Quality measures in non-statistical sampling: MFI Interest Rates Statistics (MIR)

J. Huerga, S. Pérez-Duarte and J.M. Puigvert[1]
European Central Bank, Frankfurt, Germany

Abstract

Traditional literature on sampling techniques mainly focuses on statistical samples and only marginally covers non-random (non-statistical) samples. Nevertheless, there has been a recent revival in interest on non-statistical samples due to their widespread use in certain fields like government surveys, marketing research, or for audit purposes. This paper tries to set up common rules for non-statistical sample in which only those largest institutions within each stratum are collected. This is done by focusing on the regular collection of Monetary Financial Institutions (MFIs) interest rate statistics from countries of the European Union as they are collected by the European System of Central Banks (ESCB). The paper concludes by proposing a way to establish common rules for non-statistical samples based on the basis of a synthetic measurement of a mean of absolute errors.

Keywords: sampling, interest rates, non-random and non-statistical samples.

1. Introduction

Non-statistical samples are commonly used in several fields, for example in US Federal surveys[2], market research, audit and tax inspection.[3] The extended use of non-statistical samples has recently revived the interest in the theoretical properties of the results obtained through these methods.[4] Given its popularity, cut-off sampling has also been given quite some attention; see for example Yorgason et al. (2011), Landry (2011), Benedetti et al. (2010) or Haziza et al. (2010).

A cut-off-style non-statistical sample is also used in several countries for the statistics collected by the ESCB on Monetary and Financial Institutions (MFIs) interest rates (MIR) applied on a range of deposits and loans from/to households and non-financial corporations.[5] The MIR statistics collection system in most of the EU countries is implemented by, first, stratifying the potential reporting population (all MFIs) in homogenous strata and then selecting the overall largest institutions within each stratum. For MIR there is no possibility to replace the current non-statistical sample either by a census or by a statistical sample because of the costs implied especially to cover smaller MFIs. The smallest MFIs are likely to contribute very little to lending and taking deposits, and the burden imposed on them would be excessive. Moreover, reporting in the MIR requires adaptations of the MFIs own reporting system, with a high one-off implementation cost. MFIs randomly selected for the purposes of the MIR statistics would therefore have significant and measurable start-up costs, impeding the possibility of rotating samples. In that respect, the case of MIR, with the selection of the largest institutions, resembles the cut-off sampling applied in US

[1] This paper is based on the work carried out and on a not yet published paper by the Technical Expert Group on MIR Statistics from October 2011 to June 2012. The Technical Expert Group was chaired by the ECB and included representatives from 13 EU national central banks. We are grateful for comments by Aurel Schubert, Jean-Marc Israël and Patrick Sandars. We would also like to thank Piotr Bojaruniec for his technical assistance. All views expressed are those of the authors alone and do not necessarily reflect those of the ECB or the Eurosystem. Authors contact e-mail: Javier.huerga@ecb.europa.eu, sebastien.perez_duarte@ecb.europa.eu, josep_maria.puigvert@ecb.europa.eu.

[2] See for instance, Yorgason, D. et al. (2011).

[3] See among others Fox (2010), Goodman (1961), and Babbie (1999).

[4] See for example Guarte (2006) where a theoretical exercise is performed on purposive samples on the central part of a population distribution by using bootstrap applied to different distribution functions.

[5] For a detailed reference on MIR statistics see Regulation (EC) No 63/2002.

Federal surveys as described in Yorgason et al. (2011). The reasons for using cut-off sampling include physical efficiencies, limitation of costs and reporting burden and ensuring data quality. In the same line as some of the recent literature described above, this paper explores how to establish a possible common quality measure on MIR statistics. The starting point for the analysis contained in this paper is the fact that data on MIR are collected at national level by each national central bank in the EU on the basis of different national stratifications of the potential reporting population (MFIs) and by selecting the largest institutions within each stratum as the actual reporting population. In particular, ECB Regulation 2001/18 establishes that stratification criteria "*should allow the subdivision of the potential reporting population into homogeneous strata. Strata are considered homogeneous if the sum of the intra-stratum variances of the sampling variables is substantially lower than the total variance in the entire actual reporting population*". Interest rates are then compiled by weighting interest rates by the respective business volumes attached to the loans and deposits involved. On that basis, this paper assumes that the stratification results in strata composed of institutions with similar features and therefore that the selection of the largest institutions is non-biased. However, the problem remains on how good this representation can be considered and how to establish a minimum quality threshold to ensure sufficient quality and homogeneity in the collection of these statistics country by country in the euro area and in the EU, which should also permit the collection of meaningful euro area aggregates. In order to establish a common quality measure this paper proposes an estimation of the Mean Absolute Error (MAE) as follows. First, it is assumed that each stratum can be theoretically divided in two sub-strata, a substratum from which all institutions are sampled ("take-all substratum") and a substratum from which no institution is sampled ("take-none substratum"). This approach is common practice in business surveys; see for example Landry (2011). Second, it is assumed that a measure of dispersion of the take-all substratum can serve to estimate the possible divergence of the take-none stratum from the take-all substratum within each strata and for each particular MIR indicator. A similar approach is used in Benedetti et al. (2010). However a crucial difference arises. While Benedetti et al. (2010) apply the cut-off on the basis of the variable to be measured, in the case of MIR the cut-off is applied in terms of total balance sheet while the data being collected refers to interest rates. The first and third quartiles on the actually sampled MIR data were obtained for each stratum. These two dispersion measures are used to estimate two alternative scenarios of the estimated rates for the take-none substrata and used to obtain different MAEs for each stratum, weighting sampled and non-sampled estimated rates by the respective business volumes. As a further step, the alternative scenarios of MAE for each MIR indicator in each stratum are combined to alternative scenarios for each MIR indicator at the level of the whole population and the most appropriate scenario, namely the use of a combination of first and third quartiles, was chosen. The final step consists in the establishment of a formula (synthetic MAE) that combines several MIR indicators in a single measure, on which a threshold can be established. A threshold is proposed on the basis of the data. Subject to the usual caveats on the non-representativeness of non-statistical samples, our empirical findings are that, under the conditions above described, it is possible to establish a common measure of the quality of samples on the basis of a synthetic MAE.

2. Assessing sampling quality in MIR based on MAE measures

This section describes in further detail the chosen alternative on how to measure the overall data quality on MIR sampling. This approach consists in the construction of a synthetic indicator based on the estimated Mean Absolute Error (MAE) for a particular estimator and for each MIR indicator.[6]

[6] For a given country, this synthetic indicator would provide an aggregated quality measure for all reported series in all the different strata in which the data is collected under some assumptions of the data

2.1 MAE indicator construction

A sample which is divided in j strata and in which only the largest institutions within each stratum comprise the actual reporting population can also be theoretically sub-divided into two sub-strata: j_o for non-reporting institutions, i.e. the so-called take-none substratum, and j_1 for reporting institutions, i.e. the so-called take-all sub-stratum of stratum j. The interest rate of the whole population in stratum j would be ideally obtained as

$$i_j = \frac{(i_{j1} * B_{j1} + i_{j0} * B_{j0})}{(B_{j1} + B_{j0})} \qquad (1)$$

Where i_j is the mean interest rate for stratum j, calculated as the mean interest rate for the take-all sub-stratum i_{j1} and the mean interest rate for the take-none sub-stratum, i_{j0}, weighted by the business volumes of take-all substratum, B_{j1} and the business volumes of take-all substratum B_{j0}. Given that in the sampling context the interest rate for the take-none substratum is not reported, a number of assumptions are needed to estimate the average interest rate for the stratum and subsequently calculate the estimated error. The core assumption is that the stratification, as required in the MIR Regulation results in homogeneous strata and therefore no bias results from the selection of the largest institutions. Under this assumption, the reported interest rate for the take-all substratum, i_{j1}, is used as the best estimation of the non-reported interest rate for the take-none substratum i_{j0}

$$\hat{i}_{j0} = i_{j1} \qquad (2)$$

where $\hat{i}_{j0}$ is the estimated interest rate for take-none substratum, which results in

$$\hat{i}_j = \frac{B_{j1}\, i_{j1} + B_{j0}\, \hat{i}_{j0}}{B_{j1} + B_{j0}} = i_{j1} \qquad (3)$$

where $\hat{i}_j$ is the estimated interest rate of the whole stratum j. The actual error of the estimator $\hat{i}_j$ of the interest rate would be the difference between the real and the estimated value of the interest rate within a stratum j as follows:

$$\text{error}(\hat{i}_j) = i_j - \hat{i}_j = \frac{B_{j1}\, i_{j1} + B_{j0}\, i_{j0}}{B_{j1} + B_{j0}} - i_{j1} = \frac{B_{j0}}{B_{j1} + B_{j0}}(i_{j0} - i_{j1}) \qquad (4)$$

The total error within stratum j gives an approximation of the error made to calculate the interest rate i_j considering the estimated business volume for the non-reporting institutions B_{j0}. In case that i_{j0} is known and coincides with i_{j1} or that B_{j0} is zero, i.e. there are no non-reporting institutions, then the total error within the stratum j, i.e. $\text{error}(\hat{i}_j)$, is exactly equal to zero. Here it is assumed that the business volume corresponding to the non-reporting institutions is known (for MIR outstanding

distribution of non-sampled institutions. Finally, the values of the synthetic indicator applied to the data reported by the NCBs give possible thresholds for considering the data reported as of good quality.

amounts) or at least expected to be estimated with a negligible error (for MIR new business). However, given that the interest rate for take-none substratum is not known, the error for the estimator $\hat{i}_j$ cannot be directly obtained but rather should be estimated. As the error cannot be directly calculated, because i_{j0} is not reported, only an estimate of the error can actually be calculated on the basis of the take-all distribution function, which in turn is assumed to be representative of the super-population distribution. In other words, for each value $\hat{\theta}_j$, it is possible to calculate the weighted number of observations in the take-all stratum included in an interval around the estimated average, e.g. $(i_{j1} - \hat{\theta}_j, i_{j1} + \hat{\theta}_j)$ or an interval around the estimated mode. Inversely, it is also possible to first define a desired level of confidence, apply it to the take-all substratum and calculate the value of $\hat{\theta}_j$ that complies with such confidence level. Then $\hat{\theta}_j$ would be input into the estimated error equation for the stratum:

$$\text{error}(\hat{i}_j) = \frac{B_{j1}\, i_{j1} + B_{j0}\, \hat{\theta}_j}{B_{j1} + B_{j0}} - i_{j1} = \frac{B_{j0}}{B_{j1} + B_{j0}} (\hat{\theta}_j - i_{j1}) \qquad (5)$$

The estimated errors from the different strata would be aggregated in a single Mean Absolute Error (MAE) figure, by weighting the stratum estimated error by business volume of each stratum:

$$MAE(\hat{\theta}) = \frac{\sum_j \left|\text{error}(\hat{i}_j)\right| (B_{j0} + B_{j1})}{B} \qquad (6)$$

with $\text{error}(\hat{i}_j)$ as defined in (5), B_{j1} and B_{j0} are the volumes previously defined in (1) and $B = \sum_j B_{j0} + B_{j1}$ is the total volume of all institutions in the whole population. The MAE can be interpreted as a measure for which all the individual differences per stratum are weighted by the volume within each stratum. The values that were initially considered for $\hat{\theta}$ in a stratum j included the first and third quartiles, i.e. $\hat{\theta}_j = Q1_{j1}$ and $\hat{\theta}_j = Q3_{j1}$, defined as the lowest interest rate reported for the MIR category that is higher than 25 per cent and 75 per cent, respectively, of the reported volumes in that category by the institutions in the stratum

2.2 A synthetic indicator based on the MAE

The MAE as defined in section 2.1 is calculated at the level of each individual indicator and depends on the volatility and the magnitude of each series. A possible solution to overcome the problem of having an individual MAE for each particular series and to provide at the same time with a single MAE figure would be to construct a synthetic MAE by weighting each series by its respective volume and dividing it by its interest rate. More in detail, the synthetic MAES for a given estimator $\hat{\theta}$ in a particular period can be defined as:

$$MAE_S(\hat{\theta}) = \sum_j \frac{MAE(\hat{\theta}_j) * B_j}{\sum_k B_k} * \frac{1}{(i_{j1} + (1/(1 + i_{j1}))} \qquad (7)$$

$$MAE_S(\hat{\theta}) = \sum_j \frac{B_j}{\sum_k B_k} \frac{MAE(\hat{\theta}_j)}{i_{j1} + \frac{1}{1+i_{j1}}} \tag{8}$$

where for each series j, $MAE(\hat{\theta}_j)$ is the MAE as defined in the previous section, B_j is the total volume reported $B_j = B_{j0} + B_{j1}$ for this series, and i_{j1} is the (reported) aggregated interest rate of this series for this particular period.

2.3 Results for the synthetic indicator at national level

Each country that participated in this exercise calculated the $MAE(\hat{\theta})$, as defined in section 2.1, on the 43 MIR regulation series for new business and outstanding amounts as defined in Regulation ECB/2001/18 and then aggregated for each MIR series to the national figures. The calculations were done on five different periods: September 2010, December 2010, March 2011, June 2011, and September 2011. A posteriori, these calculations were used to construct a synthetic $MAE_S(\hat{\theta})$, as defined in section 2.2, for new business and for outstanding amounts based on the first quartile (Q1) and third quartile (Q3) estimators, on the basis of robustness.[7]The interest rate in this case may end up in outcomes lying outside the distribution of the reported rates and this could imply a very large MAE value for a particular stratum. This outcome may be magnified when the calculus refer to strata composed by a relatively small number of reporting agents. Moreover, rather than taking separately the first and third quartile, it was observed that it is better to use the average of both estimators as a measure of central tendency and to avoid possible asymmetries between the two estimators. Tables 1 and table 2, provide the results of the synthetic MAE for the mean of the first quartile (Q1) and third quartile (Q3) estimators applied for each country. These figures are shown in pure units, as the synthetic MAE in (8) does not have a particular unit of measurement.

Table 1. New Business MAE synthetic indicators (average of the 5 periods)

	AT	DE	ES	FR	GR	IE	IT	LT	NL	PL
Q1	1.43	1.40	0.67	1.41	0.14	0.70	1.02	0.51	1.19	0.46
Q3	1.35	1.11	0.60	1.17	0.16	0.88	1.02	0.37	0.56	0.81
Q1 and Q3 mean	1.39	1.25	0.63	1.29	0.15	0.79	1.02	0.44	0.88	0.64

Table 2. Outstanding amounts MAE synthetic indicators (average of the 5 periods)

	AT	DE	ES	FR	GR	IE	IT	LT	NL	PL
Q1	2.40	1.98	0.45	0.87	0.16	1.11	1.87	0.19	0.58	3.76
Q3	2.18	1.95	0.46	0.58	0.18	1.06	1.62	0.23	0.56	2.73
Q1 and Q3 mean	2.29	1.97	0.46	0.73	0.17	1.09	1.74	0.21	0.57	3.25

4. Conclusions

Statistics on interest rates applied by MFIs on deposits from and loans to households and non-financial corporations (MIR statistics) are collected in the EU member states by NCBs by dividing the potential reporting population into strata and selecting the largest institutions within each stratum. This paper has revisited the quality of MIR sampling from a new perspective, namely not strictly applying sampling theory but rather using some simpler assumptions on the possible estimation of errors. That approach is somewhat similar to the methods used in other cut-off samples in the

[7] The reasons why the rest of the indicators were discarded directly depended on the estimator in consideration. As ‘outliers’ often play a significant role in the construction of statistics, the min and the max indicators should be interpreted as a conceptual extreme term of reference for the MAEs (and not exploited as actual measures of accuracy); a similar caution in the interpretation of the results should be adopted when the 2-standard deviation indicator is considered, as in the weighted distributions the average interest rate increases, by two times the standard deviation.

recent literature. However, different from these methods, in the case of MIR there is no information on the variable studied, in this case the interest rate, for the overall population. In order to address that issue, a number of assumptions are used in this paper, namely that the stratification in MIR statistics results in homogeneous strata, that within each of them there is no correlation between interest rate and size of the institution, that the possible error due to the use of estimated new business volumes is small and does not need to be considered. On that basis, the paper considers that the selection of the largest institutions can be considered in a scheme with two sub-strata take-all and take-none. Furthermore, both sub-strata can be considered as samples of a super population and therefore the statistics from the take-all sub-stratum can be applied to the take-none substratum. Several possible measures of dispersion obtained from the take-all substratum were considered and applied to national data. These measures were applied at the level of each interest rate indicator, expressed in relative terms to the level of the interest rate and then aggregated by weighting by the respective business volume into a single indicator for the main categories, new business and outstanding amounts. The results showed that the usability of the proposed synthetic MAE in particular when based on the combined use of the first and third quartile, which offer the most robust behaviour along time and across countries.

REFERENCES

Babbie, E. (1999), The basics of social research, Thomson, Wadsworth.

Baillargeon, S., Rivest, L.P. (1995), A general algorithm for univariate stratification, Université Laval, Quebec City.

Benedetti, R., Bee, M., Espa, G. (2010), A framework for cut-off sampling business survey design, Journal of Official Statistics, 26, no. 4, pp. 651-671.

Cochran, W. (1977), Sampling techniques, John Wiley.

Fox, R. J. (2010), Non-probability sampling, Wiley International Encyclopaedia of Marketing.

Goodman, L.A. (1961), Snowball sampling, Annals of Mathematical Statistics, vol. 32, no. 1, pp. 148–170.

Guarte, J. M., Barrios, E.B. (2006), Estimation under purposive sampling, Communications in Statistics-Simulations and Computation, vol. 35, pp. 277-284.

Haziza, D.,Chauvet, G., Deville, J.C. (2010), Sampling and estimation in the presence of cut-off sampling, Australian & New Zealand Journal of Statistics, vol. 52, no. 3, pp. 303-319.

Kish, L. (1995), Survey sampling, John Wiley & Sons.

Landry, S. (2011), Managing response burden by controlling sample selection and survey coverage, Joint Statistical Meetings 2011, American Statistical Association.

Tallarini, T. D. Jr. (2000), Risk-sensitive real business cycles, Journal of Monetary Economics, vol. 45, no.3, pp. 507-532.

Regulation (EC) No 63/2002 of the ECB of 20 December 2001 concerning statistics on interest rates applied by monetary financial institutions to deposits and loans vis-à-vis households and non-financial corporations (ECB/2001/18)

Regulation (EC) No 290/2009 of the ECB of 31 March 2009 amending Regulation (EC) No 63/2002 (ECB/2001/18) concerning statistics on interest rates applied by monetary financial institutions to deposits and loans vis-à-vis households and non-financial corporations (ECB/2009/7)

U.S. Office of Management and Budget (2006), Standards and Guidelines for Statistical Surveys, September 2006.

Yorgason, D., B. Bridgman, Y. Cheng, A. H. Dorfman, J. Lent, Y. K. Liu, J. Miranda, S. Rumburg (2011), Cuttoff sampling in Federal surveys: An inter-agency review, Joint Statistical Meetings 2011, American Statistical Association.