BANCA D'ITALIA

EUROSISTEMA

Temi di discussione

(Working papers)

Accounting for sampling design in the SHIW

by Ivan Faiella

April 2008

Number 662

ACCOUNTING FOR SAMPLING DESIGN IN THE SHIW

by Ivan Faiella*

Abstract

This paper analyses how sampling design affects variance estimates and inference using the data collected by the Survey on Household Income and Wealth (SHIW). The SHIW combines three basic features: stratification, clustering, and weighting to correct for unequal probabilities of selection among sampling units. A model to assess variance is presented and a Jackknife Repeated Replication method is employed to estimate variance. Empirical evidence shows that: 1) simple random sampling formula for variance underestimates by a factor of between 3 and 2 the estimates that take into account all the design features; 2) the bias of unweighted estimates may be fairly substantial; 3) all these factors can seriously mislead inference based on SHIW data.

JEL Classification: C42.
Keywords: survey methods.

Contents

* Bank of Italy, Economic and Financial Statistics Department.

Acronyms

BR Bias ratio.

BRR Balanced repeated replications.

BRS Rescaling bootstrap.

Deff Design effect.

Deft Square root of the Deff.

Epsem Equal probability selection method.

fpc Finite population correction.

GVF Generalized variance function.

HT Horvitz-Thompson estimator.

JRR Jackknife repeated replications.

Meff Misspecification effect.

Meft Square root of the Meff.

MSE Mean square error.

NSRU Non-self-representing Unit.

PPS Probability proportional to size selection method.

PSU Primary sampling unit. Cluster of population elements.

RHO Rate of homogeneity. A measure of intra-cluster correlation.

RMSE Square root of the mean square error.

SECU Standard error computation unit.

SHIW Survey on Household Income and Wealth.

SRS Simple random sampling without replacement.

SRU Self-representing unit.

URS Uniform random sampling. Simple random sampling with replacement.

1 Introduction

Standard statistical analysis generally does not take sampling design into account, thus neglecting important features of survey data. Many statistical methods are developed assuming that sample information comes from a population model where the sampling scheme plays no role (Chambers and Skinner 2003).

In particular, sampling variability cannot be properly assessed under the hypothesis that survey variables are independently and identically distributed across all possible samples (i.e. they follow an *iid* process). This implies that survey data come from a sample obtained through Simple Random Sampling (SRS) with replacement (also known as Uniform Random Sampling - URS). Sample design typically involves specific techniques such as clustering and stratification which, if ignored, generally lead to inaccurate estimation of the variance.

Furthermore, when the process of sample selection and the response mechanism is *not ignorable*,[1] univariate and multivariate analyses disregarding survey weights can be biased.

The impact exerted by these three factors (clustering, stratification and weighting) on the variance estimates differs according to the estimator, but has some common features (Purdon and Pickering 2001):

- the effect of stratification is a reduction in sampling variance;
- the effect of clustering is an increase in sampling variance;
- the effect of weighting is generally a reduction in bias and an increase in sampling variance.

These features of sampling design must be taken into account in evaluating the variance of the estimator of interest. Straightforward formulas are available in sampling textbooks only for simple estimators, such as totals, and simple designs. As two seminal papers of Frankel (1971) and Kish and Frankel (1974) show, more difficult problems arise when the objective of the research is to assess the variance of non-linear estimators such as ratios, linear regression coefficients and order statistics in presence of complex designs.

[1] If the selection of the theoretical sample and the response mechanism that leads to the actual sample depend only on the observed data, the design is *ignorable*. More formally, given the definition of a ξ model to estimate a parameter θ, the concept of design ignorability implies that, under ξ model validity, the data collection process and response mechanism do not provide any additional information to estimate θ (for an analytical description see Chapter 7 of Gelman et al. 2003) .

Many of these problems, however, have lately found solutions shared by the statistical community (see Binder 1983; Rust, 1985; Wolter, 1985). Moreover, the increasing computational power of modern PCs and the availability of several survey software packages (e.g. R, Stata and SAS) greatly help the analyst to properly compute the variance of the estimators.

The objective of the present study is to assess how ignoring sampling design can affect the results of statistical inference regarding univariate and multivariate statistics, using data collected from the Survey on Household Income and Wealth (SHIW).

For this purpose, a flexible strategy to estimate sampling variance is required in order to assess properly the statistical reliability of a wide class of estimators. A model to assess variance is presented and the *Jackknife Repeated Replication* is suggested as a strategy for variance estimation. Empirical evidence shows that:

- SRS formula for variance underestimates by a factor of between 3 and 2 the estimates that take into account all the design features.
- In the majority of cases, the increase in bias associated with unweighted estimates is not compensated for by the decrease in the standard errors
- When one takes into account all the design features, the effective coverage probabilities of the estimators are lower than the nominal threshold.[2]
- For regression models, the increase in the standard errors that include information on sampling design can reverse the significance level of some coefficients.

The study is structured as follows. Section 2 describes the main features of complex surveys and their impact on inference. Section 3 introduces a strategy for variance estimation for SHIW data and evaluates the effect of disregarding sampling information in computing the confidence intervals of the estimators. Finally, the main conclusions are drawn.

2 Features of complex surveys

The purpose of sample surveys is to obtain an estimate of a population parameter from the sample. If the sample is a probability sample – i.e. if

[2]This implies that, from a statistical point of view, confidence intervals built using SRS formula for variance are not *valid* (Särndal, Swensson, and Wretman 1992).

all the population units have a non-zero probability to be included – and if it is *measurable*, it is possible to determine the sampling distribution of the estimator.[3] When households are the target population, for example because we want to study their average income, expenditure or wealth, information is typically collected through complex surveys.

In fact, the use of SRS is often unpractical and uneconomical: frequently there are no lists that exactly enumerate all the units of the population. Apart from the several problems afflicting such a list, if it exists at all, (blanks, duplications, etc., see Groves et al. 2004), it may not be economically sustainable to sample statistical units directly. Large-scale surveys are then generally based on *cluster sampling*. Under this design, a cluster of element is selected (*primary sampling unit* - PSU), and from the selected clusters a sub-sample of elements or of other clusters (in case of *multi-stage design*) is drawn. The PSUs are usually divided into mutually exclusive *strata* before the selection in order to reduce sampling variability and to control for certain domains of analysis.[4]

Schematically, complex surveys have some common features (Lohr 1999):

1. Mainly for cost-effectiveness they sample *clusters* of elements. For example, they sample municipalities or counties and then households or individuals pertaining to the cluster (*multi-stage sampling*).

2. They employ stratification of the clusters to mimic certain population sub-groups within the population.

3. Clustering, stratification and the response process implies that in the end statistical units are selected into the sample with unequal probabilities. Furthermore, unequal probabilities of selection can also reflect the response process or are implied by the design when clusters are not proportionally allocated to the strata, due to the oversampling of some sub-classes, or are sampled with probability proportional to cluster size (*PPS sampling*) in order to reduce sampling error. With these procedures the use of weight is required to obtain unbiased estimators.

[3] A sample is *measurable* if an unbiased estimate of the sampling variance can be derived from the variability observed between units within the sample. See Särndal et al (1992), Chapter 14.3.

[4] Intuitively, stratification reduces the possibility to draw samples that are "very different" from the survey population. The choice of how allocate the sample units to the strata is essential to maximize the gain of efficiency of this sampling strategy.

2.1 Variance estimation for complex sample survey data

The analysis of survey data involves the estimation of one or more population parameters (e.g. totals, means, regression coefficients, etc.). An estimator of the variance is required in order to assess the statistical reliability of the point estimates through the construction of confidence intervals.[5] [6]

Most of the statistical methods implies that survey data come from a sample obtained through URS.

Under URS design the variance of the sampling mean $\bar{y}$ can be estimated through:

$$var(\bar{y}) = \frac{1}{n}s^2. \tag{1}$$

According to (1) the sampling variance of an estimator depends upon two factors:

1. The sample size n;

2. The sampling element variance s^2.

As previously mentioned, however, samples are selected from finite populations and SRS is almost never used.

Neglecting the finite population correction[7] has often a small impact on the estimates because in large sample surveys - such as household surveys - the sample is a negligible fraction of the population.

The impact of implementing a complex survey design is potentially greater. In particular, the consequence of sampling clusters instead of elements is an increase in the variance of the estimators for two main reasons:

1. the number of independent choices is reduced because the randomization process is based on clusters and not on elements (with a loss of degrees of freedom);

[5] The importance of the subject for official statistics is confirmed by a Eurostat publication focusing on variance estimation strategies (Eurostat 2002).

[6] In the presence of *item non-response*, a further variance component to be assessed, not considered in the present study, is the one induced by the process of imputation. See for example Särndal and Lundström (2005), Chapter 13.

[7] The *finite population correction* is the reduction of sampling variance due to the fact that sampling is *without replacement*. (1) is then multiplied by $\left(1 - \frac{n}{N}\right)$. Not considering this factor greatly simplifies variance formula and produces more conservative variance estimates (i.e. positively biased).

2. the distribution of the elements in each cluster is not at random because, within a cluster, environmental and socio-economic factors induce an aggregation of elements with similar characteristics. This positive intra-cluster correlation (also popularized as *rate of homogeneity* – *rho* – by Kish, 1965) reduces the effective size of the sample because acquiring data on more elements of the same cluster does not increase proportionally the amount of information on the object of study.

Note that in a *multi-stage* sampling framework the *sampling error computing units* (SECUs) are the PSUs only. These are known as *ultimate clusters* and, in the presence of small sampling fractions in the subsequent stages, they provide all the necessary information about the variance of the sampling process.[8]

To begin with, computing the variance requires a variance model. This means that we have to include all the features of the selection process in the relevant formula (i.e. stratum and PSU IDs). Then a method, possibly with optimality properties, can be applied as a feasible way to obtain variance estimates.

2.1.1 Strategies to estimate sampling variance

To obtain estimates for the variance of estimators in a complex survey framework, two broad strategies are suggested: the *linearization method* and the *replication-based method*.

According to the former, non-linear estimators are expressed as a function of linear estimators (such as totals) and the delta method is then applied to obtain a biased but consistent estimate of the variance.[9] This method has the advantage of providing an analytical formula for the variance, but it requires computation of numerical derivatives (for the first order Taylor expansion) and rests on the hypothesis that higher order terms are negligible.[10]

[8]“ [...] from the values of the primary variates y_α alone we can compute the entire variance, including both the between-cluster and within-cluster components of variation.” (Kish, 1965, p.158).

[9]Both the bias and the variance are characteristics of an estimator and not of one of its possible realizations (i.e. the estimate). In what follows I adhere to the interpretation of Särndal et al (1992), p. 41, indicating for biased estimate “an estimate calculated from an estimator that is biased”.

[10]As pointed out by Wolter, a statement about the order of the “remainders” is possible only in the context of infinite populations. For finite populations an evaluation of the order of higher Taylor terms is possible only under the hypothesis of a superpopulation model (Wolter 1985).

Replication methods rest on the idea of selecting k samples from the population (*replicates*) by sub-sampling the original sample and computing the estimates of interest for each sub-sample. The variance can then be computed as a measure of the deviance of the *k-th* estimate from the estimate for the original sample (an average of the estimates across the k replicates or the value of the estimate for the whole sample). Among the random group methods, two are usually adopted to estimate variance within complex surveys: the *Jackknife Repeated Replications* (JRR) and the *Balanced Repeated Replications* (BRR)(Rust 1985).

Table 1. Replication methods for complex surveys

	JRR	BRR
Step1	Delete a PSU from the first stratum: this is equivalent to zero-weighting it	Delete a PSU from strata according to a Hadamard matrix: this is equivalent to zero-weighting the deleted PSU
Step2	Increase the weight of the other PSUs in the stratum by (PSU in the stratum)/(PSU in the stratum - 1)	Increase the weight of the other PSUs in the stratum by 2
Step3	Compute the statistic of interest with the set of replicated weights	
Step4	Compute the deviance (sum of squared distances) from each replicate statistic (step 3) from the total sample statistic	
Step5	Multiply the deviance by the factor (replicates-1)/(replicates)	Divide the deviance by the number of replicates

The difference between the two techniques lies in how sub-samples are drawn. The BRR implies that, for variance computation, there are two PSUs per stratum; in each replication a cluster is selected from a given stratum according to a pattern that ensures "balanced replicates" using a Hadamard matrix (for details see Wolter, 1985, Chapter 3).[11] The JRR uses a drop-out procedure: in each replication it drops a cluster from a given stratum, appropriately re-weighting the remaining clusters in the same stratum, a method also known as JKn jackknife.

Another replication method widely used in practice is the *Bootstrap*. This method needs to be properly modified to deal with survey data (*Rescaling*

[11]A Hadamard matrix has columns pairwise orthogonally and gives a pattern such that the variance estimated with the BRR has equal asymptotic properties to the same estimator evaluated over all the possible $2^{\#ofstrata}$ samples.

Bootstrap - BRS) (Rao and Wu 1984). The method is not treated here because it does not possess good empirical properties. In a simulation study comparing linearization and the replication methods (JRR, BRR and two modified Bootstrap methods), Shao and Tu (1995) conclude that

> *Overall, the linearization and the jackknife variance estimators have the best performance; the BRR variance estimator is the second best; and the BRS variance estimators are the worst [p.252] [...] Furthermore, the bootstrap variance estimator usually requires more computations than the jackknife or the BRR [p.281].*

The linearization and replication methods present similar features: they provide biased estimates of the variance, the order of the bias being usually negligible for large samples (the bias of replication methods depends on the number of replications), but the Taylor expansion method generally shows a lower mean square error (MSE); on the other hand, replication methods have a better performance in terms of confidence intervals and coverage probabilities (Wolter 1985, Chapter 8). Since both methods produce basically the same results, they can be used interchangeably according to the nature of the estimator[12] and the software procedures available.[13]

An indirect way to assess the variance of an estimator using complex survey data is the Generalized Variance Function approach (GVF). It consists in estimating the variance through an analytical relation that links the expected value of the estimator with its variance. This relation can be estimated using data on previous surveys or it can be computed for a few key variables of the actual survey and then applied to other statistics presenting "similar characteristics" (i.e. statistics that follow a common model).

Although the GVF does not possess well-behaved asymptotic properties (as pointed out by Wolter, 1985, it can involve non-normal confidence intervals), a class of these models can be extremely useful to provide a rough idea of how the inference based on standard variance formulas (i.e. assuming URS sampling) changes considering the complexity of sampling design. This is known as the design effect (*Deff*) model (Kish, 1965, Chapter 5). Given the estimator $\hat{\theta}$ of the population parameter θ the design effect is the ratio

[12] An important difference concerns the estimate of the variance of a domain estimator. In each domain the sampling design must be preserved for this estimate to be valid. This is most easily accomplished by replication methods (the JRR in particular).

[13] A comprehensive review of the available software can be find in Chantala (2003).

between its variance computed according to the complex survey design and that computed under the hypothesis of SRS.

$$Deff(\hat{\theta}) = \frac{var(\hat{\theta})_{compl}}{var(\hat{\theta})_{srs}}. \tag{2}$$

These relations can be used to adjust the SRS variance estimates and the critical level of the confidence intervals and to compute the effective sample size (i.e. the sample size needed to achieve the same level of precision as using SRS). [14] This formula is useful because it is "portable". For example, it can be used to approximate correct standard errors in domains of study. When the analysis focuses on sub-classes (such as gender, geographical area, etc.) the total sample *Deff* for the mean of the $g - th$ domain $\bar{y_g}$ can be adjusted for the relative cluster size with the following formula:

$$Deff(\overline{y}_g) = 1 + \frac{n_g}{a} * rho, \tag{3}$$

where n_g is the domain size, the intracluster correlation *rho* is estimated using the $Deff(\overline{y})$ for the total sample as $(deff - 1)/[(n/a) - 1]$, n is the total sample size and a is the number of clusters. Skinner et al. (1989) show that (3) can easily be approximated as:

$$Deff(\overline{y}_g) = 1 + \frac{n_g}{n} * [Deff(\overline{y}) - 1], \tag{4}$$

where the only information necessary to compute $Deff(\overline{y_g})$ is $Deff(\overline{y})$ and the proportion of the sample in the g domain $\left(\frac{n_g}{n}\right)$.

The design effect appraises the impact of sampling design on the variance of the true estimator $\hat{\theta}$. Skinner et al. (1989, pp. 24-31) focus instead on the effect of the design on the estimator of variance. According to these authors it is appropriate to compare the complex variance not with SRS variance (the unbiased variance under SRS) but with a biased estimator of variance obtained ignoring all the design features, i.e. weights, clustering and stratification. This measure is usually known as the misspecification effect (*Meff*):

$$Meff(\hat{\theta}) = \frac{var(\hat{\theta})_{compl}}{var(\hat{\theta})_{unw}}, \tag{5}$$

[14] For example, given a measure of the intracluster correlation,*rho*, it is possible to derive a series of relations linking *Deff* and the average cluster size (Kish, 1965). In particular, for the sampling mean $Deff = 1+(b-1)*rho$, where b=average cluster size, *rho*=intracluster correlation.

where the numerator is the same as in *Deff*, while the denominator is the unweighted variance of $\hat{\theta}$ under the hypothesis of SRS (Eltinge and Sribney 1996).

Finally, to estimate the variance of population quantiles where the JRR performs poorly (see Kovar, Rao, and Wu 1988; Wolter, 1985), it is possible to employ an indirect method (Woodruff 1952). This procedure requires the following steps:

1. estimate y_{50}, the median value of y;

2. define an indicator variable P_{y_i} equal to 1 when $y_i < y_{50}$ and 0 otherwise and compute the statistics $P_{y_{50}} = \frac{\sum_{i=1}^{n} w_i P_{y_i}}{\sum_{i=1}^{n} w_i}$;

3. estimate the standard error of $P_{y_{50}}$ and derive the associated boundaries for a given confidence interval;

4. compute the empirical value of the cumulative distribution and take the inverse corresponding to the boundaries of $P_{y_{50}}$ calculated in step 3. These are the boundaries of y_{50};

5. given a confidence level and the associated *z-score* the standard error of the median will be equal to the width of the confidence interval (the difference between the boundaries computed in step 4) divided by twice the *z-score*.[15]

2.1.2 Inference with complex survey data

The concepts of the previous section can be used to evaluate the impact of disregarding the sample design on the inference from sample data, computing the effective coverage probabilities of the (1-α)% probability statements. To correct for the bias of the unweighted estimates it can be useful to refer to the concept of bias ratio, i.e. the bias of the estimator normalized with its standard error.

$$BR(\hat{\theta}) = \frac{E(\hat{\theta}) - \theta}{\sqrt{V(\hat{\theta})}} \tag{6}$$

In fact, given the standard normal *z-score* Z, it can be shown that the effective probabilities are given by

$$P\left\{-z_{1-\alpha/2} - BR(\hat{\theta}) < Z < z_{1-\alpha/2} - BR(\hat{\theta})\right\}. \tag{7}$$

[15] Francisco and Fuller somehow refined this method using the inversion of a robust score test (Francisco and Fuller 1991).

The coverage probabilities are therefore computed by taking the difference of two cumulative standard normal curves evaluated at the critical value $\left(z_{1-\alpha/2}\right)$ "netted" by the bias ratio (Särndal, Swensson, and Wretman 1992). If no information on the design is considered, the modified *z-score* is divided by the square root of *Meff* (also called *Meft*). When survey weights are used and variance is estimated using the SRS, the normal score can be corrected dividing it by the *Deff* square root (a measure called *Deft*), so that $z\prime = z/\sqrt{Deff} = z/Deft$ (see Skinner et al., 1989, p. 30).

Note that, to compute confidence intervals, a standard normal distribution is only an approximation of the effective distribution of the estimators. In the case of an estimator with unknown variance, a t statistic should be used. Making use of the t distribution can be a conservative approach, particularly when there are few degrees of freedom for variance computation. Furthermore, in computing degrees of freedom the appropriate level of randomization should be taken into consideration: for example, in presence of a stratified cluster sample, the degrees of freedom equal to the number of clusters - number of strata.[16] If a replication technique is adopted, the degrees of freedom are related to the number of replications (see Skinner et al., 1989, pp. 56-57).

2.2 The role of survey weights

If, according to the selection process, all the elements of the population have an equal chance to be included in the sample, the design is known as *epsem* (equal probability selection method). If the design is *epsem* and the sample size is fixed, given the sampling scheme, then the unweighted estimator of a mean (or a total) is unbiased (i.e. its expected value is equal to the true population value). In practice, units included in the sample have unequal probabilities of selection. To correct for this possible source of bias, each observation is weighted using the inverse of the sampling fraction (sampling weight) adjusted for the non-response mechanism (non-response weight) and often incorporates auxiliary information about the population (post-stratification).

The survey weight w_i is used to obtain unbiased estimates from the sample through the Horvitz-Thompson (HT) estimator, also known as the π-estimator (Särndal, Swensson, and Wretman 1992). The rationale of this class of estimators is to inflate each sample observation y_i dividing it by its

[16] The degrees of freedom are determined by the randomization process that involves the random selection of clusters within each H strata.

inclusion probability π_i. For example, the survey-weighted sample mean is an approximately unbiased estimator of the population mean given a fixed sample size.

$$\overline{y} = \frac{\sum_{i=1}^{n} w_i y_i}{\sum_{i=1}^{n} w_i}, w_i \propto \frac{1}{\pi_i} \tag{8}$$

The estimator (8) is technically biased (because it is a ratio of two random variables) but its relative bias is bounded by the coefficient of variation of the denominator (the sample total of survey weights), usually very small for large samples (Kish,1965, Chapter 2).

Note that using survey weights implies that the weighted mean is not a linear estimator and the sampling variance of (8) under SRS is not correctly estimated by (1) but we must resort to a formula for the variance of two random variables.[17] If we indicate with W and Y the estimators of the sample totals of the numerator and the denominator of (8), its variance can be estimated by:

$$var\left(\overline{y}\right) = \frac{1}{W^2}\left[var(Y) + \overline{y}^2 var(W) - 2\overline{y}cov(Y,W)\right]. \tag{9}$$

The unbiasedness of the π-estimator is not without costs: usually, there is a loss of efficiency, L_w, due to the use of weights. In the case of the estimate of the variance of the mean and complete incorrelation between y_i and w_i, this loss is proportional to the squared coefficient of variation of the weights (Kish 1992).

$$L_w = \left(1 + CV_w^2\right), \text{where } CV_w = \frac{s_w}{\bar{w}} \tag{10}$$

In fact, the unweighted estimators can be viewed as weighted estimators whose attached weights are all constant, thereby unaffecting the variance (note that in this case (9) reduces to (1)). On the contrary, the π-estimator of a mean is a cross-product of the study variable and the relative weight. If this cross-product does not vary much across observations, this can limit the loss of efficiency due to weighting.[18]

A possibility to reduce the impact of weights on the variance of the estimator is to make use of the *Generalized Regression Estimator* (GREG). The idea behind the GREG is to supplement the HT estimator with a set of auxiliary information (in a multivariate context) correlated to the study variable

[17] Some authors refer to (8) as a *ratio mean*. See (Kish 1965).

[18] This reason is the rationale of PPS sampling as a tool to limit the variance of the estimators. See Särndal et al. (1992), Chapter 3.6.

(Särndal and Lundström 2005). The HT estimator is corrected using the gap between the sample estimate and information on the value of the auxiliary vector in the population (available from larger surveys or census data). In general, the aim of calibration is to increase precision as it serves both to reduce differences between the sample and the population distribution with respect to some auxiliary variables and to reduce the variability of weights (in the case of post-stratification this happens only if the post-strata have a smaller within-stratum variance with respect to the overall variance).

Purdon and Pickering (2001), analysing the Workplace Employee Relations Survey, assess whether the increase in bias associated with unweighted estimates is compensated for by the decrease in the standard error, estimating the mean square error (i.e. the square of the bias plus the design-based variance). With the exception of a variable, all the figures are smaller for the weighted estimate than for the unweighted estimate. This suggests that weighted estimates, even with their increased standard errors, are almost always preferable to the unweighted estimates. A replication of this analysis on SHIW data confirms these results (Faiella and Gambacorta 2007).

When evaluating confidence intervals using weighted estimates it is always safer to check the degrees of freedom. In fact, following a standard practice, survey weights are considered an expansion factor summing up to the total survey population. Therefore, it is very important to compute appropriately the denominator of the variance estimates, otherwise the number of degrees of freedom has the same order of magnitude as the population, pushing the probability to reject the null hypothesis towards one. A solution is to properly specify degrees of freedom for variance computation or to normalize survey weights by re-scaling the original weights to their mean so they have mean equal to 1 and sum up to the sample size.[19]

In conclusion, it is established practice in the statistical community to use weights in the analysis of survey data. If some portions of the population are oversampled and these factors are somehow related to the analysis variables, then point estimates taken from unweighted analyses could be seriously biased.

2.2.1 A short digression on the use of weights in regression analysis

Often survey data users are interested in using multivariate analysis, such as regression techniques, to assess the association between the study variables

[19]Even if the majority of the statistical packages (Stata, SAS, R,..) handle the problem correctly, it is always better to check for the sum of weights when performing inference using survey weights.

and sets of sample information used as controls. In a *design-based* framework this involves the use of weights to obtain design-unbiased estimates[20] and a measure of the variance of those estimates that correctly includes design features. Kott shows that weighted estimates are more robust to omitted variable problems and to the heteroskedasticity that normally characterizes sample survey data (Kott 1991).

In a *model-based* framework, under the hypothesis that the model is correctly specified, using survey weights in regression analysis involve a loss of efficiency. In the case of regression analysis, the effect of the design is expected to be limited because of the controls used (especially if design variables are included among the covariates). Nathan and Smith (1989) show that unless the selection of the sampling units is *ignorable* (see footnote 1), conditional on the covariates of the model, OLS estimates are biased and inconsistent. Note that this selection pattern depends both on the actual sampling scheme (i.e. how the population elements are included the sample) and on the response process. When this information is not relevant for the model a condition of *design ignorability* is met.

Särndal et al. (1992), criticize a purely *model-based* approach, where design unbiasedness is neglected, underlining that the parameters estimated using sampling weights are more robust because they are model-unbiased if the model is true, and design-consistent if it is not (on this point see also Little 1981).

Entering the debate between *model-based* and *design-based* approach to regression analysis is outside the scope of this paper. Still, it is possible to refer to the conclusions of Hansen, Meadow and Tepping (1983) and Lohr (1999) suggesting use all the design features in regression models when sample size is large and the sample size helps to mitigate the possible loss of efficiency due to survey weights. When the sample size is small, the design feature can be neglected, but it is anyway advisable to check the consistency of these results with those including sampling information and to make use of a robust measure of variance.

[20]Non-linear estimators such as the regression or the ratio estimator are only approximately unbiased, presenting a bias of order $n - 1$, negligible in large samples (Wolter 1985).

3 The Survey on Household Income and Wealth (SHIW)

The SHIW[21] has been conducted by the Bank of Italy since 1965 to collect information on the economic behaviour of Italian households and specifically to measure the income and wealth components. The main objective is to estimate how these are distributed across Italian households.

The basic statistical unit is the household, defined as a group of individuals linked by ties of blood, marriage or affection, sharing the same dwelling and pooling all or part of their incomes. Institutionalized population is not included. The sample comprises about 8,000 households. In this paper we used data from the 2002 wave of the SHIW (8,011 households and 21,148 individuals).

The sample for the survey is drawn in two stages (municipalities and households), with the stratification of the primary sampling units (municipalities) by region and demographic size.

Until 1987 the survey was conducted with time-independent samples (cross sections) of households. Since 1989 part of the sample has comprised households interviewed in previous surveys (panel households) in order to facilitate the analysis of changes in the phenomena being investigated. This design is known as a *split panel survey* (Kish 1987) and has the advantage that it is flexible in providing both cross-sectional and longitudinal measures (Duncan and Kalton 1987).

Data are collected by means of personal interviews conducted by professionally trained interviewers and using computer-assisted devices (computer assisted personal interviewing). Data collection is entrusted to a specialized company and the interview stage is preceded by a series of meetings at which officials from the Bank of Italy and representatives of the company give instructions directly to the interviewers. The households contacted for interviews, who are guaranteed complete anonymity, receive a booklet describing the purpose of the survey and giving a number of examples of the ways in which the data are used. The participating households may request a copy of the results of a previous survey.

The core sections of the questionnaire remain basically unchanged. In order to reduce the response burden, two monographic sections, administered to a random subset of the sample, are added in each wave.

Microdata, documentation and publications (in Italian and English) can

[21]Further details on the SHIW are given in Faiella et al. (2006).

be downloaded free of charge from the Bank of Italy's website.[22]

3.1 Variance estimation in the SHIW

As mentioned, the SHIW design uses a two-stage stratified sample with the stratification of the primary sampling units (municipalities) by region and demographic size. Within each stratum, the municipalities are selected by including all municipalities with a population of more than 40,000 inhabitants (SRUs) and by random selection of smaller towns (NSRUs) with probability proportional to the resident population. The individual households to be interviewed are then randomly selected.

In 2002, the original design involved 50 primary stage strata with 344 PSUs (175 SRUs and 169 NSRUs). The panel municipalities that are home to households that have taken part in at least two surveys are placed in a separate stratum because they make no contribution to random variability in the first stage. Thus, including these municipalities we end with 307 SRUs. Because these municipalities enter the sample with certainty, they are assigned a self-representing stratum in the primary stage of selection (Särndal et al. 1992, pp. 137-138).[23] Within each SRU the elements are then randomized to form two SECUs (also called "pseudo" PSUs - PPSUs). The remaining 37 NSRUs, after collapsing the adjacent strata with just one PSU, are combined in their stratum to form a pair of SECUs per stratum for a total of 26 NSRUs. This set-up is also known as paired selections of clusters (Kish, 1965), a design used to gain the most from stratification and to simplify variance formulas.

As shown in Table 2, the chosen variance model departs from the original design by somehow "averaging" the different intra-cluster coefficients in each PSU, but in doing so it increases the average cluster size providing more stable variance estimates.[24]

In variance computations finite population correction is not considered, thus producing slightly conservative estimates, and the JRR is used to produce standard errors for the SHIW.[25] The reasons for preferring replication techniques to linearization methods are twofold.

[22] `www.bancaditalia.it/statistiche/indcamp/bilfait`.

[23] In this way they will not contribute to between-cluster variance.

[24] This choice aims to limit the "variance" of the variance estimator (the second order variance).

[25] This method is also known as a particular form of delete-a-group jackknife with two PSU per stratum (JK2).

Table 2. A model for variance estimation in the SHIW (2002)

	Strata (h)	PSU (a_h)	Sample size (n_{a_h})	Average PSU size
SRUs	307	614	7,194	12
NSRUs	13	26	817	31
Total	320	640	8,011	13

Linearization methods entail a high degree of complexity in defining the variance estimators of complex statistics and present a limited possibility of accounting for weighting adjustment (such as post-stratification). Furthermore, to allow survey data users to apply this method it is necessary to provide them with the design variables (i.e. stratum and cluster IDs). These are often related to geographical information (e.g. the IDs of the municipalities) that is usually not disseminated in the public datasets due to confidentiality. Large-scale surveys then disseminate *replication weights*, thus avoiding the inclusion of design variables in the public datasets to balance confidentiality protection and the users' possibility to properly compute variance.[26]

In order to generate replication weights the sample is broken up into subsamples, called replicates. The estimate of interest is calculated from both the full sample and from each replicate and the deviance of each replicate-based estimate from the full-sample estimate (or from the average of the replicates) is used to derive the variance of the estimator. These weights are produced according to the complex survey plan and to the chosen variance estimation techniques (such as the BRR, JRR or *Bootstrap*).

The user provided with such a set of replication weights can easily compute the variance of very complex non-linear estimators and can often rely on statistical packages where it is possible to load the replication weights that are then automatically used in computing the variance of a variety of estimators.[27] In the case of the SHIW, following the variance computation model design devised and applying the JRR method,[28] the replication

[26]This is the strategy adopted by many large scale US surveys such as the National Comorbidity Survey (`www.hcp.med.harvard.edu/ncs`), the Survey of Income and Program Participation (`www.sipp.census.gov/sipp/index.html`) and the Survey of Consumer Finances (`www.federalreserve.gov/pubs/oss/oss2/scfindex.html`).

[27]This possibility is straightforward with packages such STATA and also with software available on the public domain (R `www.r-project.org`, WESVAR `www.westat.com`, VPLX `www.census.gov/sdms/www/vwelcome.html`).

[28]The JRR was preferred over BRR and *Bootstrap* because it does not need the cumbersome computation of an Hadamard matrix - like the BRR - and it possesses good

weights are computed as follows. For each of the 320 "pseudo" strata (h) with two "pseudo" PSU (j) the replicated weight is:

$$w^r_{hji} = \begin{cases} 0, & \text{if the } i-th \text{ unit belongs to PSU 1} \\ w_{h2i} \frac{\sum_{j=1}^{2} \sum_{i=1}^{n_j} w_{hji}}{\sum_{j=1}^{2} \sum_{i=1}^{n_j} w_{hji} - \sum_{i=1}^{n_1} w_{h1i}}, & \text{if the } i-th \text{ unit belongs to PSU 2} \\ w_{hji}, & \text{for all the other units.} \end{cases}$$

The JRR variance is finally calculated using the following steps:

1. the number R of replications is equal to the number of "pseudo" strata,[29] $R = \sum_h (a_h - 1) = 320$;

2. in each replicate the weight of the first "pseudo" primary sampling unit is set equal to zero and the sampling weight of the other is raised by a factor equal to the weight of the cancelled unit on the total weight in the stratum (see above);

3. this weight is used to calculate, for each replicate, the relevant estimators $\hat{\theta}_{(r)}$;

4. since the design for variance estimation contains two units per stratum, the estimate of the standard error is calculated as the square root of the sum of the square deviations of the estimate of the replications from the estimate on the total sample $\hat{\theta}$ (see Kish and Frankel 1974). [30]

$$stderr_{JRR} = \sqrt{\sum_{r=1}^{R} (\hat{\theta}_{(r)} - \hat{\theta})^2} \tag{11}$$

3.2 SHIW weights

Faiella and Gambacorta (2007) provide a detailed description of the SHIW weighting process and its impact on the estimates. Their main findings are that the increasing variability induced by using weighted estimators is

empirical properties - unlike the *Bootstrap*.

[29] These are called "pseudo" strata because the original strata, determined by the sampling design, are modified according to the model used to compute sampling variance.

[30] Note that this variance estimator (v_4 in the Wolter classification) is also robust to the possible bias of the replicates, and is also known as an MSE formula for variance estimation(Wolter 1985).

compensated for by the bias reduction even when performing analysis on sample domains.[31]

3.3 Design and misspecification effects for some influential statistics in the SHIW

In this section the concepts analysed previously are applied to assess the effect of not considering sampling design on inference. A number of univariate and multivariate statistics computed on SHIW 2002 data are considered.

The design and misspecification effects are evaluated under three different scenarios:

1. that the variables were collected with an equal probability SRS, i.e. estimators and their variance do not include survey weights and SRS formula for variance is applied (UNW);

2. that the variables were collected with an unequal probability SRS, i.e. estimators and their variance use survey weights coupled with SRS formula (SRS);

3. that the variables were collected according to the actual sampling plan, i.e. the estimates include both survey weights and use the correct estimation formula for variance (COMPL).

To clarify the hypotheses under these different scenarios in Table 3 we report the formulas used to estimate the sample mean and its variance.

Table 3. Sampling mean and its variance under three different scenarios

Scenarios	Sampling mean - $\bar{y}$	Variance of the sampling mean - $var\,(\bar{y})$
UNW	$\bar{y}_{unw} = \frac{\sum_{i=1}^{n} y_i}{n}$	$var(\bar{y}_{unw}) = \frac{\sum_{i=1}^{n}(y_i-\bar{y}_{unw})^2}{n}$
SRS	$\bar{y}_{srs} = \frac{\sum_{i=1}^{n} w_i y_i}{\sum_{i=1}^{n} w_i}$	$var(\bar{y}_{srs}) = \frac{\sum_{i=1}^{n} w_i (y_i-\bar{y}_{srs})^2}{\sum_{i=1}^{n} w_i}$
COMPL$^{(1)}$	$\bar{y}_{compl} = \frac{\sum_{i=1}^{n} w_i y_i}{\sum_{i=1}^{n} w_i}$	$var(\bar{y}_{compl}) = \sum_{r=1}^{R} (\bar{y}_{(r)} - \bar{y}_{compl})^2$

$^{(1)}\bar{y}_{(r)}$ is computed for each r, using the set of replicated weights computed as described in Section 3.1.

[31] The study also proposes for the first time a set of longitudinal weights for the SHIW, i.e. weights that, giving their enhanced description of the "panel population", are better suited to perform longitudinal analysis.

3.3.1 Univariate analysis

Producing their survey reports, sample survey statisticians mainly focus on univariate statistics (means, ratios, proportions or quantiles). In Table A1 means and proportions for a subset of SHIW variables are reported. These 15 variables are household income, expenditure and wealth (A-C), age of the head of household and of household members (D and H), household size (F) and the proportion of people having certain characteristics (according to gender, working status and education) (G-Q). The table reports the mean, the standard error of the mean and the 95% confidence limit for each of these variables under the three hypotheses previously illustrated (UNW, SRS, COMPL).

Table A2 contains information on *Deffs*, *Meffs* and other measures of the statistical reliability of the estimates. In the first column the bias of the unweighted estimates (approximated by the difference between weighted and unweighted estimates) is reported and in the second column the bias is normalized with the standard error of the estimates (the *Bias Ratio*). A measure of the empirical counterpart of bias is computed, following Kish (1992, p.191), as the difference between the unweighted and the weighted estimate.[32] The bias can be fairly large ranging in absolute value from 1 to 14 times the standard error, according to the measured variable.

In columns 3 and 4, the reported *Meff* and the *Deff* present an average value of 3.2, thus indicating that not taking into account sampling weights and other design features leads to downward biased variance estimates. In the next three columns for each design the *Root Mean Square Error* (RMSE) is computed (the square root of the variance for SRS and COMPL and of the sum of the squared bias and the variance for UNW). As previously mentioned, these results confirm those of Purdon and Pickering (2001) and show that, in general, the increase in bias associated with unweighted estimates is not compensated for by their lower standard errors.

Finally, in the last three columns the 95% probability statements are confronted with the effective coverage probabilities when one takes into account the inflation factor due to the design variance and to the bias of the unweighted estimates. The probabilities under UNW design are adjusted using the $BR\left(\widehat{\theta}\right)$ and the *Meft* ($\sqrt{Meff}$) while those under SRS are adjusted dividing the *z-score* by the *Deft* ($\sqrt{Deff}$).

For example, rather than being included in the standard error range for

[32] A refinement of this approach, based on Little et al. (1997), is to correct for the possible "overestimation" of the empirical measure of the squared bias. This approach is used extensively in Faiella and Gambacorta (2007).

95% of the possible samples, the mean of income is included in the z*standard error range in 64% under UNW design and in 92% under SRS design. The same percentages are respectively 73% and 93% for wealth and household expenditure.

The median is a robust estimator in the presence of skewed distributions. This statistic is widely used (with other quantiles) to analyse income and wealth distribution (see for example Faiella and Neri 2004, and Brandolini et al. 2004). In order to have an alternative to the JRR, which performs poorly with order statistics, an estimate of the variance of the median is obtained by the procedure due to Woodruff described earlier (Woodruff 1952). The results of this procedure are reported in Table A3. The standard errors of the median are, as expected, lower than those of the mean for very skewed variables such as household wealth. Moreover, for this statistic the effect of the design heavily influences variance estimates. The *Deff* ranges from 1 for expenditure and income to 1.28 for wealth. The *Meff* goes from 1.69 in the case of income to 2.39 for wealth.[33]

Another measure often used in income and wealth distribution analysis is the Gini index. The standard error of this index can be approximated using an asymptotic formula (Cowell 1989). In the hypothesis that the *Deff* formula for the mean is also valid for this measure of concentration, its variance can be adjusted to account for complex design using the simple GVF[34]

$$Std.err(Gini)_{compl} = Std.err(Gini)_{asymptotic} * Deft(Mean). \tag{12}$$

Suppose we want to adjust the asymptotic standard error of the Gini of household income to account for complex design. Considering the square root of the *Deff* for the mean income in Table A2 and using it as an adjustment factor, the asymptotic standard error is increased by 38% ($\sqrt{1.893}$).

This procedure is straightforward to apply but must be carefully evaluated. For example, in the case of the Gini index, the estimate is 0.357, the JRR standard error is 0.0042 (with a coefficient of variation - CV - of

[33] A study by Hansen and Tepping (1985) points out that when the sample size is large and the sample is a multistage cluster design, then the JRR method produces satisfactory estimates of variances for medians. With SHIW data, the use of the JRR to estimate median standard errors (row COMPL2 in Table A3) seems to confirm the instability of this method in appraising the standard error of order statistics.

[34] The same procedure can be applied for other variables (such as wealth, financial assets, expenditure) and other highly non-linear statistics to improve the reliability of survey-based inference.

1.2%), the asymptotic standard error is 0.0058 (with a CV of 1.6%) and after the correction through the Deft it is 0.008 (with a CV of 2.2%). These differences probably originate from asymptotic standard error formula that are too conservative or from the fact that the *Deff* formula, derived for the sample mean, is not really portable for very different statistics (such as the Gini index). This example is useful as a reminder that the GVF is only an imperfect tool to give a rough idea of the variance of an estimator, but it cannot replace a proper variance model.

3.3.2 Multivariate analysis

In a preceding section we briefly mentioned that in a multivariate context there is no agreement on the need to incorporate survey information in the analysis. This section measures the consequences of disregarding this information in regression analysis using very simple examples. As the purpose of the comparison is only to illustrate the effect of the complex design on variance estimates, the models presented are simple and goodness-of-fit measures are not reported.

In Tables A4-A7 the outcomes of a linear and a logistic regression are presented.

Table A4 shows the results of a linear regression for estimating an earning function model. The log of the labour income is regressed on a set of covariates, including age, experience, education, marital status, gender and household size. Table A5 reports, as in the previous section, the bias (absolute and relative), the Meffs and the Deffs and the coverage probabilities.

The effect of bias seems to be smaller than in the univariate analysis, ranging from 2 to 0.06. Both the Meffs and the Deffs appear to be pretty high, with values of around 3. This increase in the standard error of estimates can reverse the significance level of some coefficients: the "age" coefficient, significant at the 5% level under the UNW design, is not different from zero in the other two set-ups; the t statistic of "experience" is almost halved once one accounts for design features. Coverage probabilities of the nominal 95% confidence level can be as low as 50% for the UNW and average around 89% for SRS.

Table A6 reports the estimates of a logistic model used to predict the probability of participating in the labour force for individuals aged between 14 and 65. A labour force indicator is the response variable and the set of predictors is age, education, marital status, gender, number of income earners and of dependants (children less than 18 years old). Table A7 contains the same information as Table A5 (bias, Meffs, Deffs, coverage probabilities).

For the logistic model the relative bias of the unweighted estimates ranges from 3.6 to 0.005. Both the Meffs and Deffs are between 2 and 3. The effect of the number of dependents is significant at the 3% level under the UNW design, while it is significant at the 6% level using complex design adjustment. The z statistic of most of the regressors is halved once one accounts for design features. Coverage probabilities of the nominal 95% confidence level average 60% for the UNW and around 86% for SRS.

4 Conclusions

The study analysed the rationale of incorporating the sampling features when making inference using survey data.

A correct estimate of sampling variance is required in order to properly assess the statistical reliability of the point estimates through the construction of confidence intervals. The analysis points out the following.

- The standard estimators for sampling variance assume URS, i.e. that sampling units are independently and identically distributed. On the contrary, sample data are collected from finite populations and seldom using simple random sampling, but using some kind of stratification and clustering. Neglecting the finite population hypothesis has a small impact on the estimates when the sample is a negligible fraction of the population, but the consequence of sampling clusters instead of elements is to increase the variance of the estimates, because the randomization process is based on clusters of elements and the distribution of the elements in each cluster is not random but rather influenced by positive intra-cluster correlation. This higher variance can be partially mitigated by stratifying the PSUs and using controlled selection methods such as probability proportional to size (PPS).

- These features of sampling design must be taken into account in evaluating the variance of the estimator of interest. Because straightforward variance formula are often not available in a complex survey framework, the research must resort to linearization- or replication-based methods. Both methods provide biased estimates of the variance, the order of the bias being usually negligible for large samples. Given that they basically produce the same results, they can be used interchangeably according to the nature of the estimator and the software procedures available. Always check that the number of degrees of freedom for variance estimates is coherent with the method used to estimate sampling variance.

- The GVF can also be useful to get an idea of how the inference drawn with standard estimation formula for variance is affected by the complex design. In survey practice, a popular GVF is the Kish design effect, i.e. the ratio of the variance computed according to the complex survey design to the variance computed under the hypothesis of SRS.

Regarding the use of sampling weights the following conclusions are drawn.

- It is suggested to use survey weights for descriptive statistics. If some portion of the population is over-sampled or the list presents problems of imperfect coverage, or unit non-response is significant, and these factors are somehow related to the analysis variables, then point estimates from unweighted analyses could be seriously biased. This is going to affect the coverage of the confidence level statements through the bias ratio.

- The use of survey weights generally increases the variance of the estimates, approximately in proportion to the squared coefficient of variation of the weights. But, at least for descriptive inference in the SHIW, the net effect on the MSE shows that the increased variance of the weighted estimator is more than compensated for by the bias reduction (Faiella and Gambacorta 2007).

- Using survey weights in regression analysis gives design unbiased parameters that are robust to model misspecification.

- To avoid misleading inference on the parameters, the survey weights should be normalized dividing them by their mean, so that they sum up to the sample size.

A model to estimate the variance in the SHIW is proposed, using JRR. This method makes it possible to estimate straightforwardly the variance of even highly non-linear statistics and to account for weighting adjustment, and it can be applied without providing the design variables in the public datasets (thus protecting confidentiality).

The study then briefly reviews the impact of these factors using variance estimates in the SHIW both in an univariate and multivariate context. The main findings are as follows.

- The bias of unweighted univariate estimates can be rather considerable, ranging from 1 to 14 times the standard error, depending on the variable measured.
- The computed RMSE confirms that, in the majority of cases, the increase in bias associated with unweighted estimates is not compensated for by the decrease in the standard error.
- The Meffs and Deffs present an average value of 3.2 for univariate estimates. The 95% probability statements are confronted with the effective coverage probabilities when one takes into account both the inflation factor due to the design variance and the bias of the unweighted estimates. The mean of the income is included in the z*standard error range in only 64% of the samples, if weights and design features are ignored (92% using weights under the SRS assumption).
- For regression models, coverage probabilities can be as low as 50% if we include both the weights and the variance correction, while they are around 86-89% if we use weights and an SRS variance estimate. The increase in the standard error of estimates can reverse the significance level of some coefficients: some coefficients significant at 5% level under the UNW design are not different from zero in the other two set-ups and the t statistic computed on other covariates is almost halved once one accounts for design features. This result must be treated cautiously given the simple structure of the models presented. Further research is needed to evaluate the impact of survey design on regression analysis.

References

BINDER, D. (1983): "On the Variance of Asimptotically Normal Estimators from Complex Surveys," *International Statistical Review*, 51, 279–272.

BRANDOLINI, A., L. CANNARI, G. D'ALESSIO, AND I. FAIELLA (2004): "Household Wealth Distribution in Italy in the 1990s," *Temi di Discussione del Servizio Studi*, 504.

CHAMBERS, R. L., AND C. J. SKINNER (2003): *Analysis of survey data*. Wiley, New York.

CHANTALA, K. (2003): "Introduction to Analyzing Add Health Data," Discussion paper, Carolina Population Center, http://www.cpc.unc.edu/projects/addhealth/files/analyze.pdf.

COWELL, F. (1989): "Sampling Variance and Decomposable Inequality Measures," *Journal of Econometrics*, 42, 27–41.

DUNCAN, G. J., AND G. KALTON (1987): "Issues of design and analysis of surveys across time," *International statistical review*, 55, 97–117.

ELTINGE, J., AND W. SRIBNEY (1996): "Survey Sampling," Stata technical bulletin reprints, Stata Corp.

EUROSTAT (2002): "Variance estimation methods in the European Union," Monographs of official statistics, Eurostat, http://epp.eurostat.ec.europa.eu/pls/portal/.

FAIELLA, I., AND R. GAMBACORTA (2007): "The weighting process in the SHIW," *Temi di Discussione del Servizio Studi*, 636.

FAIELLA, I., R. GAMBACORTA, S. IEZZI, AND A. NERI (2006): "Italian Household Budgets in 2004," Supplementi al Bollettino Statistico - Indagini Campionarie 7, Bank of Italy.

FAIELLA, I., AND A. NERI (2004): "La ricchezza delle famiglie italiane e americane," *Temi di Discussione del Servizio Studi*, 501.

FRANCISCO, C., AND W. FULLER (1991): "Quantile Estimation with a Complex Survey Design," *The Annals of statistics*, 19, 454–469.

FRANKEL, M. (1971): "Inference from Survey Samples," Discussion paper, ISR-The University of Michigan.

Gelman, A., J. B. . Carlin, H. S. . Stern, and D. B. . Rubin (2003): *Bayesian Data Analysis*, Chapman & Hall Texts in Statistical Science. Chapman & Hall, 2 edn.

Groves, R., F. Fowler, M. Couper, J. Lepkowsky, E. Singer, and R. Tourangeau (2004): *Survey Methodology*. Wiley.

Hansen, M., W. Madow, and B. Tepping (1983): "An evaluation of model-dependent and probability sampling inferences in sample surveys," *Journal of the American Statistical Association*, 78, 776–793.

Hansen, M., and B. Tepping (1985): "Estimation of variance in NAEP," Mimeo, NAEP.

Kish, L. (1965): *Survey Sampling*. Wiley, New York.

——— (1987): *Statistical design for research*. Wiley, New York.

——— (1992): "Weighting for Unequal Pi," *Journal of Official Statistics*, 8, 183–200.

Kish, L., and M. Frankel (1974): "Inference from complex samples," *The Journal of the Royal Statistical Society Series B*, 36, 1–37.

Kott, P. (1991): "A Model-Based Look at Linear Regression with Survey Data," *The American Statistician*, 45, 107–112.

Kovar, J. G., J. N. K. Rao, and C. F. J. Wu (1988): "Bootstrap and other methods to measure errors in survey estimates," *Canadian J. Statist.*, 16, 25–45.

Little, R. (1981): "Robust model-based inference for a finite population mean from unequally weighted samples," in *Proceedings of the Survey Research Methods Section American Statistical Association.*

Little, R. J. A., S. Lewitzky, S. Heeringa, J. Lepkowski, and R. C. Kessier (1997): "Assessment of Weighting Methodology for the National Comorbidity Survey," *American Journal of Epidemiology*, 146, 290–300.

Lohr, S. (1999): *Sampling: Design and Analysis*. Duxbury Press.

Nathan, G., and T. Smith (1989): "The Effect of Selection in Regression Analysis," in *Analysis of Complex Surveys*, pp. 149–163. Wiley.

Purdon, S., and K. Pickering (2001): "The use of sampling weights in the analysis of the 1998 Workplace Employee Relations Survey," Discussion paper, National Center for Social Research.

Rao, J., and C. Wu (1984): "Boostrap inference from sample surveys," in *Proceedings of the Survey Research Methods Section American Statistical Association*.

Rust, K. (1985): "Variance Estimation for Complex Estimators in Sample Surveys," *Journal of Official Statistics*, 1, 381–397.

Shao, J., and D. Tu (1995): *The Jackknife and Bootstrap*. Springer-Verlag.

Skinner, C., D. Holt, and T. Smith (1989): *Analysis of Complex Surveys*. Wiley, New York.

Särndal, C., and S. Lundström (2005): *Estimation on Surveys with Survey Nonresponse*. Wiley, New York.

Särndal, C., B. Swensson, and J. Wretman (1992): *Model Assisted Survey Sampling*. Springer-Verlag.

Wolter, K. (1985): *Introduction to Variance Estimation*. Springer-Verlag.

Woodruff, R. (1952): "Confidence Intervals for Medians and Other Position Measures," *Journal of the American Statistical Association*, 47, 635–646.

APPENDIX: Statistical Tables

Table A.1. Means, standard errors and 95 per cent confidence intervals*

Variables	Estimate			Standard errors			Lower bound			Upper bound		
	UNW	**SRS**	**COMPL**	**UNW**	**SRS**	**COMPL**	**UNW**	**SRS**	**COMPL**	**UNW**	**SRS**	**COMPL**
A. Income	27,962	27,608	27,608	243	235	323	27,485	27,148	26,972	28,438	28,069	28,244
B. Expenditure	20,419	20,243	20,243	152	150	189	20,120	19,948	19,870	20,717	20,537	20,616
C. Wealth	179,998	176,412	176,412	3,272	3,304	4,331	173,584	169,936	167,892	186,413	182,888	184,932
D. Age of head of HH	55.23	53.41	53.41	0.1845	0.1873	0.3441	54.87	53.04	52.73	55.60	53.77	54.08
E. Head of HH male	69.90	70.63	70.63	0.5125	0.5089	0.8278	68.90	69.64	69.00	70.91	71.63	72.26
F. Household size	2.64	2.69	2.69	0.0143	0.0146	0.0319	2.61	2.66	2.63	2.67	2.72	2.75
G. Male	48.25	48.50	48.50	0.34	0.34	0.39	47.58	47.83	47.73	48.92	49.18	49.28
H. Age of individuals	43.36	41.28	41.28	0.00	0.15	0.34	43.06	40.98	40.61	43.67	41.58	41.95
I. Employee	27.51	28.55	28.55	0.31	0.31	0.62	26.90	27.94	27.33	28.11	29.16	29.77
L. Self-Employed	7.76	8.18	8.18	0.18	0.19	0.32	7.40	7.81	7.55	8.13	8.55	8.81
M. Retired	25.79	22.74	22.74	0.30	0.29	0.57	25.20	22.17	21.62	26.38	23.30	23.85
N. Unemployed	13.56	13.29	13.29	0.24	0.23	0.44	13.10	12.83	12.42	14.02	13.74	14.15
O. Housewife	12.85	12.56	12.56	0.23	0.23	0.43	12.40	12.11	11.70	13.30	13.00	13.41
P. Low level of education	65.62	66.12	66.12	0.33	0.33	0.69	64.98	65.48	64.76	66.26	66.76	67.49
Q. High level of education	6.03	5.72	5.72	0.16	0.16	0.28	5.71	5.41	5.17	6.35	6.03	6.27

** SHIW 2002. Euros, units, age, percentages.*

UNW=unweighted estimates; SRS=weighted estimates; variance estimated with SRS formulas;

COMPL=weighted estimates; variance estimated using JRR replicated weights;

JRR replicated weights loaded in STATA 9.1 with the command: `svyset _n [pw=pesofl],vce(jack) jkrweight(pwt*,multiplier(1));`

JRR replicated weights loaded in R 2.5 with the command: `svrepdesign(data=dataset,type=c("JKn"),repweights=pesirep,scale=1,rscales=1)`.

Table A.2. Deffs, Meffs and other effects of the sampling design*

Variables	Bias	Bias Ratio	Meff	Deff	RMSE			Effective coverage probabilities		
					UNW	**SRS**	**COMPL**	Nominal	**UNW**	**SRS**
A. Income	-354	-1.456	1.769	1.893	429	235	323	0.9500	0.6426	0.9229
B. Expenditure	-176	-1.154	1.545	1.589	233	150	189	0.9500	0.7356	0.9400
C. Wealth	-3,586	-1.096	1.751	1.718	4,855	3,304	4,331	0.9500	0.7326	0.9326
D. Age of head of HH	-1.83	-9.9	3.478	3.374	1.84	0.19	0.34	0.9500	0.0000	0.8570
E. Head of HH male	0.0073	1.423	2.609	2.646	0.0089	0.0051	0.0083	0.9500	0.612	0.8859
F. Household size	0.0493	3.442	4.934	4.79	0.0514	0.0146	0.0319	0.9500	0.2449	0.8147
G. Male	0.2531	0.737	1.315	1.315	0.4268	0.3437	0.3941	0.9500	0.8476	0.9563
H. Age of individuals	-2.0871	-13.606	4.943	4.97	2.0927	0.153	0.341	0.9500	0.0000	0.8103
I. Employee	1.0455	3.405	4.088	3.996	1.0897	0.3106	0.6208	0.9500	0.2334	0.8366
L. Self-Employed	0.4149	2.255	2.994	2.855	0.4539	0.1885	0.3184	0.9500	0.4249	0.8770
M. Retired	-3.053	-10.148	3.529	3.844	3.0678	0.2882	0.5651	0.9500	0.0000	0.8413
N. Unemployed	-0.2701	-1.147	3.497	3.557	0.3583	0.2334	0.4402	0.9500	0.6197	0.8506
O. Housewife	-0.2916	-1.267	3.569	3.639	0.3714	0.2279	0.4347	0.9500	0.5993	0.8479
P. Low level of education	0.5042	1.544	4.497	4.529	0.6007	0.3255	0.6927	0.9500	0.5286	0.8215
Q. High level of education	-0.3154	-1.926	2.903	3.053	0.3554	0.1597	0.279	0.9500	0.4966	0.8690

** SHIW 2002. Euros, units, age, percentages.*

Table A.3. Medians, standard errors and 95 per cent confidence intervals*

Variables	Income				Expenditure				Wealth			
Design	**UNW**	**SRS**	**COMPL1**	**COMPL2**	**UNW**	**SRS**	**COMPL1**	**COMPL2**	**UNW**	**SRS**	**COMPL1**	**COMPL2**
Estimate	23,088	22,892	22,892	22,892	17,000	16,800	16,800	16,800	108,000	102,200	102,200	102,200
Standard error	218.78	281.68	284.67	378.54	88.81	133.21	133.21	468.8	1,221.14	1,665.19	1,887.21	3,061.49
Lower Bound	22,415	22,215	22,206	22,150	16,800	16,800	16,800	15,881	100,000	99,000	98,500	96,199
Upper Bound	23,400	23,484	23,488	23,634	17,200	17,400	17,400	17,719	105,500	106,500	107,000	108,201
DEFF		1	1.0214	1.8059		1	1	12.3844		1	1.2844	3.3802
MEFF	1		1.693	2.9936	1		2.25	27.8649	1		2.3884	6.2855

** SHIW 2002. Euros, percentages.*

UNW=unweighted estimates; SRS=weighted estimates; variance estimated with SRS formulas;

COMPL1=weighted estimates; variance estimated using the Woodruff procedure;

COMPL2=weighted estimates; variance estimated using JRR replicated weights.

Table A.4. Linear Regression*

	UNW				**SRS**				**COMPL**			
Dep. variable is *ln(Labour Income)*	Estimate	StdErr	t	P-value	Estimate	StdErr	t	P-value	Estimate	StdErr	t	P-value
Intercept	9.2122	0.1196	77.01	<.0001	9.3545	0.1215	77.02	<.0001	9.3545	0.2301	40.65	<.0001
Age	0.0132	0.0067	1.96	0.0496	0.0068	0.0069	0.98	0.3263	0.0068	0.0123	0.55	0.5820
Age squared	-0.0001	0.0001	-0.89	0.3743	0.0000	0.0001	0.39	0.6970	0.0000	0.0001	0.22	0.8260
Experience (years)	0.0395	0.0033	12.04	<.0001	0.0429	0.0034	12.76	<.0001	0.0429	0.0054	7.99	<.0001
Experience squared	-0.0008	0.0001	-11.25	<.0001	-0.0009	0.0001	-12.30	<.0001	-0.0009	0.0001	-7.94	<.0001
Male	0.2974	0.0141	-21.12	<.0001	0.3121	0.0143	-21.81	<.0001	0.3121	0.0179	-17.46	<.0001
Lower education	-0.3045	0.0155	-19.59	<.0001	-0.3251	0.0158	-20.60	<.0001	-0.3251	0.0232	-14.01	<.0001
Higher education	0.2662	0.0238	11.19	<.0001	0.2646	0.0244	10.86	<.0001	0.2646	0.0340	7.77	<.0001
Married	0.0208	0.0179	1.16	0.2451	0.0316	0.0179	1.76	0.0780	0.0316	0.0276	1.14	0.2530
Household size	-0.0261	0.0061	-4.27	<.0001	-0.0386	0.0061	-6.31	<.0001	-0.0386	0.0086	-4.50	<.0001

** SHIW 2002. Earning function.*

UNW=unweighted estimates; SRS=weighted estimates; variance estimated with SRS formulas;

COMPL=weighted estimates; variance estimated using JRR replicated weights;

JRR replicated weights loaded in STATA 9.1 with the command: `svyset _n [pw=pesofl],vce(jack) jkrweight(pwt*,multiplier(1));`

JRR replicated weights loaded in R 2.5 with the command: `svrepdesign(data=dataset,type=c("JKn"),repweights=pesirep,scale=1,rscales=1).`

Table A.5. Deffs, Meffs and other effects of the sampling design

Variables	Bias	Bias ratio	Meff	Deff	Effective coverage probabilities		
					Nominal	**UNW**	**SRS**
Intercept	0.1422	1.1892	3.7004	3.5898	0.9500	0.6049	0.8495
Age	-0.0065	-0.9599	3.3116	3.1690	0.9500	0.6544	0.8646
Age squared	0.0001	1.2949	3.4086	3.1289	0.9500	0.6017	0.8661
Experience (years)	0.0034	1.0244	2.6788	2.5528	0.9500	0.6821	0.8900
Experience squared	-0.0001	-1.8634	2.7230	2.3951	0.9500	0.5131	0.8973
Male	0.0148	1.0504	1.6123	1.5609	0.9500	0.7542	0.9417
Lower education	-0.0206	-1.3256	2.2307	2.1634	0.9500	0.6506	0.9087
Higher education	-0.0016	-0.0685	2.0485	1.9521	0.9500	0.8286	0.9197
Married	0.0108	0.6051	2.3845	2.3712	0.9500	0.7615	0.8985
Household size	-0.0125	-2.0426	1.9745	1.9680	0.9500	0.4743	0.9188

Table A.6. Logistic Regression*

	UNW				SRS				COMPL			
Dep. variable is *labour force=1*	Estimate	StdErr	z	P value	Estimate	StdErr	z	P value	Estimate	StdErr	z	P value
Intercept	-7.6527	0.2115	-36.17	<.0001	-7.6538	0.2141	-35.74	<.0001	-7.6538	0.4108	-18.63	<.0001
Age	0.5924	0.0114	51.89	<.0001	0.6033	0.0116	52.15	<.0001	0.6033	0.0229	26.39	<.0001
Age squared	-0.0075	0.0001	-55.15	<.0001	-0.0076	0.0001	-55.11	<.0001	-0.0076	0.0003	-27.28	<.0001
Male	1.6937	0.0463	-36.61	<.0001	1.8624	0.0473	-39.40	<.0001	1.8624	0.0888	-20.97	<.0001
Lower education	-0.4828	0.0456	-10.58	<.0001	-0.5741	0.0455	-12.62	<.0001	-0.5741	0.0649	-8.85	<.0001
Higher education	1.0962	0.1024	10.71	<.0001	1.0777	0.1054	10.23	<.0001	1.0777	0.1236	8.72	<.0001
Married	-0.5596	0.0642	-8.71	<.0001	-0.5817	0.0636	-9.14	<.0001	-0.5817	0.0951	-6.12	<.0001
Number of earners	0.5349	0.0257	20.85	<.0001	0.6008	0.0265	22.64	<.0001	0.6008	0.0448	13.41	<.0001
Number of dependants	-0.0608	0.0278	-2.20	0.0280	-0.0697	0.0269	-2.60	0.0090	-0.0697	0.0338	-2.07	0.0400

* *SHIW 2002. Labour Force Participation.*

UNW=unweighted estimates; SRS=weighted estimates; variance estimated with SRS formulas;

COMPL=weighted estimates; variance estimated using JRR replicated weights;

JRR replicated weights loaded in STATA 9.1 with the command: `svyset _n [pw=pesofl],vce(jack) jkrweight(pwt*,multiplier(1))`;

JRR replicated weights loaded in R 2.5 with the command: `svrepdesign(data=dataset,type=c("JKn"),repweights=pesirep,scale=1,rscales=1)`.

Table A.7. Deffs, Meffs and other effects of the sampling design

Variables	Bias	Bias ratio	Meff	Deff	Effective coverage probabilties		
					Nominal	**UNW**	**SRS**
Intercept	-0.0011	-0.0052	3.7705	3.6797	0.9500	0.6872	0.8074
Age	0.0109	0.9547	4.0100	3.9052	0.9500	0.6194	0.7971
Age squared	-0.0001	-0.8130	4.2096	4.0819	0.9500	0.6237	0.7893
Male	0.1687	3.6464	3.6848	3.5308	0.9500	0.1881	0.8145
Lower education	-0.0913	-2.0014	2.0209	2.0314	0.9500	0.4857	0.9016
Higher education	-0.0185	-0.1807	1.4579	1.3766	0.9500	0.8916	0.9478
Married	-0.0221	-0.3441	2.1903	2.2310	0.9500	0.8028	0.8883
Number of earners	0.0659	2.5690	3.0509	2.8519	0.9500	0.3589	0.8503
Number of dependants	-0.0089	-0.3200	1.4794	1.5825	0.9500	0.8808	0.9323

RECENTLY PUBLISHED "TEMI" (*)

N. 636 – *The weighting process in the SHIW*, by Ivan Faiella and Romina Gambacorta (June 2007).

N. 637 – *Emerging markets spreads and global financial conditions*, by Alessio Ciarlone, Paolo Piselli and Giorgio Trebeschi (June 2007).

N. 638 – *Comparative advantage patterns and domestic determinants in emerging countries: An analysis with a focus on technology*, by Daniela Marconi and Valeria Rolli (September 2007).

N. 639 – *The generation gap: Relative earnings of young and old workers in Italy*, by Alfonso Rosolia and Roberto Torrini (September 2007).

N. 640 – *The financing of small innovative firms: The Italian case*, by Silvia Magri (September 2007).

N. 641 – *Assessing financial contagion in the interbank market: Maximum entropy versus observed interbank lending patterns*, by Paolo Emilio Mistrulli (September 2007).

N. 642 – *Detecting long memory co-movements in macroeconomic time series, by* Gianluca Moretti (September 2007).

N. 643 – *The producer service sector in Italy: Long-term growth and its local determinants*, by Valter Di Giacinto and Giacinto Micucci (September 2007).

N. 644 – *Aggregazioni bancarie e specializzazione nel credito alle PMI: peculiarità per area geografica*, by Enrico Beretta and Silvia Del Prete (November 2007).

N. 645 – *Costs and benefits of creditor concentration: An empirical approach*, by Amanda Carmignani and Massimo Omiccioli (November 2007).

N. 646 – *Does the underground economy hold back financial deepening? Evidence from the Italian credit market*, by Giorgio Gobbi and Roberta Zizza (November 2007).

N. 647 – *Optimal monetary policy under low trend inflation*, by Guido Ascari and Tiziano Ropele (November 2007).

N. 648 – *Indici di bilancio e rendimenti di borsa: un'analisi per le banche italiane*, by Angela Romagnoli (November 2007).

N. 649 – *Bank profitability and taxation,* by Ugo Albertazzi and Leonardo Gambacorta (November 2007).

N. 650 – *Modelling bank lending in the euro area: A non-linear approach,* by Leonardo Gambacorta and Carlotta Rossi (November 2007).

N. 651 – *Revisiting poverty and welfare dominance,* by Gian Maria Tomat (November 2007).

N. 652 – *The general equilibrium effects of fiscal policy: Estimates for the euro area,* by Lorenzo Forni, Libero Monteforte and Luca Sessa (November 2007).

N. 653 – *Securitisation and the bank lending channel,* by Yener Altunbas, Leonardo Gambacorta and David Marqués (November 2007).

N. 654 – *The cyclical response of fiscal policies in the euro area. Why do results of empirical research differ so strongly?*, by Roberto Golinelli and Sandro Momigliano (January 2008).

N. 655 – *What's behind "inflation perceptions"? A survey-based analysis of Italian consumers*, by Paolo Del Giovane, Silvia Fabiani and Roberto Sabbatini (January 2008).

N. 656 – *The effects of fiscal policy in Italy: Evidence from a VAR model*, by Raffaela Giordano, Sandro Momigliano, Stefano Neri and Roberto Perotti (January 2008).

N. 657 – *Excess money growth and inflation dynamics*, by Barbara Roffia and Andrea Zaghini (January 2008).

N. 658 – *R&D and market structure in a horizontal differentiation framework*, by Davide Fantino (January 2008).

N. 659 – *Housing market spillovers: Evidence from an estimated DSGE model*, by Matteo Iacoviello and Stefano Neri (January 2008).

N. 660 – *Real exchange rate volatility and disconnect: An empirical investigation*, by Riccardo Cristadoro, Andrea Gerali, Stefano Neri and Massimiliano Pisani (April 2008).

N. 661 – *The effect of investment tax credit: Evidence from an atypical programme in Italy*, by Raffaello Bronzini, Guido de Blasio, Guido Pellegrini and Alessandro Scognamiglio (April 2008).

(*) Requests for copies should be sent to:
Banca d'Italia – Servizio Studi di struttura economica e finanziaria – Divisione Biblioteca e Archivio storico – Via Nazionale, 91 – 00184 Rome – (fax 0039 06 47922059). They are available on the Internet www.bancaditalia.it.

2005

L. DEDOLA and F. LIPPI, *The monetary transmission mechanism: Evidence from the industries of 5 OECD countries,* European Economic Review, 2005, Vol. 49, 6, pp. 1543-1569, **TD No. 389 (December 2000)**.

D. Jr. MARCHETTI and F. NUCCI, *Price stickiness and the contractionary effects of technology shocks*. European Economic Review, Vol. 49, 5, pp. 1137-1164, **TD No. 392 (February 2001)**.

G. CORSETTI, M. PERICOLI and M. SBRACIA, *Some contagion, some interdependence: More pitfalls in tests of financial contagion,* Journal of International Money and Finance, Vol. 24, 8, pp. 1177-1199, **TD No. 408 (June 2001)**.

GUISO L., L. PISTAFERRI and F. SCHIVARDI, *Insurance within the firm*. Journal of Political Economy, Vol. 113, 5, pp. 1054-1087, **TD No. 414** (**August 2001**)

R. CRISTADORO, M. FORNI, L. REICHLIN and G. VERONESE, *A core inflation indicator for the euro area,* Journal of Money, Credit, and Banking, Vol. 37, 3, pp. 539-560, **TD No. 435 (December 2001)**.

F. ALTISSIMO, E. GAIOTTI and A. LOCARNO, *Is money informative? Evidence from a large model used for policy analysis,* Economic & Financial Modelling, Vol. 22, 2, pp. 285-304, **TD No. 445 (July 2002)**.

G. DE BLASIO and S. DI ADDARIO, *Do workers benefit from industrial agglomeration?* Journal of regional Science, Vol. 45, (4), pp. 797-827, **TD No. 453 (October 2002).**

G. DE BLASIO and S. DI ADDARIO, *Salari, imprenditorialità e mobilità nei distretti industriali italiani,* in L. F. Signorini, M. Omiccioli (eds.), *Economie locali e competizione globale: il localismo industriale italiano di fronte a nuove sfide,* Bologna, il Mulino, **TD No. 453 (October 2002).**

R. TORRINI, *Cross-country differences in self-employment rates: The role of institutions*, Labour Economics, Vol. 12, 5, pp. 661-683, **TD No. 459 (December 2002)**.

A. CUKIERMAN and F. LIPPI, *Endogenous monetary policy with unobserved potential output,* Journal of Economic Dynamics and Control, Vol. 29, 11, pp. 1951-1983, **TD No. 493 (June 2004)**.

M. OMICCIOLI, *Il credito commerciale: problemi e teorie,* in L. Cannari, S. Chiri e M. Omiccioli (eds.), *Imprese o intermediari? Aspetti finanziari e commerciali del credito tra imprese in Italia,* Bologna, Il Mulino, **TD No. 494 (June 2004)**.

L. CANNARI, S. CHIRI and M. OMICCIOLI, *Condizioni di pagamento e differenziazione della clientela,* in L. Cannari, S. Chiri e M. Omiccioli (eds.), *Imprese o intermediari? Aspetti finanziari e commerciali del credito tra imprese in Italia*, Bologna, Il Mulino, **TD No. 495 (June 2004)**.

P. FINALDI RUSSO and L. LEVA, *Il debito commerciale in Italia: quanto contano le motivazioni finanziarie?,* in L. Cannari, S. Chiri e M. Omiccioli (eds.), *Imprese o intermediari? Aspetti finanziari e commerciali del credito tra imprese in Italia*, Bologna, Il Mulino, **TD No. 496 (June 2004)**.

A. CARMIGNANI, *Funzionamento della giustizia civile e struttura finanziaria delle imprese: il ruolo del credito commerciale,* in L. Cannari, S. Chiri e M. Omiccioli (eds.), *Imprese o intermediari? Aspetti finanziari e commerciali del credito tra imprese in Italia,* Bologna, Il Mulino, **TD No. 497 (June 2004)**.

G. DE BLASIO, *Credito commerciale e politica monetaria: una verifica basata sull'investimento in scorte,* in L. Cannari, S. Chiri e M. Omiccioli (eds.), *Imprese o intermediari? Aspetti finanziari e commerciali del credito tra imprese in Italia,* Bologna, Il Mulino, **TD No. 498 (June 2004)**.

G. DE BLASIO, *Does trade credit substitute bank credit? Evidence from firm-level data.* Economic notes, Vol. 34, 1, pp. 85-112, **TD No. 498 (June 2004).**

A. DI CESARE, *Estimating expectations of shocks using option prices,* The ICFAI Journal of Derivatives Markets, Vol. 2, 1, pp. 42-53, **TD No. 506 (July 2004).**

M. BENVENUTI and M. GALLO, *Il ricorso al "factoring" da parte delle imprese italiane,* in L. Cannari, S. Chiri e M. Omiccioli (eds.), *Imprese o intermediari? Aspetti finanziari e commerciali del credito tra imprese in Italia,* Bologna, Il Mulino, **TD No. 518 (October 2004)**.

L. CASOLARO and L. GAMBACORTA, *Redditività bancaria e ciclo economico,* Bancaria, Vol. 61, 3, pp. 19-27, **TD No. 519 (October 2004)**.

F. PANETTA, F. SCHIVARDI and M. SHUM, *Do mergers improve information? Evidence from the loan market,* CEPR Discussion Paper, 4961, **TD No. 521 (October 2004)**.

P. DEL GIOVANE and R. SABBATINI, *La divergenza tra inflazione rilevata e percepita in Italia,* in P. Del Giovane, F. Lippi e R. Sabbatini (eds.), *L'euro e l'inflazione: percezioni, fatti e analisi,* Bologna, Il Mulino, **TD No. 532 (December 2004).**

R. TORRINI, *Quota dei profitti e redditività del capitale in Italia: un tentativo di interpretazione*, Politica economica, Vol. 21, 1, pp. 7-41, **TD No. 551 (June 2005)**.

M. OMICCIOLI, *Il credito commerciale come "collateral"*, in L. Cannari, S. Chiri, M. Omiccioli (eds.), Imprese o intermediari? Aspetti finanziari e commerciali del credito tra imprese in Italia, Bologna, il Mulino, **TD No. 553 (June 2005)**.

L. CASOLARO, L. GAMBACORTA and L. GUISO, *Regulation, formal and informal enforcement and the development of the household loan market. Lessons from Italy*, in Bertola G., Grant C. and Disney R. (eds.) The Economics of Consumer Credit: European Experience and Lessons from the US, Boston, MIT Press, **TD No. 560 (September 2005)**.

S. DI ADDARIO and E. PATACCHINI, *Lavorare in una grande città paga, ma poco*, in Brucchi Luchino (ed.), *Per un'analisi critica del mercato del lavoro*, Bologna , Il Mulino, **TD No. 570 (January 2006)**.

P. ANGELINI and F. LIPPI, *Did inflation really soar after the euro changeover? Indirect evidence from ATM withdrawals,* CEPR Discussion Paper, 4950, **TD No. 581 (March 2006)**.

S. FEDERICO, *Internazionalizzazione produttiva, distretti industriali e investimenti diretti all'estero,* in L. F. Signorini, M. Omiccioli (eds.), *Economie locali e competizione globale: il localismo industriale italiano di fronte a nuove sfide*, Bologna, il Mulino, **TD No. 592 (October 2002).**

S. DI ADDARIO, *Job search in thick markets: Evidence from Italy*, Oxford Discussion Paper 235, Department of Economics Series, **TD No. 605 (December 2006)**.

2006

F. BUSETTI, *Tests of seasonal integration and cointegration in multivariate unobserved component models,* Journal of Applied Econometrics, Vol. 21, 4, pp. 419-438, **TD No. 476 (June 2003).**

C. BIANCOTTI, *A polarization of inequality? The distribution of national Gini coefficients 1970-1996*, Journal of Economic Inequality, Vol. 4, 1, pp. 1-32, **TD No. 487 (March 2004)**.

L. CANNARI and S. CHIRI, *La bilancia dei pagamenti di parte corrente Nord-Sud (1998-2000),* in L. Cannari, F. Panetta (a cura di), Il sistema finanziario e il Mezzogiorno: squilibri strutturali e divari finanziari, Bari, Cacucci, **TD No. 490 (March 2004).**

M. BOFONDI and G. GOBBI, *Information barriers to entry into credit markets,* Review of Finance, Vol. 10, 1, pp. 39-67, **TD No. 509 (July 2004).**

FUCHS W. and LIPPI F., *Monetary union with voluntary participation,* Review of Economic Studies, Vol. 73, pp. 437-457 **TD No. 512 (July 2004).**

GAIOTTI E. and A. SECCHI, *Is there a cost channel of monetary transmission? An investigation into the pricing behaviour of 2000 firms*, Journal of Money, Credit and Banking, Vol. 38, 8, pp. 2013-2038 **TD No. 525 (December 2004)**.

A. BRANDOLINI, P. CIPOLLONE and E. VIVIANO, *Does the ILO definition capture all unemployment?,* Journal of the European Economic Association, Vol. 4, 1, pp. 153-179, **TD No. 529 (December 2004)**.

A. BRANDOLINI, L. CANNARI, G. D'ALESSIO and I. FAIELLA, *Household wealth distribution in Italy in the 1990s,* in E. N. Wolff (ed.) International Perspectives on Household Wealth, Cheltenham, Edward Elgar, **TD No. 530 (December 2004)**.

P. DEL GIOVANE and R. SABBATINI, *Perceived and measured inflation after the launch of the Euro: Explaining the gap in Italy,* Giornale degli economisti e annali di economia, Vol. 65, 2 , pp. 155-192, **TD No. 532 (December 2004).**

M. CARUSO, *Monetary policy impulses, local output and the transmission mechanism*, Giornale degli economisti e annali di economia, Vol. 65, 1, pp. 1-30, **TD No. 537 (December 2004).**

A. NOBILI, *Assessing the predictive power of financial spreads in the euro area: does parameters instability matter?,* Empirical Economics, Vol. 31, 1, pp. 177-195, **TD No. 544 (February 2005)**.

L. GUISO and M. PAIELLA, *The role of risk aversion in predicting individual behavior, In P. A. Chiappori e C. Gollier (eds.) Competitive Failures in Insurance Markets: Theory and Policy Implications, Monaco, CESifo*, **TD No. 546 (February 2005).**

G. M. TOMAT, *Prices product differentiation and quality measurement: A comparison between hedonic and matched model methods*, Research in Economics, Vol. 60, 1, pp. 54-68, **TD No. 547 (February 2005)**.

F. LOTTI, E. SANTARELLI and M. VIVARELLI, *Gibrat's law in a medium-technology industry: Empirical evidence for Italy*, in E. Santarelli (ed.), Entrepreneurship, Growth, and Innovation: the Dynamics of Firms and Industries, New York, Springer, **TD No. 555 (June 2005).**

F. BUSETTI, S. FABIANI and A. HARVEY, *Convergence of prices and rates of inflation*, Oxford Bulletin of Economics and Statistics, Vol. 68, 1, pp. 863-878, **TD No. 575 (February 2006).**

M. CARUSO, *Stock market fluctuations and money demand in Italy, 1913 - 2003*, Economic Notes, Vol. 35, 1, pp. 1-47, **TD No. 576 (February 2006)**.

S. IRANZO, F. SCHIVARDI and E. TOSETTI, *Skill dispersion and productivity: An analysis with matched data*, CEPR Discussion Paper, 5539, **TD No. 577 (February 2006).**

R. BRONZINI and G. DE BLASIO, *Evaluating the impact of investment incentives: The case of Italy's Law 488/92.* Journal of Urban Economics, Vol. 60, 2, pp. 327-349, **TD No. 582 (March 2006).**

R. BRONZINI and G. DE BLASIO, *Una valutazione degli incentivi pubblici agli investimenti*, Rivista Italiana degli Economisti , Vol. 11, 3, pp. 331-362, **TD No. 582 (March 2006).**

A. DI CESARE, *Do market-based indicators anticipate rating agencies? Evidence for international banks*, Economic Notes, Vol. 35, pp. 121-150, **TD No. 593 (May 2006).**

L. DEDOLA and S. NERI, *What does a technology shock do? A VAR analysis with model-based sign restrictions*, Journal of Monetary Economics, Vol. 54, 2, pp. 512-549, **TD No. 607 (December 2006)**.

R. GOLINELLI and S. MOMIGLIANO, *Real-time determinants of fiscal policies in the euro area*, Journal of Policy Modeling, Vol. 28, 9, pp. 943-964, **TD No. 609 (December 2006).**

P. ANGELINI, S. GERLACH, G. GRANDE, A. LEVY, F. PANETTA, R. PERLI,S. RAMASWAMY, M. SCATIGNA and P. YESIN, *The recent behaviour of financial market volatility*, BIS Papers, 29, **QEF No. 2 (August 2006)**.

2007

L. CASOLARO. and G. GOBBI, *Information technology and productivity changes in the banking industry*, Economic Notes, Vol. 36, 1, pp. 43-76, **TD No. 489 (March 2004)**.

M. PAIELLA, *Does wealth affect consumption? Evidence for Italy*, Journal of Macroeconomics, Vol. 29, 1, pp. 189-205, **TD No. 510 (July 2004).**

F. LIPPI. and S. NERI, *Information variables for monetary policy in a small structural model of the euro area*, Journal of Monetary Economics, Vol. 54, 4, pp. 1256-1270, **TD No. 511 (July 2004)**.

A. ANZUINI and A. LEVY, *Monetary policy shocks in the new EU members: A VAR approach*, Applied Economics, Vol. 39, 9, pp. 1147-1161, **TD No. 514 (July 2004)**.

R. BRONZINI, *FDI Inflows, agglomeration and host country firms' size: Evidence from Italy*, Regional Studies, Vol. 41, 7, pp. 963-978, **TD No. 526 (December 2004).**

L. MONTEFORTE, *Aggregation bias in macro models: Does it matter for the euro area?*, Economic Modelling, 24, pp. 236-261, **TD No. 534 (December 2004)**.

A. DALMAZZO and G. DE BLASIO, *Production and consumption externalities of human capital: An empirical study for Italy*, Journal of Population Economics, Vol. 20, 2, pp. 359-382, **TD No. 554 (June 2005).**

M. BUGAMELLI and R. TEDESCHI, *Le strategie di prezzo delle imprese esportatrici italiane,* Politica Economica, v. 3, pp. 321-350, **TD No. 563 (November 2005)**.

L. GAMBACORTA and S. IANNOTTI, *Are there asymmetries in the response of bank interest rates to monetary shocks?*, Applied Economics, v. 39, 19, pp. 2503-2517, **TD No. 566 (November 2005).**

S. DI ADDARIO and E. PATACCHINI, *Wages and the city. Evidence from Italy*, Development Studies Working Papers 231, Centro Studi Luca d'Agliano, **TD No. 570 (January 2006)**.

P. ANGELINI and F. LIPPI, *Did prices really soar after the euro cash changeover? Evidence from ATM withdrawals*, International Journal of Central Banking, Vol. 3, 4, pp. 1-22, **TD No. 581 (March 2006)**.

A. LOCARNO, *Imperfect knowledge, adaptive learning and the bias against activist monetary policies*, International Journal of Central Banking, v. 3, 3, pp. 47-85, **TD No. 590 (May 2006)**.

F. LOTTI and J. MARCUCCI, *Revisiting the empirical evidence on firms' money demand*, Journal of Economics and Business, Vol. 59, 1, pp. 51-73, **TD No. 595 (May 2006)**.

P. CIPOLLONE and A. ROSOLIA, *Social interactions in high school: Lessons from an earthquake*, American Economic Review, Vol. 97, 3, pp. 948-965, **TD No. 596 (September 2006).**

A. BRANDOLINI, *Measurement of income distribution in supranational entities: The case of the European Union*, in S. P. Jenkins e J. Micklewright (eds.), Inequality and Poverty Re-examined, Oxford, Oxford University Press, **TD No. 623 (April 2007).**

M. PAIELLA, *The foregone gains of incomplete portfolios*, Review of Financial Studies, Vol. 20, 5, pp. 1623-1646, **TD No. 625 (April 2007).**

K. BEHRENS, A. R. LAMORGESE, G.I.P. OTTAVIANO and T. TABUCHI, *Changes in transport and non transport costs: local vs. global impacts in a spatial network*, Regional Science and Urban Economics, Vol. 37, 6, pp. 625-648, **TD No. 628 (April 2007).**

G. ASCARI and T. ROPELE, *Optimal monetary policy under low trend inflation*, Journal of Monetary Economics, v. 54, 8, pp. 2568-2583, **TD No. 647 (November 2007).**

R. GIORDANO, S. MOMIGLIANO, S. NERI and R. PEROTTI, *The Effects of Fiscal Policy in Italy: Evidence from a VAR Model*, European Journal of Political Economy, Vol. 23, 3, pp. 707-733, **TD No. 656 (December 2007).**

2008

S. MOMIGLIANO, J. Henry and P. Hernández de Cos, *The impact of government budget on prices: Evidence from macroeconometric models*, Journal of Policy Modelling, v. 30, 1, pp. 123-143 **TD No. 523 (October 2004).**

P. DEL GIOVANE, S. FABIANI and R. SABATINI, *What's behind "inflation perceptions"? A survey-based analysis of Italian consumers, i*n P. Del Giovane e R. Sabbatini (eds.), *The Euro Inflation and Consumers' Perceptions. Lessons from Italy*, Berlin-Heidelberg, Springer, **TD No. 655 (January 2008).**

FORTHCOMING

S. SIVIERO and D. TERLIZZESE, *Macroeconomic forecasting: Debunking a few old wives' tales*, Journal of Business Cycle Measurement and Analysis, **TD No. 395 (February 2001)**.

P. ANGELINI, *Liquidity and announcement effects in the euro area*, Giornale degli economisti e annali di economia, **TD No. 451 (October 2002).**

S. MAGRI, *Italian households' debt: The participation to the debt market and the size of the loan*, Empirical Economics, **TD No. 454 (October 2002)**.

P. ANGELINI, P. DEL GIOVANE, S. SIVIERO and D. TERLIZZESE, *Monetary policy in a monetary union: What role for regional information?*, International Journal of Central Banking, **TD No. 457 (December 2002)**.

L. MONTEFORTE and S. SIVIERO, *The Economic Consequences of Euro Area Modelling Shortcuts*, Applied Economics, **TD No. 458 (December 2002).**

L. GUISO and M. PAIELLA,, *Risk aversion, wealth and background risk*, Journal of the European Economic Association, **TD No. 483 (September 2003).**

G. FERRERO, *Monetary policy, learning and the speed of convergence*, Journal of Economic Dynamics and Control, **TD No. 499 (June 2004).**

F. SCHIVARDI e R. TORRINI, *Identifying the effects of firing restrictions through size-contingent Differences in regulation*, Labour Economics, **TD No. 504 (giugno 2004)**.

C. BIANCOTTI, G. D'ALESSIO and A. NERI, *Measurement errors in the Bank of Italy's survey of household income and wealth*, Review of Income and Wealth, **TD No. 520 (October 2004)**.

D. Jr. MARCHETTI and F. Nucci, *Pricing behavior and the response of hours to productivity shocks,* Journal of Money Credit and Banking, **TD No. 524 (December 2004).**

L. GAMBACORTA, *How do banks set interest rates?,* European Economic Review, **TD No. 542 (February 2005).**

P. ANGELINI and A. Generale, *On the evolution of firm size distributions*, American Economic Review, **TD No. 549 (June 2005).**

R. FELICI and M. PAGNINI,, *Distance, bank heterogeneity and entry in local banking markets*, The Journal of Industrial Economics, **TD No. 557 (June 2005).**

M. BUGAMELLI and R. TEDESCHI, *Le strategie di prezzo delle imprese esportatrici italiane*, Politica Economica, **TD No. 563 (November 2005).**

S. DI ADDARIO and E. PATACCHINI, *Wages and the city. Evidence from Italy*, Labour Economics, **TD No. 570 (January 2006)**.

M. BUGAMELLI and A. ROSOLIA, *Produttività e concorrenza estera*, Rivista di politica economica, **TD No. 578 (February 2006).**

PERICOLI M. and M. TABOGA, *Canonical term-structure models with observable factors and the dynamics of bond risk premia*, **TD No. 580 (February 2006).**

E. VIVIANO, *Entry regulations and labour market outcomes. Evidence from the Italian retail trade sector*, Labour Economics, **TD No. 594 (May 2006)**.

S. FEDERICO and G. A. MINERVA, *Outward FDI and local employment growth in Italy*, Review of World Economics, Journal of Money, Credit and Banking, **TD No. 613 (February 2007).**

F. BUSETTI and A. HARVEY, *Testing for trend*, Econometric Theory **TD No. 614 (February 2007).**

V. CESTARI, P. DEL GIOVANE and C. ROSSI-ARNAUD, *Memory for Prices and the Euro Cash Changeover: An Analysis for Cinema Prices in Italy*, In P. Del Giovane e R. Sabbatini (eds.), The Euro Inflation and Consumers' Perceptions. Lessons from Italy, Berlin-Heidelberg, Springer, **TD No. 619 (February 2007).**

B. ROFFIA and A. ZAGHINI, *Excess money growth and inflation dynamics*, International Finance, **TD No. 629 (June 2007)**.

M. DEL GATTO, GIANMARCO I. P. OTTAVIANO and M. PAGNINI, *Openness to trade and industry cost dispersion: Evidence from a panel of Italian firms*, Journal of Regional Science, **TD No. 635 (June 2007)**.

A. CIARLONE, P. PISELLI and G. TREBESCHI, *Emerging Markets' Spreads and Global Financial Conditions*, Journal of International Financial Markets, Institutions & Money, **TD No. 637 (June 2007).**

S. MAGRI, *The financing of small innovative firms: The Italian case*, Economics of Innovation and New Technology, **TD No. 640 (September 2007)**.