

Evaluation of the impact of substitutions for unit nonresponse in the labour force survey of the municipality of Florence

Valutazione dell'effetto sostituzioni per non risposta completa nell'indagine sulle forze lavoro del comune di Firenze

Andrea Giommi (*), Emilia Rocco (*)

(*) Dipartimento di Statistica, Università degli Studi di Firenze

Riassunto: Il lavoro ha per oggetto l'analisi dell'effetto delle sostituzioni nell'indagine sulle forze di lavoro (FL) svolta dal Comune di Firenze. Dal 2002 il Comune di Firenze, a differenza dell'ISTAT, svolge l'indagine sulle FL seguendo un disegno basato sulla selezione di singoli individui anziché di famiglie. Il campione è stratificato per sesso età e zona di residenza e i non rispondenti sono sostituiti da individui appartenenti allo stesso strato. La pratica della sostituzione viene analizzata comparativamente ad alternative possibili basate sull'utilizzazione dei soli rispondenti.

Keywords: Unit non response in sample surveys, Substitution for non response, Calibrated estimators.

1. Introduction

Since 1995, the Municipality of Florence designed a LF survey, parallel to that of ISTAT, to cope with the unavailability, at municipality level, of accurate estimates for the main items studied in the survey. In a first phase, which goes from 1995 to 2001 the coordinator of the Municipal Statistical Office, simply chose for the enlargement of the ISTAT sample. Beginning from 2002, the same office thought it right to carry out an its own survey with a sampling design different from that of the ISTAT. The new survey, designed in collaboration with the Department of Statistics of the University of Florence, is based on a stratified proportional random sample of 1200 single persons for each quarter. Strata are defined on the basis of sex, age (six classes: 15 – 25, 26 – 35, 36 – 45, 46 – 55, 56 – 65, 66 – 75) and twenty residence areas. As far as unit nonresponse and unit rotation are concerned, the same strategy of ISTAT is followed, that is, substitution for nonresponse and 2-2-2 rotation. From 2004, after a first face to face interview, sample units have been reinterviewed by telephone and from the subsequent year all the interviews have been performed by CATI technique. The introduction of a complete CATI technique involved the reduction of strata from 240 to 60, substituting the residence areas with the five districts of the Municipality just in order to make easier the substitutions for nonresponse inside the same strata.

2. Substitutions vs. calibration adjustment for nonresponse

Both the use of the telephone file for sample selection and the substitution for unit nonresponse may produce some bias in the estimates. The portion of the bias due to the incompleteness of the telephone file with respect to the registry office list, which must

be considered the best frame for a survey on residents, is obviously worth to be studied. However in this work we focus on the effect of substitution for nonresponse which is not, as well known, the best tool for facing unit nonresponse and as well as the list problems can be responsible for a large part of the bias.

Here we compare estimates obtained by means of simple direct estimators, after replacing for nonresponse, with estimates which are based only on respondents after adjusting for nonresponse. We use adjustments for nonresponse that relate mainly to calibration estimation methods, with the addition of logical corrections which are allowed by the rotation structure of the survey. That is, when we estimate a parameter, such as the employment rate or the unemployment rate, we assume that a person who does not respond at a certain wave of the survey could maintain the same position reported in the previous wave, if this has been declared as a stable state.

The differences, often important, emerging in the comparisons between the two strategies of estimation must be discussed by the light of the diverse assumptions underlying the adjustment techniques.

Though the substitutions are executed in a very large number of strata (240 or 60) and the strata are defined on the basis of variables that seem strictly correlated with the main study variables as well as the response behaviour, they tend to introduce in the sample persons that are easier to contact with both face to face and telephone interview. The relationship between the higher availability and the employment position may be a source of bias. Moreover, we observed that generally the substitutions are not uniformly distributed on sex and age classes. Regarding age, for example, the most of substitutions fall in the “central” classes, which are typically those with the highest level of employment. Substitutes in these classes show a slightly lower rate of employment and for this reason the calibration adjustment estimates tend to increase the employment rate with respect to the direct estimates.

Deeper analyses are certainly needed in order to assess the best strategy but our first study shows that, at least for the Municipal LF survey, the use of calibrated estimators based on respondents, instead of substitutions, can produce both good estimates and savings in the collection of the information.

References

- Deville, J.C. and Särndal, C.E. (1992). Calibration estimators in survey sampling. *Journal of the American Statistical Association* 87, 376-382.
- Ekholm, A. and Laaksonen, S. (1991). Weighting via response modeling in the Finnish Household Budget Survey. *Journal of Official Statistics* 3, 325-337.
- Kalton, G. and Kasprzyk, D. (1986). The treatment of missing data. *Survey Methodology* 12, 1-16.
- Lundström, S. and Särndal, C.E. (1999). Calibration as a standard method for treatment of nonresponse. *Journal of Official Statistics*.
- Oh, H.L. and Scheuren, F.J. (1983). Weighting adjustment for unit nonresponse. In: W.G. Madow, I. Olkin and D.B. Rubin (eds.), *Incomplete Data in Sample Surveys*, Vol. 2. New York: Academic Press, 143-184.
- Särndal, C.E., Swensson, B. and Wretman, J.H. (1992). *Model Assisted Survey Sampling*. New York: Springer-Verlag.