Survey Methodology, December 1994
Vol. 20, No. 2, pp. 171–176
Statistics Canada

Empirical Comparison of Small Area Estimation Methods for the Italian Labour Force Survey

P.D. FALORSI, S. FALORSI and A. RUSSO[1]

ABSTRACT

The study was undertaken to evaluate some alternative small areas estimators to produce level estimates for unplanned domains from the Italian Labour Force Sample Survey. In our study, the small areas are the Health Service Areas, which are unplanned sub-regional territorial domains and were not isolated at the time of sample design and thus cut across boundaries of the design strata. We consider the following estimators: post-stratified ratio, synthetic, composite expressed as linear combination of synthetic and of post-stratified ratio, and sample size dependent. For all the estimators considered in this study, the average percent relative biases and the average relative mean square errors were obtained in a Monte Carlo study in which the sample design was simulated using data from the 1981 Italian Census.

KEY WORDS: Small area estimators; Unplanned domains; Bias; Mean Square Error; Simulation study.

1. INTRODUCTION

In Italy, as in many other countries, there is a growing need for current and reliable data on small areas. This information need concerns most sample surveys realised by the Italian National Statistical Institute (ISTAT), especially the Labour Force Survey (LFS), which has been studied to warrant accuracy in regional estimates.

In the past, ISTAT's solution to this problem was to broaden the sample without changing the estimation method (Fabbris *et al.* 1988). In the last few years, however, in order to find a solution to the negative aspects of over-sized samples, research has been launched to identify estimation methods to improve the accuracy of small areas estimates (Falorsi and Russo 1987, 1989, 1990 and 1991).

In our study, the small areas are the Health Service Areas (HSA), which are unplanned sub-regional territorial domains and were not isolated at the time of sample design and thus cut-across the boundaries of the design strata. The sizes of these territorial domains are such that the reliability of regular estimates would have been satisfactory had these domains been designed with separate fixed sample sizes from individual domains.

The study was undertaken to evaluate some of the alternative small areas estimators to produce HSA level estimates from the LFS.

We consider the following estimators: post-stratified ratio, synthetic, composite (expressed as linear combination of the synthetic and of the post-stratified ratio), and sample size dependent.

For all the estimators considered in this study, the average percent relative biases and the average relative mean square errors were obtained in a Monte Carlo study in which the LFS design was simulated using data from the 1981 Italian Census.

2. BRIEF DESCRIPTION OF THE LFS SAMPLE STRATEGY

2.1 Design

The LFS is based on a two stage sample design stratified for the primary sampling units (PSU). The PSUs are the municipalities, while the secondary sampling units (SSU) are the households. In the framework of each geographical region the PSUs are divided according to the provinces. In each province the PSUs are divided into two main area types: the self-representing area consisting of the larger PSUs, and the non self-representing area consisting of the smaller PSUs.

All PSUs in the self-representing area are sampled, while the selection of PSUs in the non self-representing area is carried out within the strata that have approximately equal measures of size. Two sample PSUs are selected from each stratum without replacement and with probability proportional to size (total number of persons). The SSUs are selected without replacement and with equal probabilities from the selected PSUs independently. All members of each sample household are enumerated.

2.2 Estimator of Total

With reference to the generic geographical region, we introduce the following subscripts: h, for stratum ($h = 1, \ldots, H$); i, for primary sampling unit; j, for secondary sampling units; g, for age-sex groups ($g = 1, \ldots, G$).

[1] P.D. Falorsi, Senior Researcher, National Statistical Institute, Rome, Italy; S. Falorsi, Researcher, National Statistical Institute, Rome, Italy; Aldo Russo, Associate Professor, University of Molise, Campobasso, Italy.

In the present study we consider the following age classes 14-19, 20-29, 30-59, 60-64, and over 65.

A quantity referring to stratum h, primary sampling unit i, and secondary sampling unit j will be briefly referred to as the quantity in hij; and a quantity referring to stratum h and primary sampling unit i will be referred to as the quantity in hi.

The following notations are also used: N_h, for number of PSUs in h; P_h, for total number of persons in h; n_h, for number of sample PSUs selected in h; M_{hi} for number of SSUs in hi; P_{hi}, for total number of persons in hi; m_{hi}, for number of sample SSUs selected in hi; P_{ghij}, for number of persons in group g belonging to hij; P_{hij}, for number of persons in hij.

Further let

$$Y = \sum_{g=1}^{G} \sum_{h=1}^{H} \sum_{i=1}^{N_h} \sum_{j=1}^{M_{hi}} Y_{ghij}$$

be the total of the characteristic y for regional population, where Y_{ghij} denotes total of the characteristic of interest y for the P_{ghij} persons. Actually, the estimate of Y is obtained by a post-stratified estimator. This estimator is given by:

$$\hat{Y} = \sum_{g=1}^{G} \frac{\hat{Y}_g}{\hat{P}_g} P_g,$$

where

$$\hat{Y}_g = \sum_{h=1}^{H} \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} K_{hij} Y_{ghij}; \; \hat{P}_g = \sum_{h=1}^{H} \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} K_{hij} P_{ghij}$$

represent unbiased estimates of

$$Y_g = \sum_{h=1}^{H} \sum_{i=1}^{N_h} \sum_{j=1}^{M_{hi}} Y_{ghij} \; ; \quad P_g = \sum_{h=1}^{H} \sum_{i=1}^{N_h} \sum_{j=1}^{M_{hi}} P_{ghij}.$$

In the above formulas, the symbol K_{hij}, that denotes the basic weight, is expressed by:

$$K_{hij} = \frac{P_h}{n_h P_{hi}} \frac{M_{hi}}{m_{hi}}.$$

3. SMALL AREA ESTIMATORS

With reference to the generic geographical region, we suppose that the population P is divided into D non-overlapping small areas $1, \ldots, d, \ldots, D$ for which estimates are required. Each area is obtained by an aggregation of municipalities. The problem considered is the estimation the total of a y-variable for all units belonging to the small area d. In practice, the small area d will have a non-null intersection with only a certain number of design strata which we denote as $\tilde{H} = \{h \mid {}_dP_h > 0\}$, where ${}_dP_h$ represents the part of P_h belonging to the small area d.

Denoting by ${}_dN_h$ the number of PSUs belonging to small area d in stratum h, we seek to estimate the small area total

$${}_dY = \sum_{g=1}^{G} \sum_{h=1}^{\tilde{H}} \sum_{i=1}^{{}_dN_h} \sum_{j=1}^{M_{hi}} Y_{ghij}.$$

The development of a particular estimation method for small areas basically depends on available information. In Italy the accessible information at small area level is very poor. At present the accessible territorial information is total population by sex for each municipality collected through register statistics. In a future context (at end of 1994), the population counts by age-sex group will be available for each municipality. For this reason, in the present study we consider only those small area estimators that utilize, as auxiliary information, the population total by age-sex group.

3.1 Post-stratified Ratio Estimator

A post-stratified ratio estimator (POS) of ${}_dY$ is given by:

$${}_d\hat{Y}_{\text{POS}} = \sum_{g=1}^{G} \frac{{}_d\hat{Y}_g}{{}_d\hat{P}_g} \, {}_dP_g, \qquad (1)$$

where

$${}_d\hat{Y}_g = \sum_{h=1}^{\tilde{H}} \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} K_{hij} Y_{ghij} \delta_{hi},$$

$${}_d\hat{P}_g = \sum_{h=1}^{\tilde{H}} \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} K_{hij} P_{ghij} \delta_{hi},$$

$${}_dP_g = \sum_{h=1}^{\tilde{H}} {}_dP_{gh} = \sum_{h=1}^{\tilde{H}} \sum_{i=1}^{{}_dN_h} \sum_{j=1}^{M_{hi}} P_{ghij},$$

in which ${}_dP_{gh}$ denotes the total population for the age/sex group g in small area d intersected by stratum h, δ_{hi} is a binary variate that equals 1 if the PSU hi belongs to the small area d and equals 0 otherwise. For a better explanation of formula (1), we observe that PSU is a subset of small area and then does not intersect it.

The post-stratified ratio estimator is unbiased except for the effect of ratio estimation bias which is usually negligible. The estimator is defined to be zero when there is no sample within the domain. This estimator is not reliable for small sample sizes.

3.2 Synthetic Estimator

For computing a synthetic estimator, it is assumed that the small area population means for given population sub-groups are approximately equal to the larger area populations means of the same sub-groups. This estimator is obtained by means of a two steps procedure: (i) with respect to an aggregated territorial level, estimates of the investigated features are determined for population sub-groups; (ii) estimates for the aggregated territorial level area are then scaled in proportion to the sub-group incidence within the small domain of interest.

The synthetic estimator has a low variance since it is based on a larger sample, but it suffers from bias depending on the distance from the assumption of homogeneity, for each subgroup, between the small area and the larger area with reference to the characteristic of interest, y. The problems associated with synthetic estimators have been documented by Purcell and Linacre (1976), Gonzalez and Hoza (1978), Ghangurde and Singh (1978), Schaible (1979) and Levy (1979) among others.

In this study we consider the following form of synthetic estimator (SYN):

$$ {}_d\hat{Y}_{\mathrm{SYN}} = \sum_{g=1}^{G} \frac{\tilde{Y}_g}{\tilde{P}_g}\, {}_dP_g, \tag{2}$$

where

$$\tilde{Y}_g = \sum_{h=1}^{\bar{H}} \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} K_{hij}\, Y_{ghij};\ \tilde{P}_g = \sum_{h=1}^{\bar{H}} \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} K_{hij}\, P_{ghij}.$$

3.3 Composite Estimator

The composite estimator (COM) considered here is obtained as a linear combination of the estimators SYN (biased with low sample variance) and POS (less biased with high sample variance):

$$ {}_d\hat{Y}_{\mathrm{COM}} = \alpha\, {}_d\hat{Y}_{\mathrm{POS}} + (1-\alpha)\, {}_d\hat{Y}_{\mathrm{SYN}}, \tag{3}$$

where α is a constant $(0 \le \alpha \le 1)$. This estimator minimizes the chances of extreme situations (both in terms of bias and sample variance). Therefore, in a given concrete situation such estimator may turn out to be more advantageous than its two components considered separately.

The optimum value for α that minimizes the MSE of the COM estimator is given by

$$\alpha_{\mathrm{opt}} = \frac{\mathrm{MSE}({}_d\hat{Y}_{\mathrm{SYN}}) - E({}_d\hat{Y}_{\mathrm{SYN}} - {}_dY)({}_d\hat{Y}_{\mathrm{POS}} - Y_d)}{\mathrm{MSE}({}_d\hat{Y}_{\mathrm{SYN}}) + \mathrm{MSE}({}_d\hat{Y}_{\mathrm{POS}}) - 2E({}_d\hat{Y}_{\mathrm{SYN}} - {}_dY)({}_d\hat{Y}_{\mathrm{POS}} - Y_d)}. \tag{4}$$

Furthermore, when neglecting the covariance term in (4), under the assumption that this term will be small relative to $\mathrm{MSE}({}_d\hat{Y}_{\mathrm{SYN}})$ and $\mathrm{MSE}({}_d\hat{Y}_{\mathrm{POS}})$, the optimal weight α can be approximated by

$$\alpha^*_{\mathrm{opt}} = \frac{\mathrm{MSE}({}_d\hat{Y}_{\mathrm{SYN}})}{\mathrm{MSE}({}_d\hat{Y}_{\mathrm{SYN}}) + \mathrm{MSE}({}_d\hat{Y}_{\mathrm{POS}})}. \tag{5}$$

This is the approach to define weights followed by Schaible (1978).

In our work the optimal values of α have been obtained from Census data using formula (5). When considering a real sample survey only an estimated value of optimum α may be used, thus resulting in a decrease in efficiency.

3.4 Sample Size Dependent Estimator

The sample size dependent estimator is a particular case of the composite estimator. The linear combination of synthetic and of the less biased estimator is made for each sub-group and depends on the outcome of the given sample. We consider the following form of sample size dependent estimator (SD) which take into account the realized sample size in the small area. It is defined as (Drew, Singh and Choudhry 1982):

$$ {}_d\hat{Y}_{\mathrm{SD}} = \sum_{g=1}^{G} \left\{ \alpha_g \left(\frac{{}_d\hat{Y}_g}{{}_d\hat{P}_g}\, {}_dP_g \right) + (1-\alpha_g) \frac{\tilde{Y}_g}{\tilde{P}_g}\, {}_dP_g \right\}, \tag{6}$$

where

$$\alpha_g = \begin{cases} 1/({}_dR_g F) & 1/{}_dR_g < F, \\ 1 & \text{otherwise} \end{cases} \tag{7}$$

with ${}_dR_g = {}_dP_g / {}_d\hat{P}_g$.

The constant F is chosen to control the contribution of the synthetic component. The reliance on the synthetic portion decreases as the value of F increases. The choice of the value for F would depend upon several factors. In our study the efficiency of sample dependent estimator has been investigated for $F = 1$. This value proved to be efficient while affording protection against the bias of synthetic estimator.

The logic behind the SD estimator is that when the sample size within domain d and group g is small, then the direct estimate for domain d and group g would be unstable and a synthetic estimate may be superior. However, if the sample in domain d and group g is larger than expected this is not a problem, since the performance of the post-stratified direct part would improve as the sample size improves. In conclusion, we observe that SD estimator may be considered as a particular form of sample size dependent regression estimator given in Särndal and Hidiroglou (1989), that has good conditional properties.

4. DESCRIPTION OF THE EMPIRICAL STUDY

4.1 Simulation of the LFS Sample Design

In our study, we have considered the 14 HSAs of the Friuli region as small areas. The variable of interest, y, is the number of unemployed.

Evaluation of the performance of the various estimators, discussed in Section 3, was done by referring to a sample design (two stages with stratification of the PSUs) identical to that adopted for the LFS in Friuli. This design is based on the selection of 39 PSUs and 2,290 SSUs from a population of 219 PSUs and 465,000 SSUs.

We have selected independently 400 Monte Carlo sample replicates each of identical size (in terms of PSUs and of SSUs) of the LFS' sample. All the information utilized in the simulation is taken from the 1981 General Population Census, so ${}_dY$ is known.

4.2 Evaluation of Small Area Estimators

We denote by ${}_d\hat{Y}(mr)$ the estimate of the total ${}_dY$ for the small area d from the rth Monte Carlo replicate when using the estimator m. The percent relative bias of estimator m for the small area d is given by

$$ {}_d\mathrm{ARB}_m = \frac{1}{R}\left(\sum_{r=1}^{R} \frac{{}_d\hat{Y}(mr)}{{}_dY} - 1 \right) 100, $$

where R is the number of samples ($R = 400$).

The average of the percent absolute relative bias of estimator m over the whole set of small areas is:

$$ \overline{\mathrm{ARB}}_m = \frac{1}{D} \sum_{d=1}^{D} |\, {}_d\mathrm{ARB}_m \,|, $$

where D is the number of small areas under observation ($D = 14$).

The percent root mean square error of estimator m for small area d is

$$ {}_d\mathrm{RMSE}_m = \frac{\sqrt{{}_d\mathrm{MSE}_m}}{{}_dY} 100, $$

where the mean square error of estimator m for the small area d is expressed by

$$ {}_d\mathrm{MSE}_m = \frac{1}{R} \sum_{r=1}^{R} ({}_d\hat{Y}(mr) - {}_dY)^2. $$

The average percent root mean square error of estimator m over all areas is

$$ \overline{\mathrm{RMSE}}_m = \frac{1}{D} \sum_{d=1}^{D} {}_d\mathrm{RMSE}_m. $$

4.3 Analysis of Results

A. Overall Performance Measures

The average percent absolute biases and the average percent root mean square errors of the small area estimators for the LFS characteristic "number of unemployed persons" are presented in Table 1. Looking at this table, the following conclusions emerge:

(i) As expected, POS presents the smallest bias. The bias of SYN is larger than the bias of the other estimators. The bias of COM is roughly 30% lower than the bias of SYN estimator. The bias of SD estimator is only slightly lower than that of POS estimator.

(ii) SYN and COM have the smallest average percent root mean square errors, but these estimators are affected by a very high bias. POS, with low bias, is, conversely, the less efficient estimator. The average percent root mean square error of SD is approximately 30% higher than those of SYN and COM estimators.

Table 1

Average Percent Absolute Relative Bias $\overline{\mathrm{ARB}}$ and Average Percent Root Mean Square Error $\overline{\mathrm{RMSE}}$ for Unemployed by Estimator

Estimator	$\overline{\mathrm{ARB}}$	$\overline{\mathrm{RMSE}}$
POS	1.75	42.08
SYN	8.97	23.80
COM	6.00	23.57
SD	2.39	31.08

B. Performance Measures by Small Area

Tables 2 and 3 present the Percent Relative Bias (${}_d$ARB) and the Percent Root Mean Square Error (${}_d$RMSE) of the estimators for each of fourteen Health Service Areas in Friuli. Furthermore, Table 2 gives the percent ratio between the population of the HSA and the population of the set $\tilde{H}$ of strata including the HSA (p_1); Table 3 shows the percent ratio between the population of the HSA and the population of the region Friuli (p_2) and the percent ratio between the population of the set $\tilde{H}$ of strata including the HSA and the population of the region Friuli (p_3). Looking at these Tables, the following conclusions emerge:

(i) SYN and COM are badly biased in some small areas, namely, in those small areas where the model underlying SYN fits poorly. Generally the small areas with low values of the ratio p_1 are affected by large bias (*e.g.*, HSAs 1, 2, 3, 4 and 6). Conversely, large values of the ratio p_1 are associated with low values of the bias (*e.g.*, HSAs 5, 9, 10 and 13). However, SYN and COM consistently have an attractively low RMSE compared to other alternatives. In three of the fourteen areas (*viz*, areas 3, 4 and 8) COM is consistently the most efficient estimator. In two areas (10 and 12)

SYN is evidently more efficient and in the remaining areas the two estimators are roughly similar from the point of view of efficiency. Furthermore, we observe that the lowest values of RMSE for SYN generally are associated with the highest values of the ratio p_3 (*e.g.*, HSAs 1, 2, 5, 6, 9 and 13). HSAs 3 and 4, while having an high value of the ratio p_3, present a high value of RMSE. This is due to the large bias.

(ii) POS shows negligible bias values in almost all small areas. The RMSE values of POS are much higher than those of the other estimators in all the small areas. We observe that the RMSE of the POS estimator is negatively correlated with the ratio p_2. This is caused by the fact that the expected sample size increases as the ratio p_2 increases. Consequently, the variance (which is the main component of MSE of POS) decreases.

(iii) The estimator SD presents a negligible bias in seven (5, 7, 9 10, 11, 12 and 13) of the fourteen small areas. In the other areas the bias is quite low. Furthemore, in nine areas (2, 3, 4, 5, 9, 10, 11, 12 and 13) SD has a bias similar to that of POS. The estimator SD is better, from the MSE point of view, in comparison with POS. In four areas (7, 8, 9, and 13) RMSE is similar to those of SYN and COM.

(iv) Finally, we notice that in the largest areas with the highest values of the ratio p_2 (*e.g.*, HSAs 9 and 5) all the estimators considered give similar results in terms of bias and MSE. For the remaining areas, where the estimators have different performances, there is a problem in the choice of the best estimator.

Table 2

Percent Relative Bias ($_d$ARB) of Each of Fourteen Health Service Areas (HSA) in Friuli for Unemployed by Estimator

		Estimator			
HSA	p_1	POS	SYN	COM	SD
1	19.1	−1.57	−10.92	−7.68	−3.01
2	16.1	−5.61	−9.21	−6.97	−4.79
3	15.3	−5.21	28.82	17.98	5.79
4	16.3	−2.50	20.92	15.02	2.99
5	47.1	−0.46	1.61	0.98	−0.28
6	24.6	−1.37	−12.24	−9.06	−3.28
7	81.8	0.05	−6.25	−3.40	−1.66
8	70.7	0.81	11.80	6.63	2.17
9	92.2	0.47	0.76	0.68	0.78
10	71.2	0.36	−1.34	0.51	−1.02
11	21.7	−1.01	−5.64	−5.00	−1.62
12	40.6	−1.52	−6.66	−6.05	−1.19
13	56.3	−0.95	−3.12	−1.11	−1.28
14	21.8	−2.51	−6.21	−3.03	−3.53

p_1 = percent ratio between the population of the HSA and the population of the set $\tilde{H}$ of strata including the HSA.

Table 3

Percent Root Mean Square Error ($_d$RMSE) of Each of Fourteen Health Service Areas (HSA) in Friuli for Unemployed by Estimator

			Estimator			
HSA	p_2	p_3	POS	SYN	COM	SD
1	3.8	19.9	52.23	20.41	21.12	32.39
2	3.1	19.2	63.36	19.45	20.81	38.30
3	3.6	23.2	57.44	36.57	30.71	42.46
4	3.8	23.2	58.19	30.09	27.02	36.88
5	20.2	42.9	18.81	13.38	14.01	17.87
6	8.5	34.8	28.09	17.49	17.00	22.69
7	6.9	8.4	23.83	21.47	21.67	22.67
8	4.8	6.8	28.75	28.54	26.35	27.40
9	21.2	22.9	17.29	16.15	16.40	16.89
10	1.8	2.5	67.00	50.12	53.31	59.27
11	3.2	14.6	49.82	18.35	19.20	30.42
12	4.3	10.7	46.40	22.10	24.04	33.18
13	12.6	22.4	20.13	15.53	15.40	17.88
14	2.3	10.1	57.80	23.58	22.94	36.81

p_2 = percent ratio between the population of the HSA and the population of the region Friuli.

p_3 = percent ratio between the population of the set $\tilde{H}$ of strata including the HSA and the population of the region Friuli.

5. CONCLUSIONS

From the point of view of bias, the post-stratified ratio estimator (POS) is essentially unbiased in almost all the small areas. Furthermore the sample size dependent estimator (SD) has negligible values of the bias in almost all small areas. Synthetic (SYN) and composite (COM) estimators present bias values much higher than those of the other estimators.

From the point of view of efficiency, SYN and COM consistently have significantly lower RMSE compared to other alternatives. The estimator SD is much more efficient than POS and furthermore in four of the fourteen areas it shows RMSE values close to those of SYN and COM. Further, when considering the estimator COM there is the problem of the computation of optimum α. In practice only an estimated value of α may be used, resulting in a decrease in efficiency of this estimator. Thus considering both, bias and efficiency, the SD estimator would seem to be preferable to other estimators examined in the context of LFS in Friuli. The sampling rates in Friuli are relatively high and the magnitudes of relative biases and efficiencies of these estimators may be different in other regions where the sampling rates are low, *e.g.*, Piemonte and Lombardia.

REFERENCES

DREW, J.D., SINGH, M.P., and CHOUDHRY, G.H. (1982). Evaluation of small area techniques for the Canadian Labour Force Survey. *Survey Methodology*, 8, 17-47.

FABBRIS, L., RUSSO, A., and SANETTI, I. (1988). Storia e proposte in tema di campionamento a livello regionale, provinciale e sub-provinciale per l'indagine sulle forze di lavoro. *Rapporto di ricerca FOLA*, 4. Dipartimento di Scienze Statistiche, Universita' di Padova.

FALORSI, P.D., and RUSSO, A. (1987). Un metodo di stima sintetica per piccoli domini territoriali nelle indagini ISTAT sulle famiglie. *Atti del Convegno della Societa' Italiana di Statistica*, Perugia, Italia, 11-20.

FALORSI, P.D., and RUSSO, A. (1989). Un' analisi comparativa di alcune tecniche di stima per piccole aree per l'indagine sulle forze di lavoro. *Rapporto di ricerca FOLA*, 18. Dipartimento di Scienze Statistiche, Universita' di Padova.

FALORSI, P.D., and RUSSO, A. (1990). La stima dell'errore quadratico medio di alcune forme di stimatore sintetico nei campioni a due stadi utilizzati nelle indagini ISTAT sulle famiglie. *Giomate di studio: Classificazione ed analisi dei dati, metodi, software, applicazioni:*, Pescara, Italia, 27-39.

FALORSI, P.D., and RUSSO, A. (1991). Evaluation of small area estimation techniques for Italian Labour Force Survey. *Proceedings of the 1991 Annual Research Conference*. U.S. Bureau of the Census, 80-106.

GHANGURDE, P.D., and SINGH, M.P. (1978). Evaluation of efficiency of synthetic estimates. *Proceedings of the Social Statistics Section, American Statistical Association*, 52-61.

GONZALEZ, M.E., and HOZA, C. (1978). Small area estimation with application to unemployment and housing estimates. *Journal of the American Statistical Association*, 73, 7-15.

LEVY, P.S. (1979). Small area estimation synthetic and other procedures, 1968-1978. *Synthetic Estimates for Small Areas*, National Institute on Drug Abuse, Research Monograph, No. 24, U.S. Government Printing Office, Washington, D.C., 4-19.

PURCELL, N.J., and LINACRE, S. (1976). Techniques for the Estimation of Small Area Characteristics. Paper presented at the 3rd Australian Statistical Conference, Melbourne.

SÄRNDAL, C.-E., and HIDIROGLOU, M.A. (1989). Small domain estimation: A conditional analysis. *Journal of the American Statistical Association*, 84, 266-275.

SCHAIBLE, W.L. (1978). Choosing weights for composite estimators for small area statistics. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 741-746.

SCHAIBLE, W.L. (1979). A composite estimator for small area statistics. *Synthetic Estimates for Small Areas*, National Institute on Drug Abuse, Research Monograph, No. 24, U.S. Government Printing Office, Washington, D.C., 36-83.