps, direct estimates, based on auxiliary variables of 2000



Balanced,  $\pi$ ps, direct estimates, based on auxiliary variables of 1999



## References

- Al-Saleh M.F., Al-Omari A.I. (2002) Multistage ranked set sampling, *Journal of Statistical Planning and Inference*, 102, 273–286.
- Bai Z., Chen Z. (2003) On the theory of ranked-set sampling and its ramifications, *Journal of Statistical Planning and Inference*, 109, 81–99.
- Bethel J. (1989) Sample allocation in multivariate surveys, *Survey Methodology*, 15, 47–57.
- Deville J.C., Särndal C.E. (1992) Calibration Estimators in Survey Sampling, *Journal of the American Statistical Association*, 87, 418, 376–382.
- Dorfman A.H., Valliant R. (2000) Stratification by size revised, *Journal of Official Statistics*, 16, 2, 139–154.
- Espa G., Benedetti R., Piersimoni F. (2001) Prospettive e soluzioni per il data editing nelle rilevazioni in agricoltura, *Statistica Applicata*, 13, 4, 363–391.
- Hidiroglou M.A. (1986) The construction of a self-representing stratum of large units in survey design, *The American Statistician*, 40, 1, 27–31.
- Li D., Sinha B.K., Perron F. (1999) Random selection in ranked set sampling and its applications, *Journal of Statistical Planning and Inference*, 76, 185–201.
- McIntyre G.A. (1952) A method for unbiased selective sampling, using ranked set, *The Australian Journal of Agricultural and Resource Economics*, 3, 385–390.
- Patil G.P., Sinha A.K., Taillie C. (1994a) Ranked set sampling, in G.P. Patil and C.R. Rao (eds) *Handbook of Statistics, Volume 12, Environmental Statistics*, North Holland Elsevier, New York, 167–200.
- Patil G.P., Sinha A.K., Taillie C. (1994b) Ranked set sampling for multiple characteristics, *International Journal of Ecology and Environmental Sciences*, 20, 94–109.
- Ridout M.S. (2003) On ranked set sampling for multiple characteristics, *Environmental and Ecological Statistics*, 10, 255–262.
- Rosén B. (1997) On sampling with probability proportional to size, *Journal of Statistical Planning and Inference*, 62, 159–191.
- Royall R.M. (1970) On finite population sampling theory under certain linear regression models, *Biometrika*, 57, 2, 377–387.
- Royall R.M. (1992) Robustness and optimal design under prediction models for finite populations, *Survey Methodology*, 18, 179–185.
- Royall R.M., Herson J. (1973a) Robust estimation in finite populations I, *Journal of the American Statistical Association*, 68, 344, 880–889.
- Royall R.M., Herson J. (1973b) Robust estimation in finite population II: stratification on a size variable, *Journal of the American Statistical Association*, 68, 344, 890–893.
- Särndal C-E, Swensson B., Wretman J. (1992) *Model Assisted Survey Sampling*, Springer Verlag, New York.