Analysis of sample size in consumer surveys

GfK Polonia, 2013, Poland

Agenda

1. Introduction
   a. Theoretical aspects
   b. Sample size in the DG ECFIN Consumer Survey

2. Quality measures vs. sample size
   a. Sample size - analysis of impact
   b. Effective sample size - analysis of impact
   c. Additional analysis

3. Summary and conclusions

4. References.

1. Introduction

Sample size: Theoretical aspects

Main aspects of determining the size of a sample

- the purpose of the study
- population size
- sampling error (the level of precision)
- the confidence level
- the degree of variability in the main measured attributes

Sample size: Theoretical aspects

Ways of determining the size of a sample

- using or conducting a census survey
- transfer a sample size from a similar study
- using published tables
- applying of formulas to calculate a sample size

Sample size: Theoretical aspects, formulas

Representative sample for proportions in large populations

(W. G. Cochran)

$$n_0 = \frac{Z^2 * p * (1 - p)}{e^2}$$

n0 - the sample size

Z^2 - the abscissa of the normal curve that cuts off an area α at the tails

e - the acceptable sampling error

p - the estimated proportion of an attribute that is present in the population

Sample size: Theoretical aspects, formulas

Finite population correction for proportions

$$n = \frac{n_0}{1 + \frac{(n_0 - 1)}{N}}$$

n0 - the initial sample size

n - adjusted sample size

N - the population size.

Sample size: Theoretical aspects, formulas

Simplified formula for proportions*

(Taro Yamane)

$$n = \frac{N}{1 + N * (e)^2}$$

n - the sample size

N - the population size

e - the acceptable sampling error

* *95% confidence level and p = 0.5 are assumed*

Sample size: Theoretical aspects, formulas

Formula for sample size for the mean

$$n_0 = \frac{Z^2 * \sigma^2}{e^2}$$

n0 - the sample size

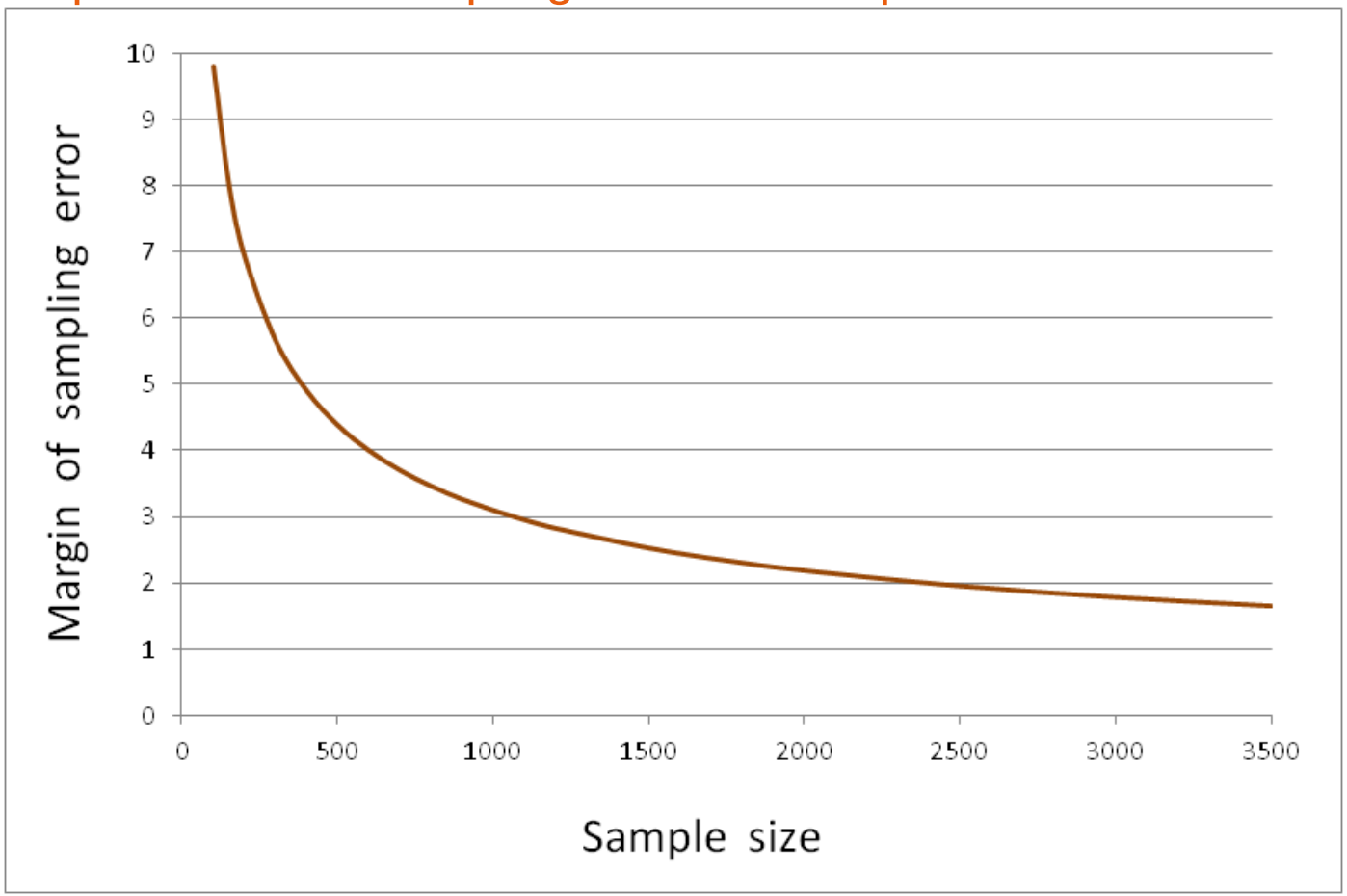
Z - the abscissa of the normal curve that cuts off an area σ at the tails

e - the acceptable sampling error

σ^2 - the variance of an attribute in the population

Sample size: Theoretical aspects

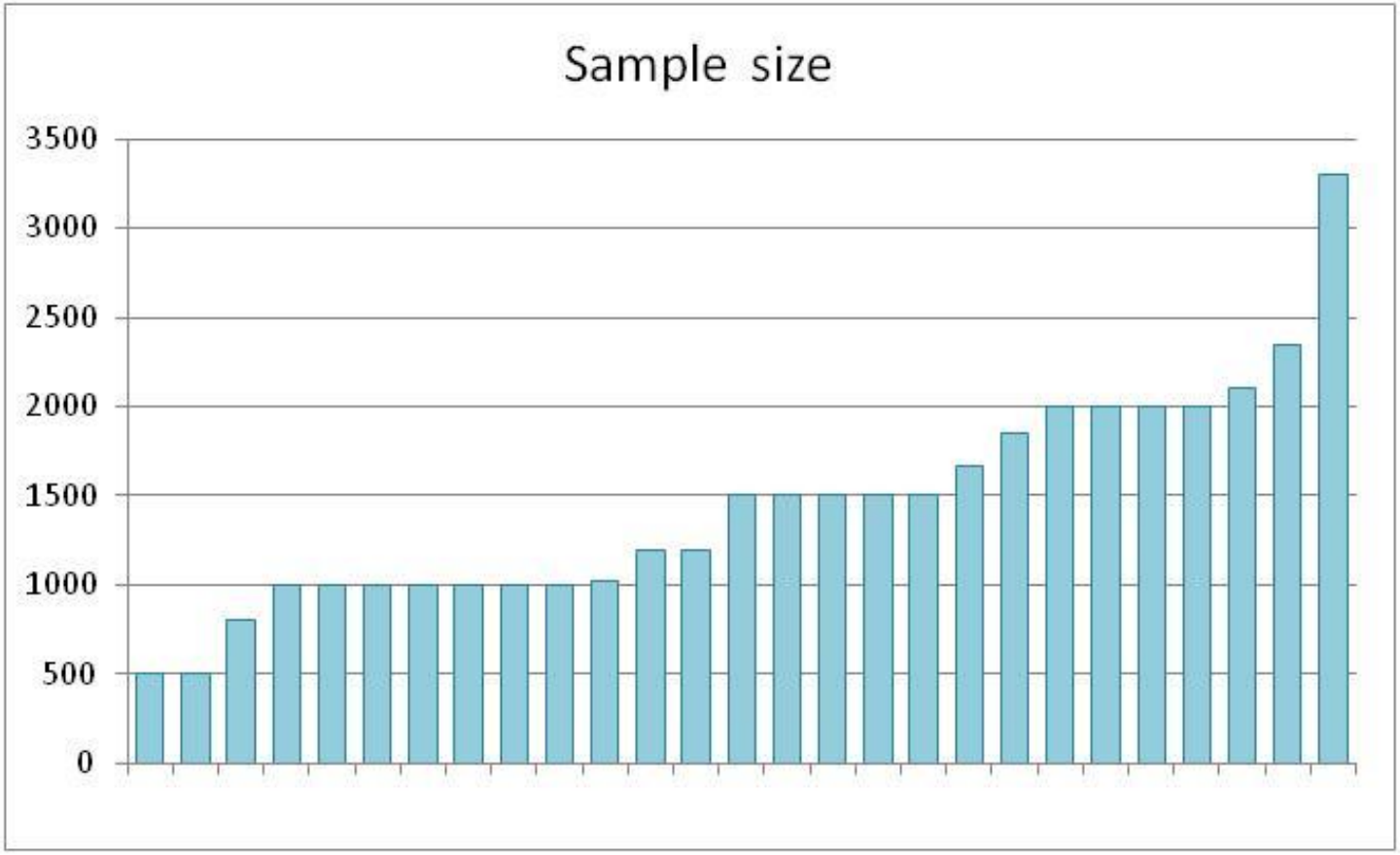
Dependence of sampling error on sample size



Margin of sampling error was calculated for 95% confidence interval, 50% fraction and infinite population

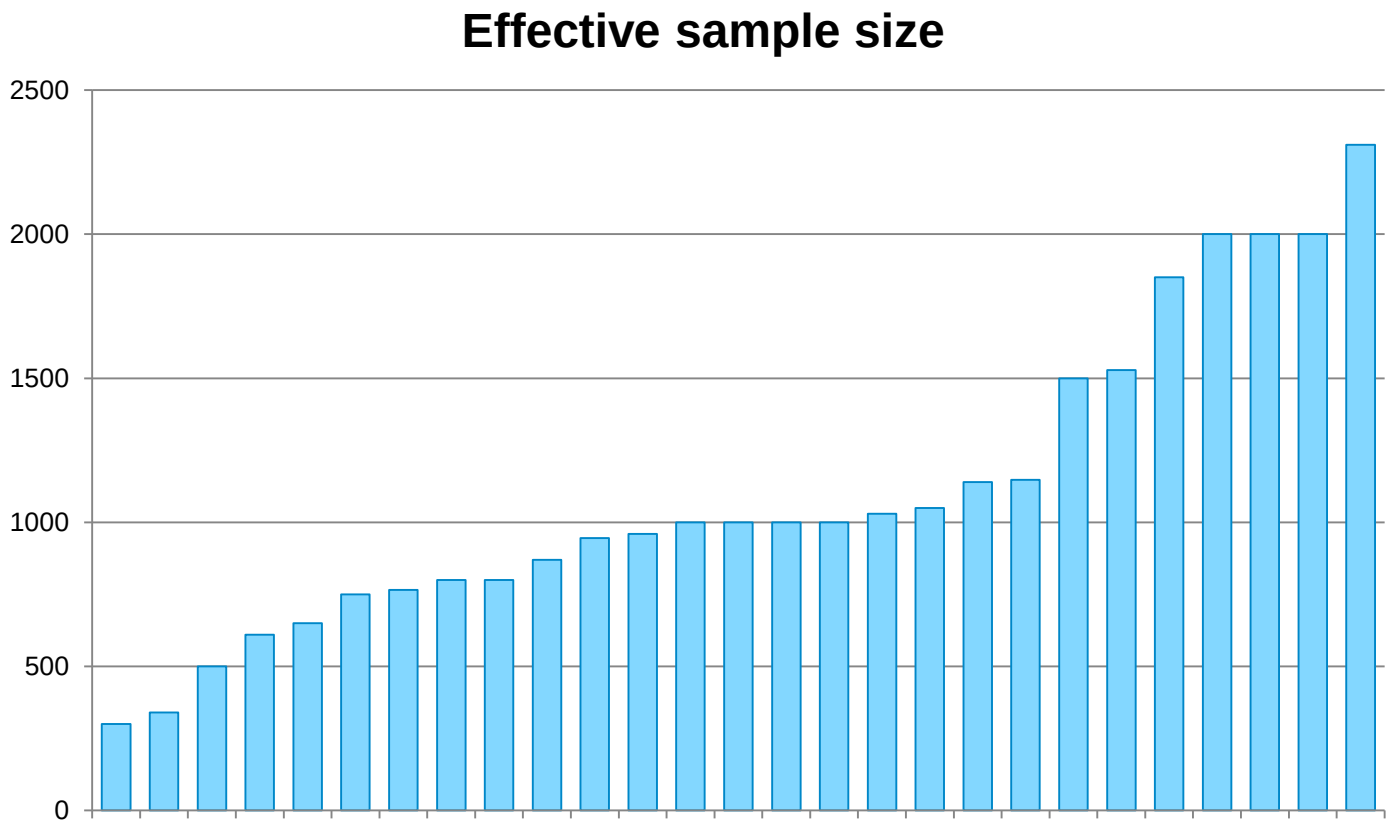
Sample size in the DG ECFIN Consumer Survey

Sample sizes in the survey



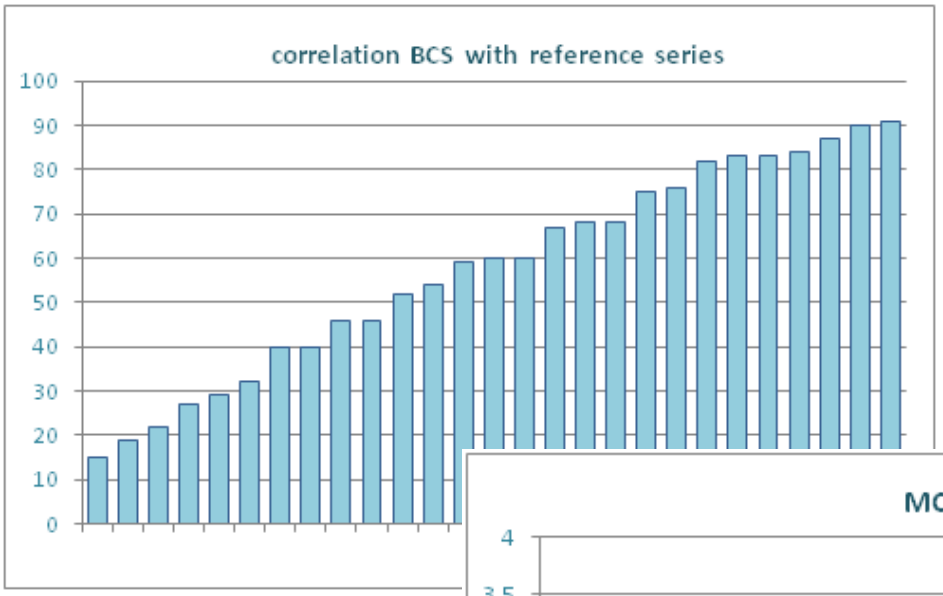
Sample size in the DG ECFIN Consumer Survey

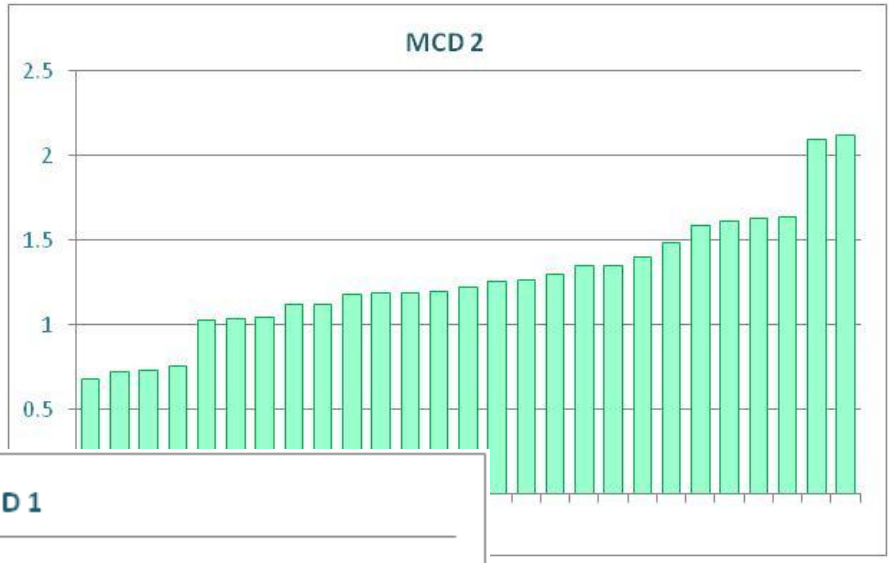
Effective sample sizes in the survey

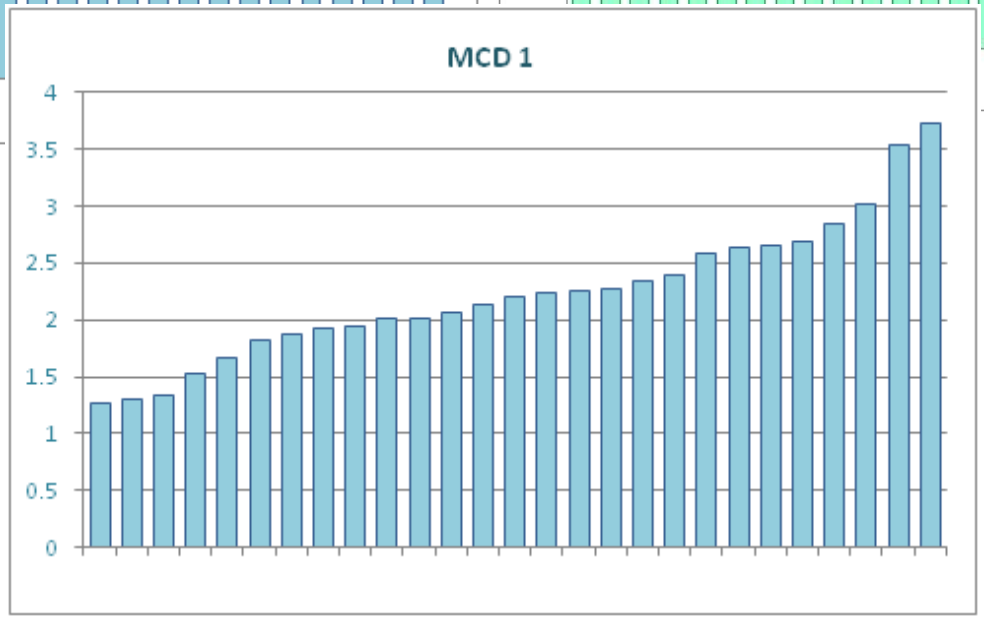


2. Quality measures vs. sample size

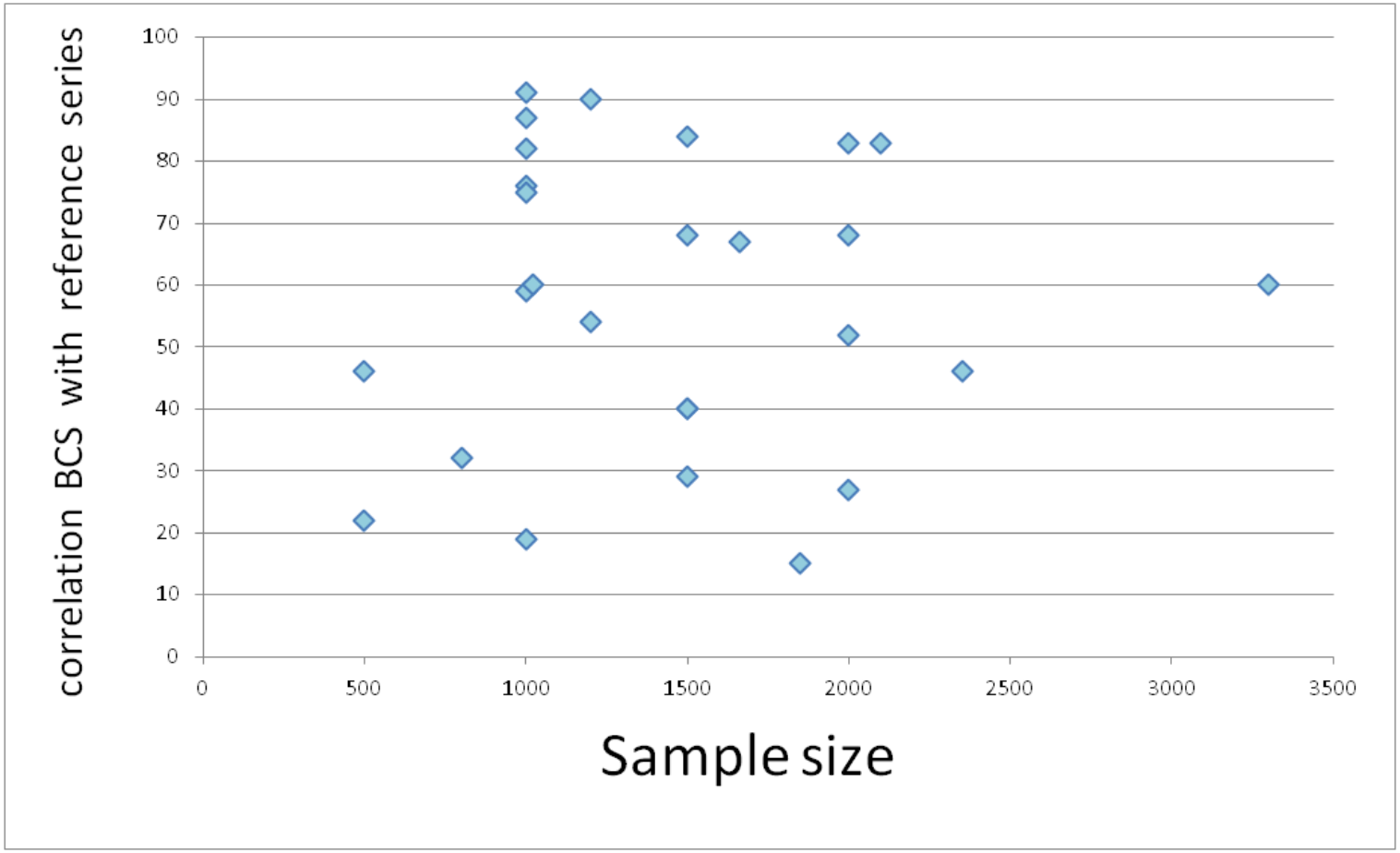
Quality measures



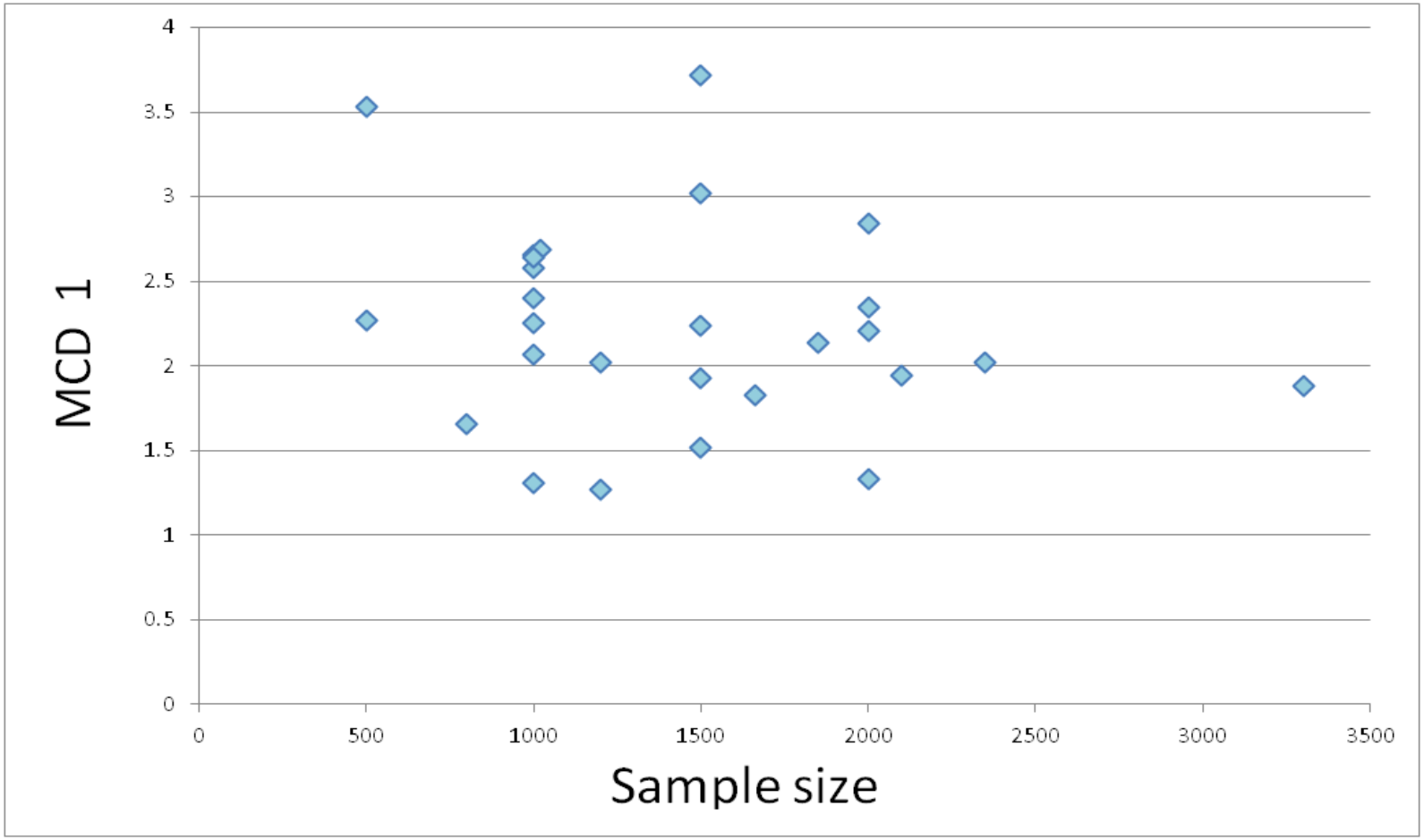




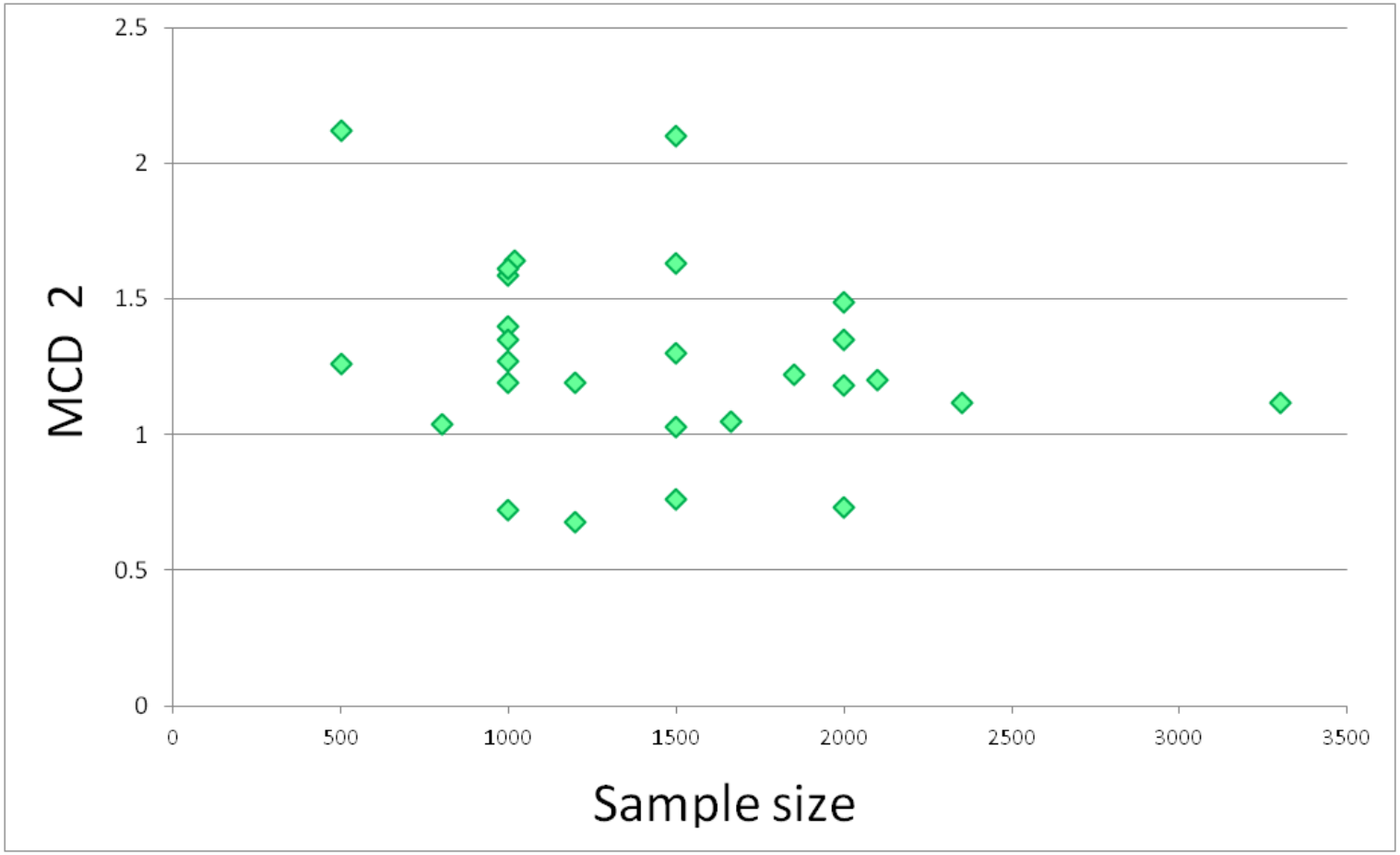
Sample size - analysis of impact on quality measures



Sample size - analysis of impact on quality measures



Sample size - analysis of impact on quality measures



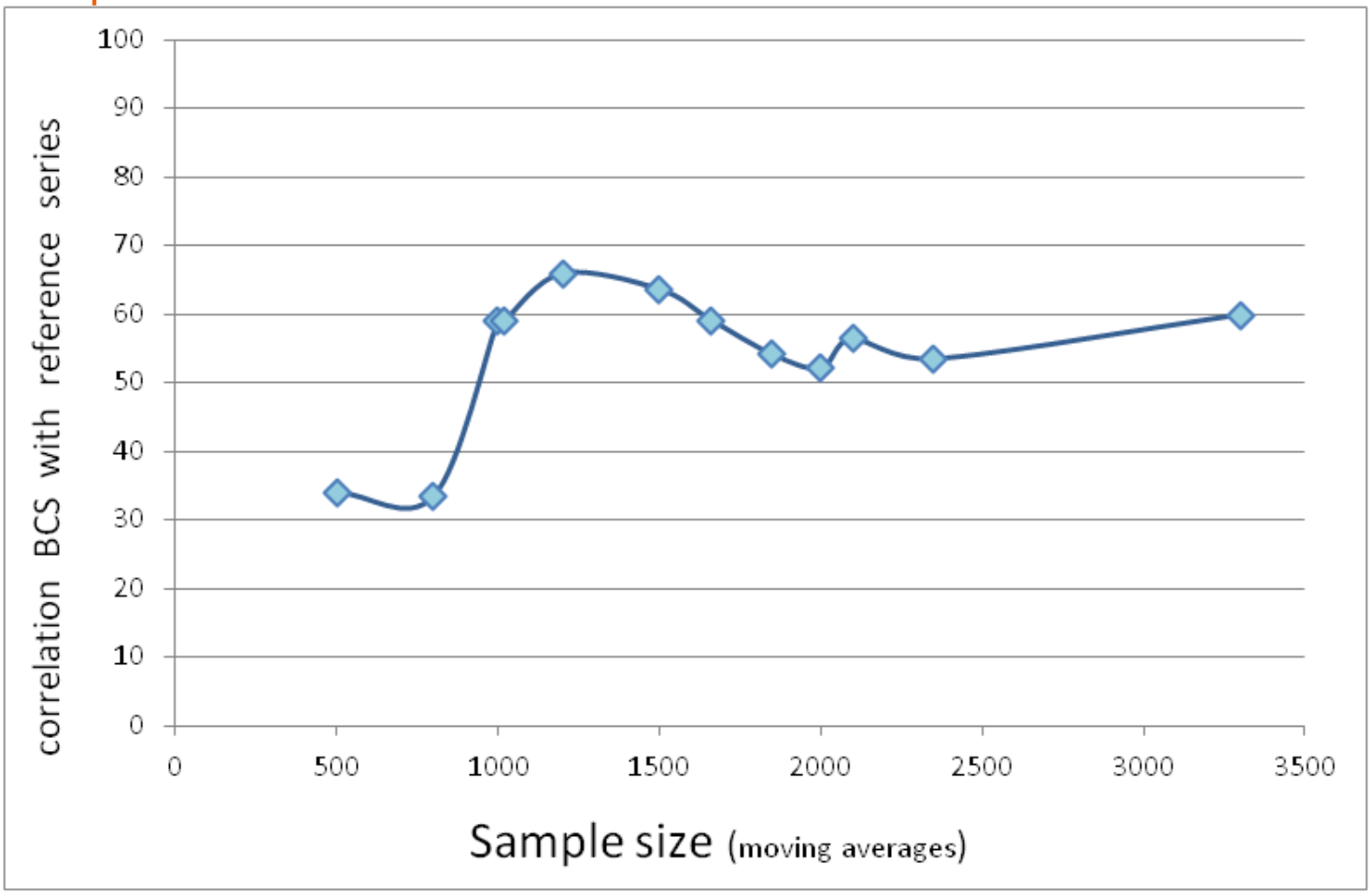
Sample size - analysis of impact on quality measures

Quality measures as moving averages for cumulated countries' samples

sample size	correlation BCS with ref.series (moving average)	MCD_1 (moving average)	MCD_2 (moving average)
500	34.0	2.90	1.69
800, 500	33.3	2.49	1.47
1000, 800, 500	58.9	2.34	1.36
1020, 1000, 800	59.0	2.37	1.38
1200, 1020, 1000, 800	65.9	2.14	1.24
1500, 1200, 1020, 1000	63.6	2.29	1.30
1661, 1500, 1200, 1020	59.1	2.25	1.26
1850, 1661, 1500, 1200	54.1	2.19	1.22
2000, 1850, 1661, 1500	52.1	2.28	1.26
2100, 2000, 1850, 1661	56.4	2.09	1.17
2350, 2100, 2000, 1850	53.4	2.12	1.18
3300, 2350, 2100, 2000	59.9	2.08	1.17

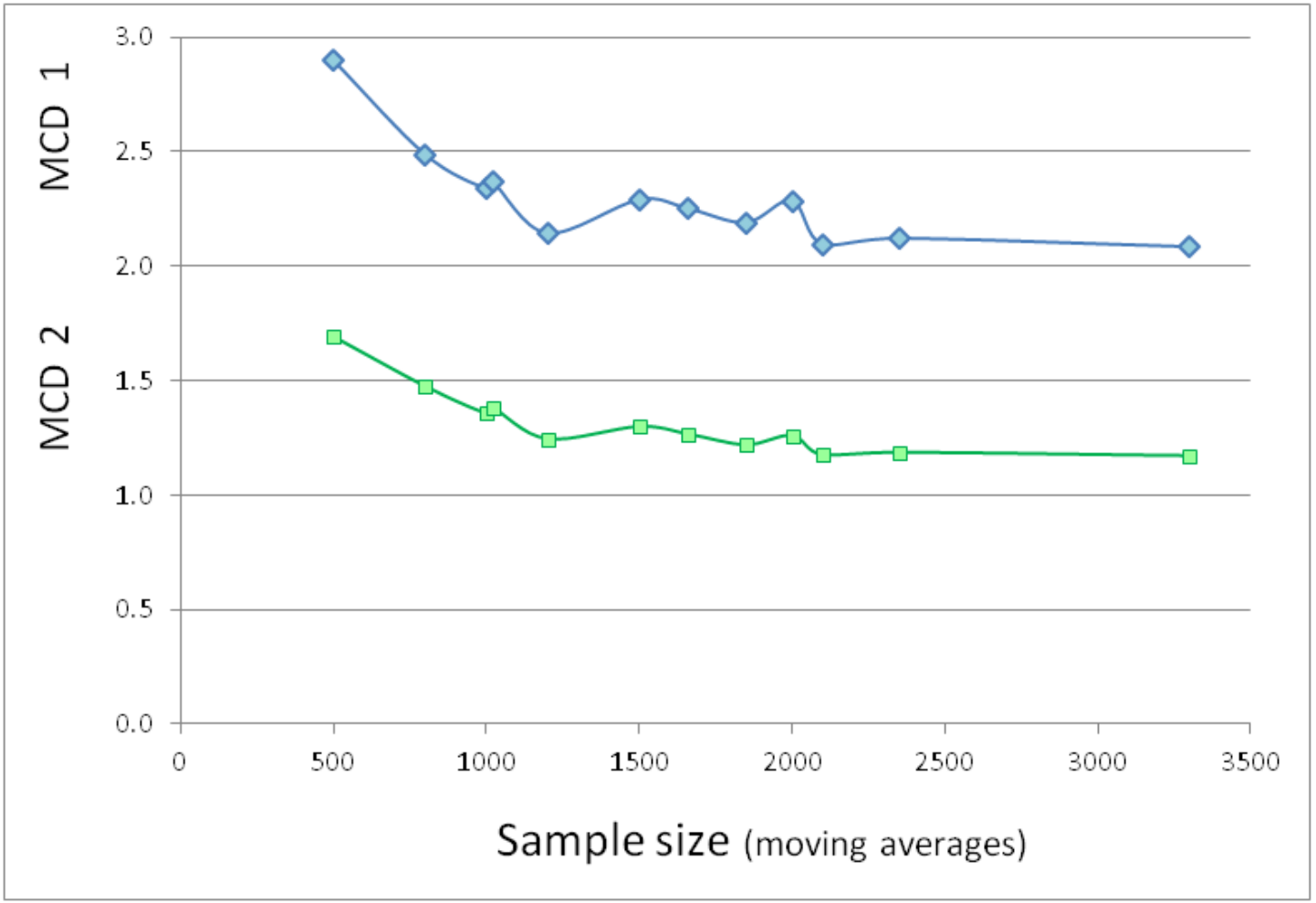
Sample size - analysis of impact on quality measures

Tendency of changes in the quality of data along with the increase of sample size

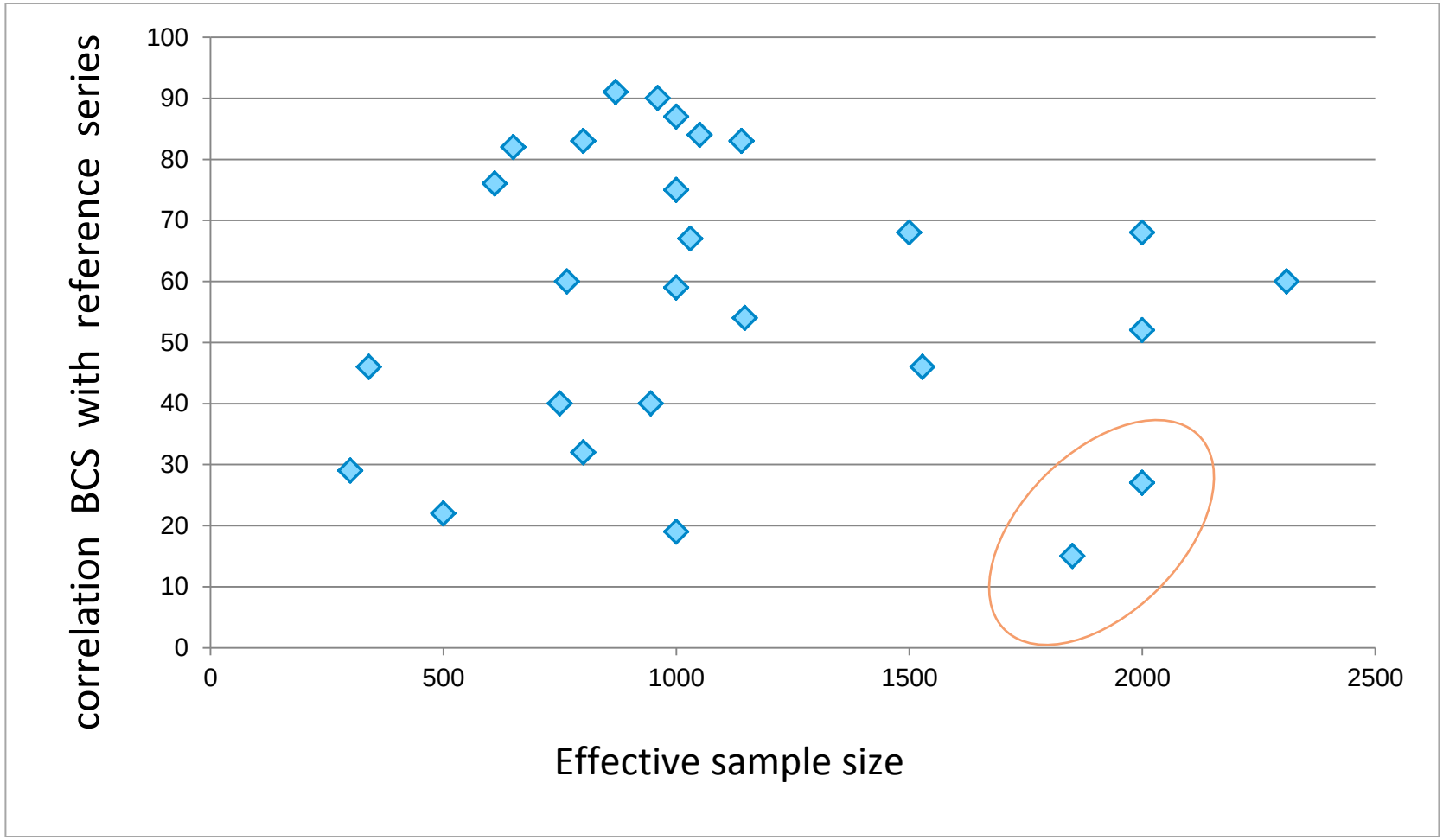


Sample size - analysis of impact on quality measures

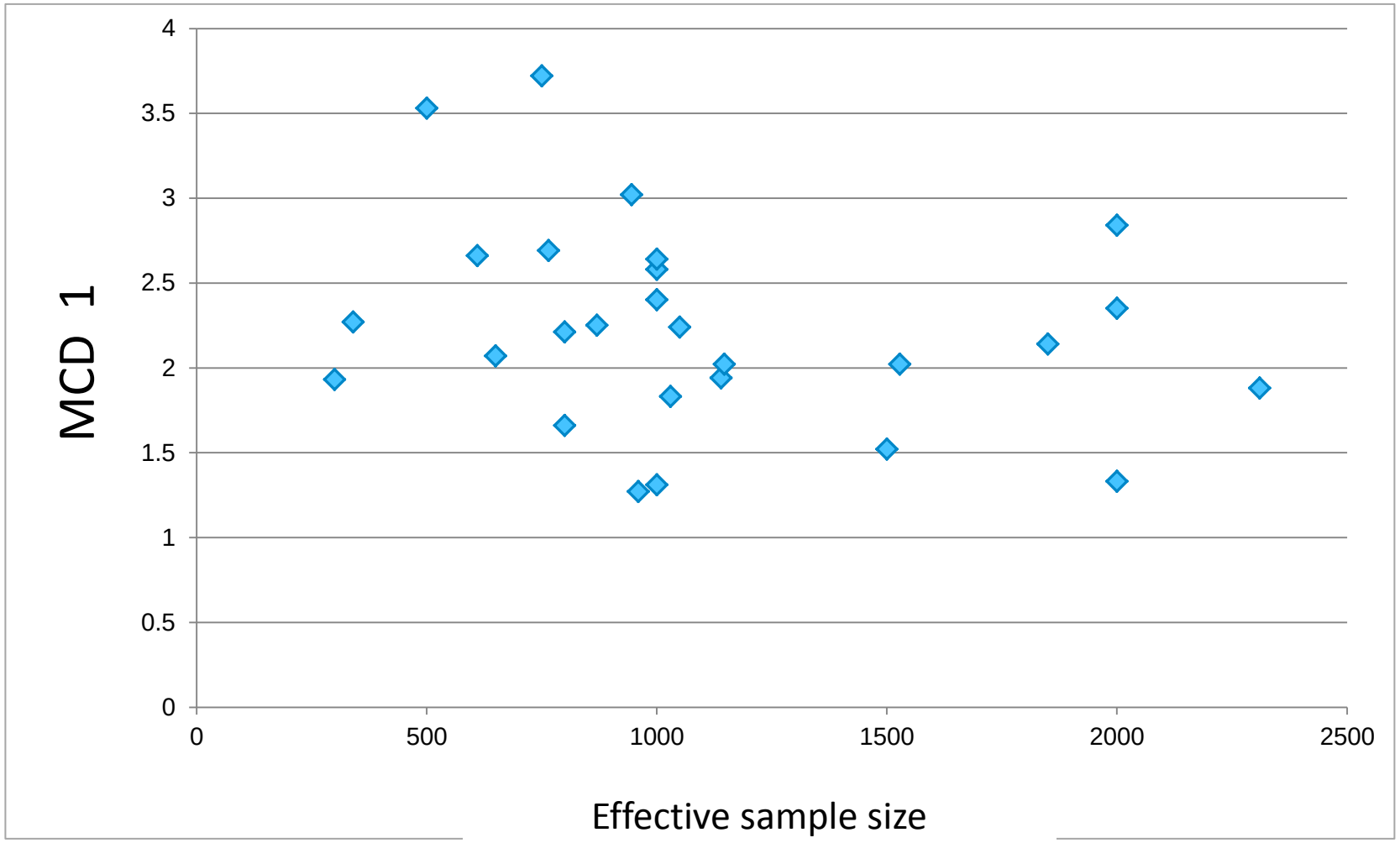
Tendency of changes in the quality of data along with the increase of sample size



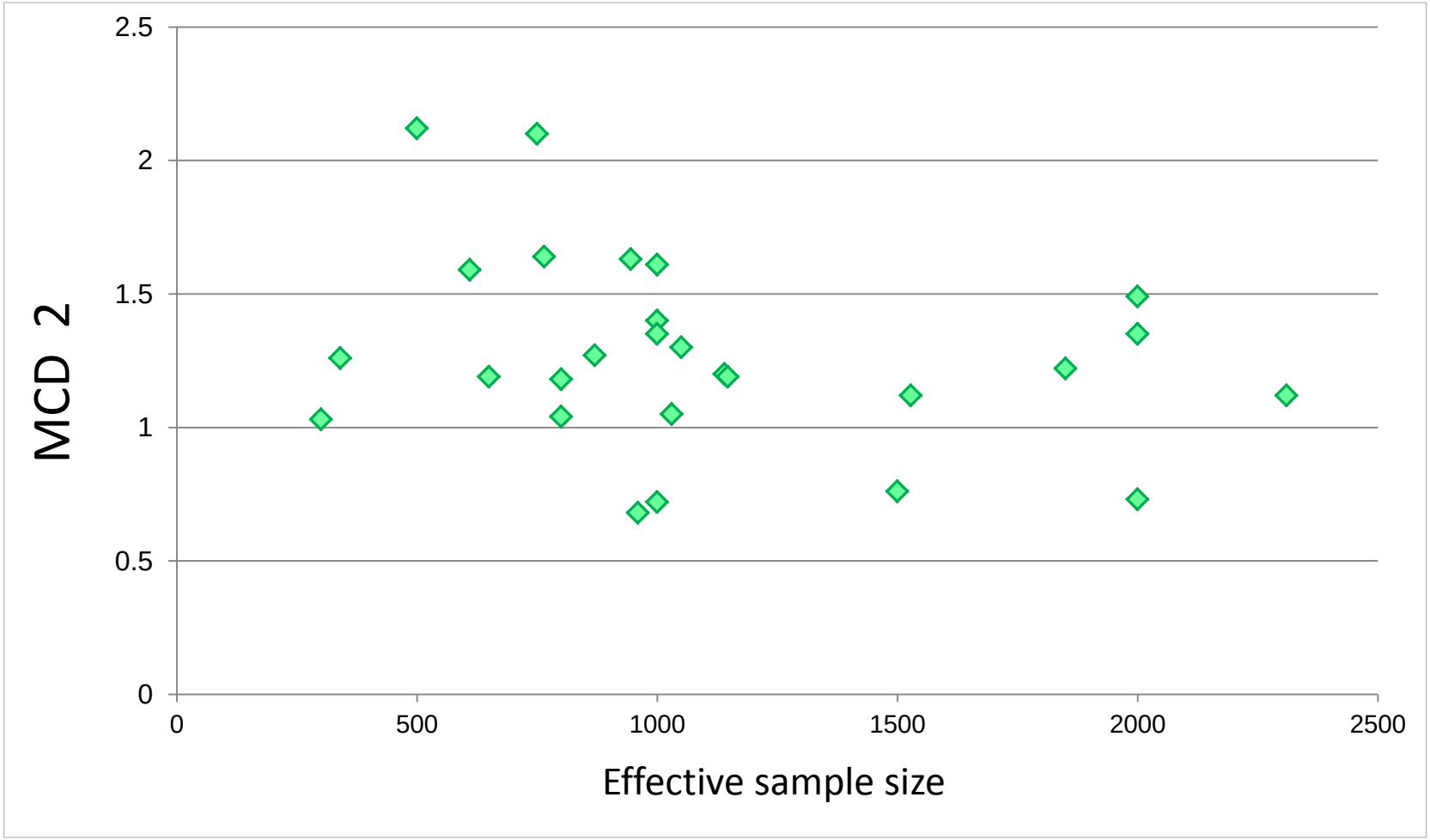
Effective sample size - analysis of impact on quality measures



Effective sample size - analysis of impact on quality measures

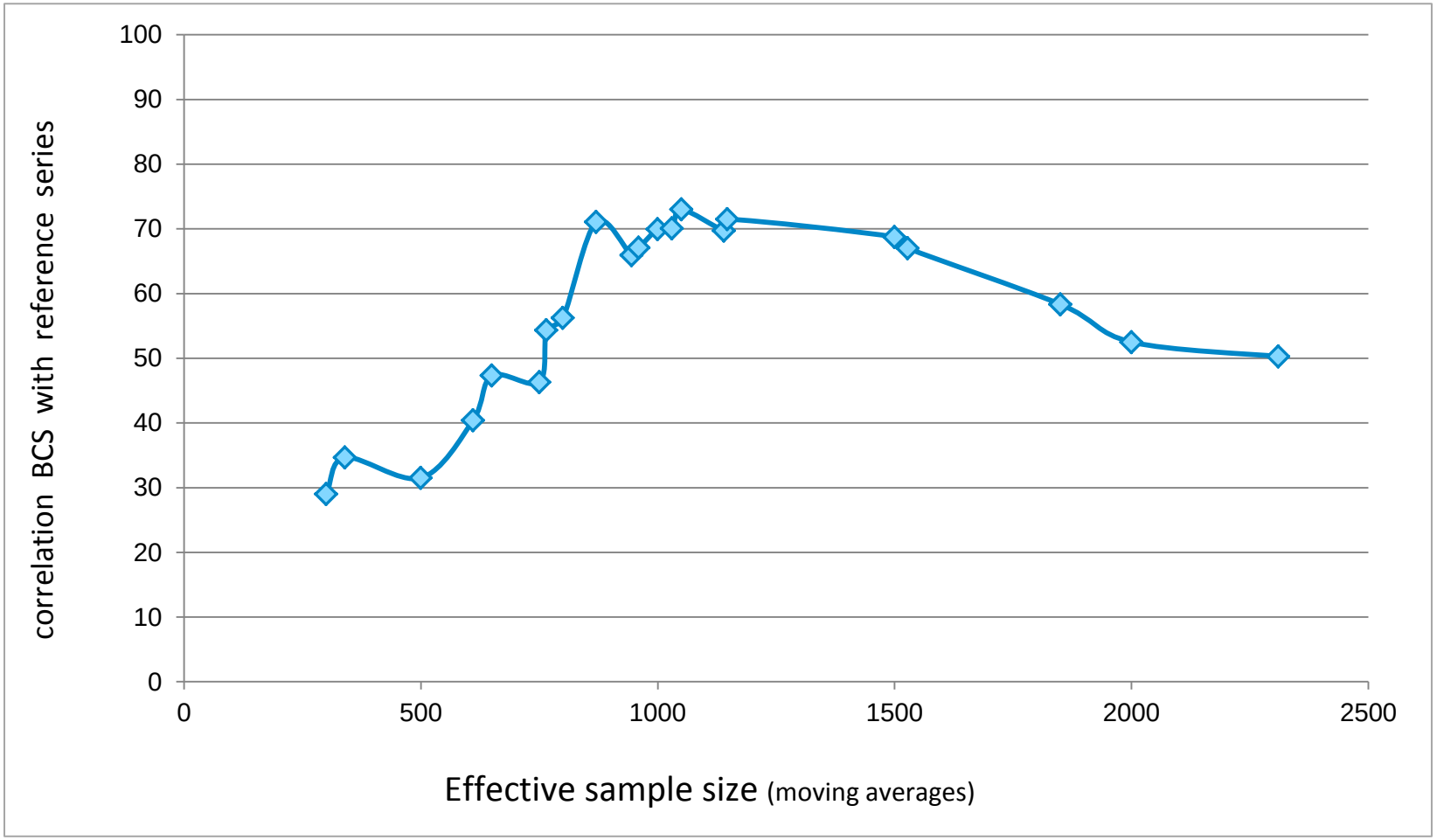


Effective sample size - analysis of impact on quality measures



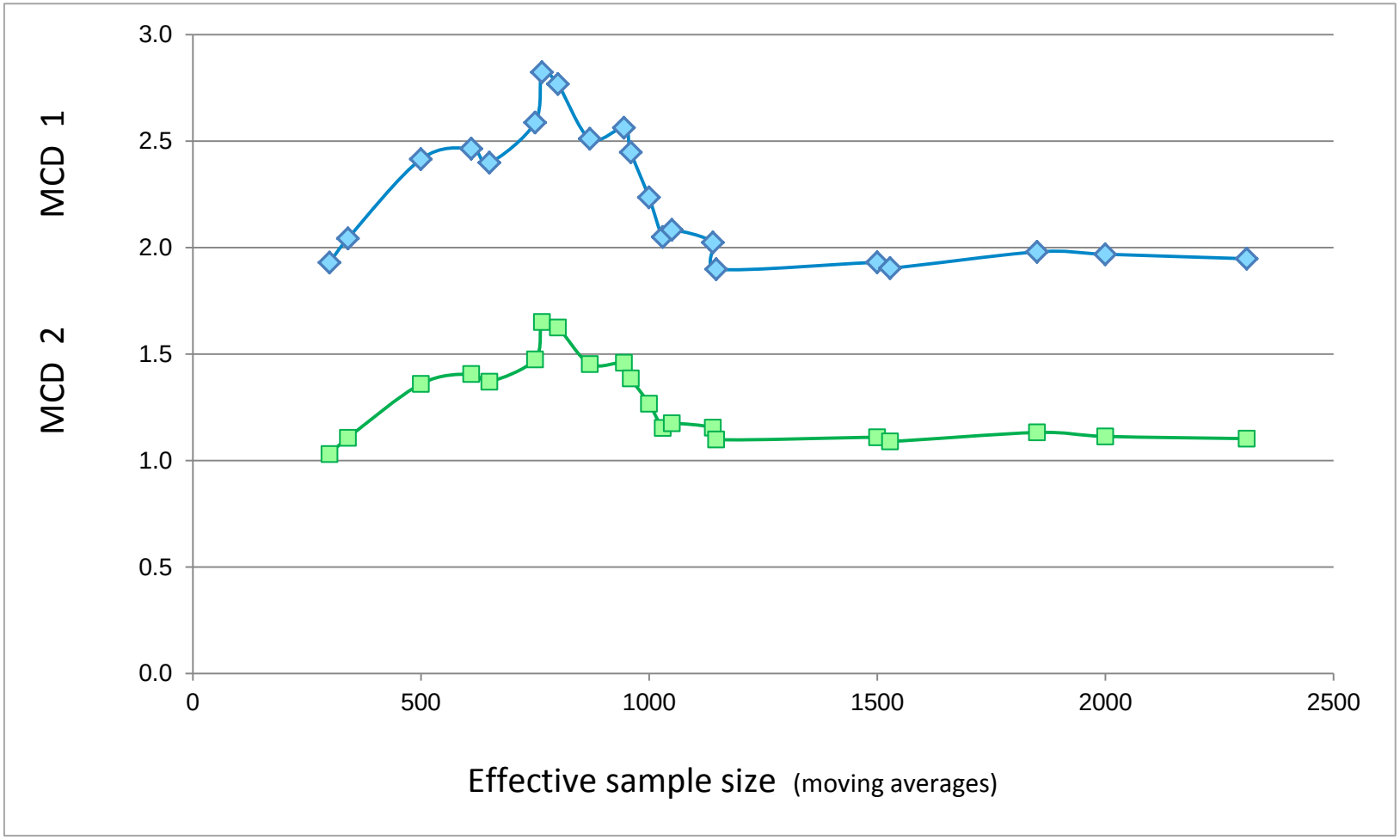
Effective sample size - analysis of impact on quality measures

Tendency of changes in the quality of data along with the increase of effective sample size



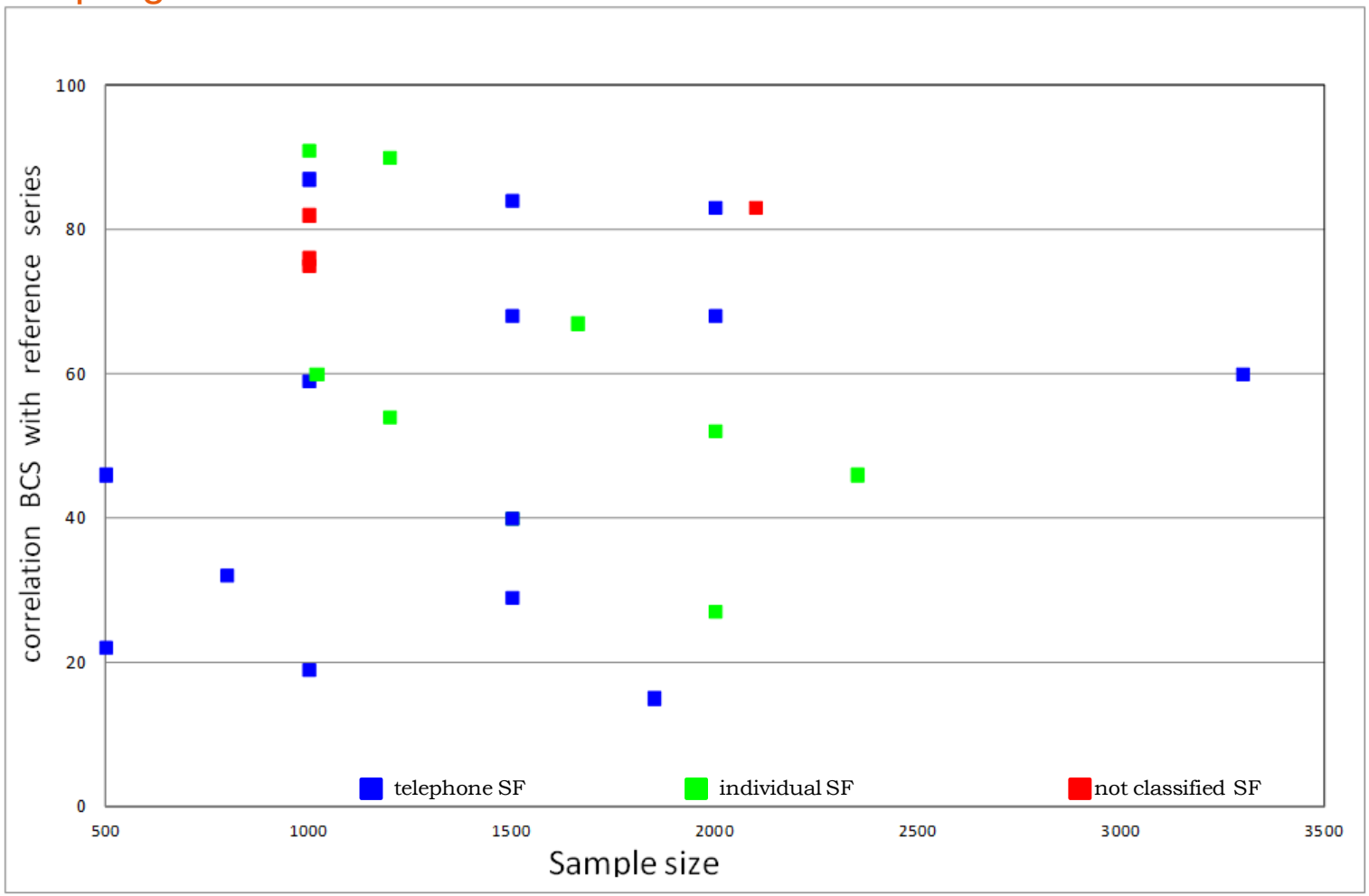
Effective sample size - analysis of impact on quality measures

Tendency of changes in the quality of data along with the increase of effective sample size



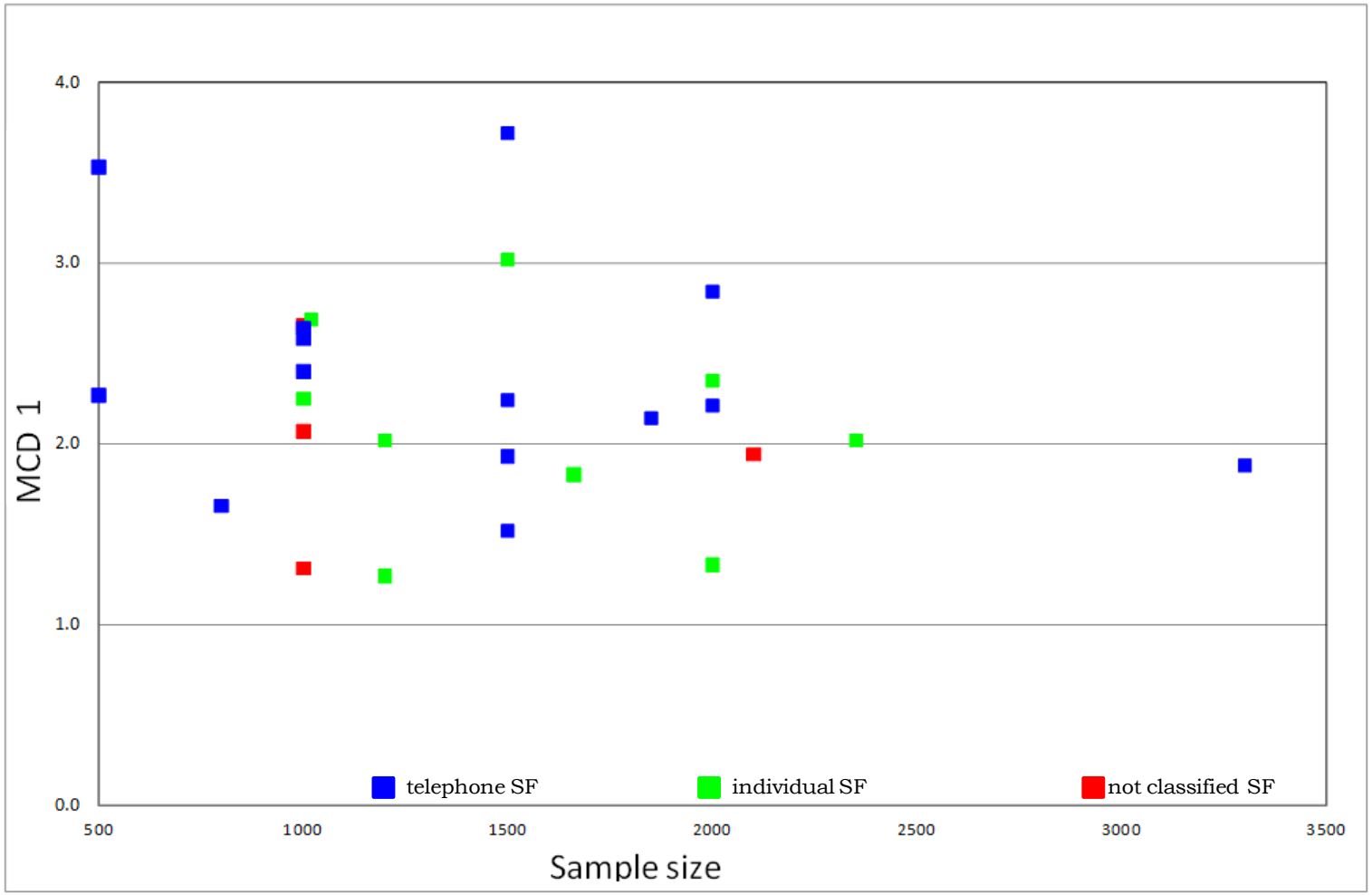
Additional analysis

Sample size and correlation BCS with reference series for 3 types of sampling frames



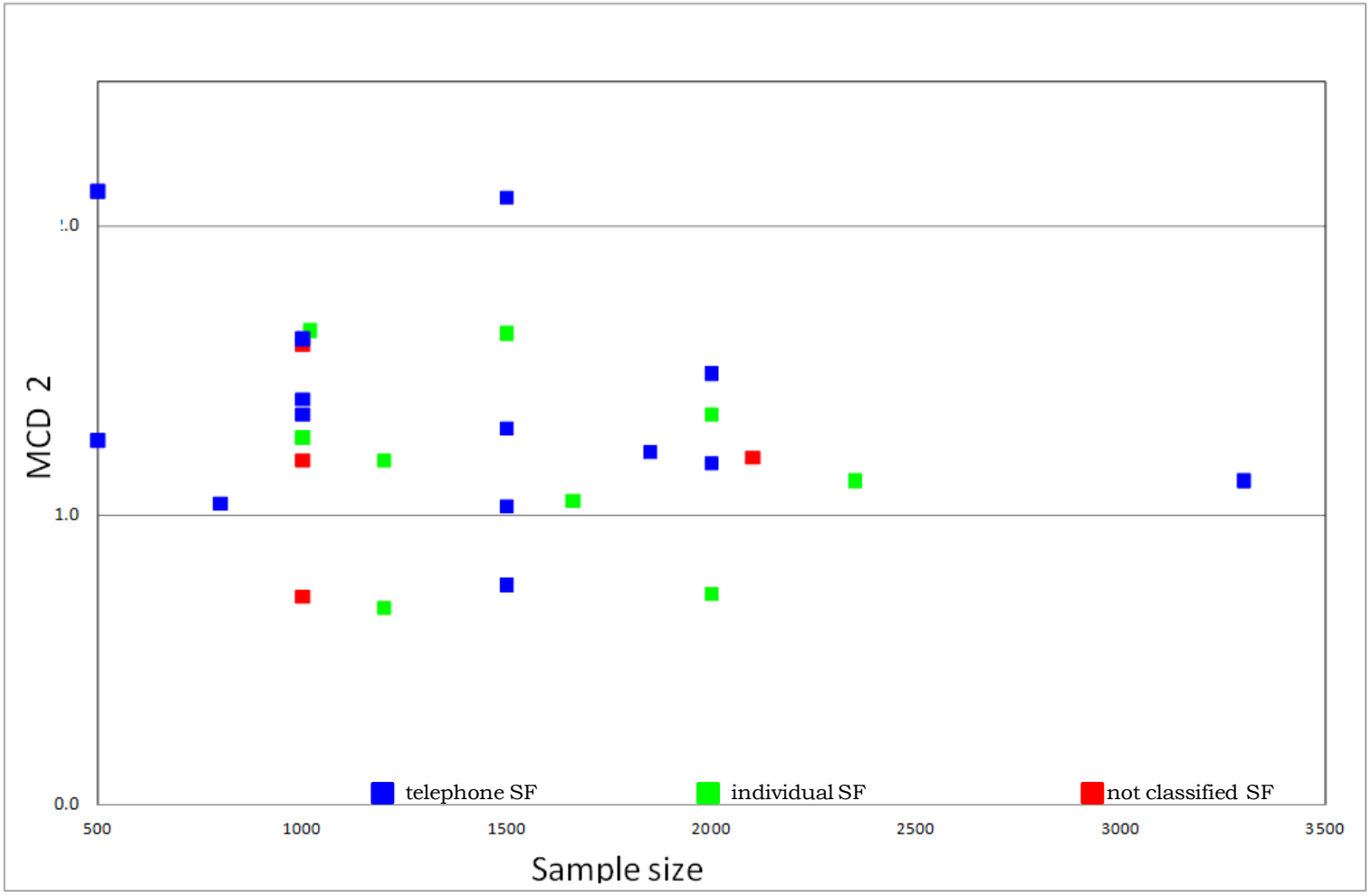
Additional analysis

Sample size and MCD 1 for 3 types of sampling frames



Additional analysis

Sample size and MCD 2 for 3 types of sampling frames



3. Summary and conclusions

Main findings

- Both sample size and effective sample size have moderate impact on quality measures
- Although tendencies are slight they are quite clear and coherent for all MCD1 and MCD2 measures. Picture for correlation is not so clear.

- Effect of sample size on data volatility and consistency with reference series is clear as tendency
- Impact of effective sample size on MCD1, MCD2 is achieved also as a trend, and for the major part of the ESS range – rather strong
- the ESS impact on correlation BCS with reference series is also clear except the behavior of few biggest samples. It should be probably analyzed in connection with other factors that influence the quality

Recommendations

- Revise the way of determining the sample size in the survey
- Verify the sampling error based on conducted sample size
- Consider possibility of increase of sample size in some countries
- Compare similar sample sizes and their likely effects on data quality

4. References

References

1. Glossary of Statistical Terms OECD

2. Särndal, Carl-Erik; Swensson, Bengt; Wretman, (2003). *Model assisted survey sampling*. Springer. (Retrieved 2 2011)

3. Israel, Glen D. (1992) *Determining Sample Size*, Agricultural Education and Communication Department, University of Florida, IFAS Extension, PEOD6 (Reviewed 2013)

4. Turner, Anthony G. *Sampling frames and master samples*. United Nations Secretariat. Retrieved 12/11/2012.

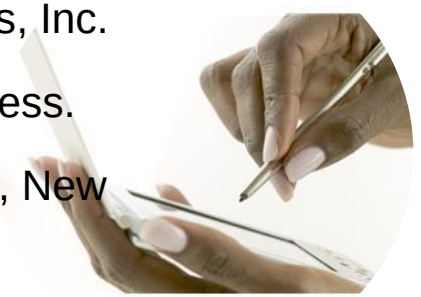
5. Data source: *Metadata_checked_by_partners - Consumers.xlsx*

6. Cochran, W. G. (1963). *Sampling Techniques*, 2nd Ed., New York: John Wiley and Sons, Inc.

7. Kish, Leslie. (1965). *Survey Sampling*. New York: John Wiley and Sons, Inc.

8. Sudman, Seymour. (1976). *Applied Sampling*. New York: Academic Press.

9. Yamane, Taro. (1967). *Statistics: An Introductory Analysis*, 2nd Edition, New York: Harper and Row.



THANK YOU FOR YOUR ATTENTION