

# Analysis of sampling frames in consumer surveys

GfK Polonia, 2013, Poland

# Agenda

1. Introduction
  - a. Definitions
  - b. Common problems with sampling frames
  - c. Telephone-based sampling frames – drawbacks and benefits
  - d. Sampling frames used in the DG ECFIN Consumer Survey
  
2. Quality measures vs. descriptive features of sampling frames
  - a. Type of units in sampling frames - analysis of impact
  - b. Update frequency of sampling frames - analysis of impact
  - c. Sampling frames coverage - analysis of impact
  - d. Additional analysis
  
3. Analyses summary and conclusions
  
4. References.

# 1. Introduction

## Definitions

Sampling frame is a list of all members of a population used as a basis for survey design

two types of frames:

- **list frames** - a list of all the units in the survey target population (e.g. administrative lists, the electoral roll)
- **area frames** - a complete and exhaustive list of non- overlapping geographic areas

# Definitions

## Requirements for sampling frames

- well- organized in a logical, systematic fashion
- all units should be accessible
- "up-to-date"
- every element of the population is present in the frame
- every element of the population is present only once
- no elements from outside the population are present in the frame

## Common problems with sampling frames

- Coverage errors
- Out-of-scope units
- Incompleteness

# Telephone-based sampling frames

## Drawbacks

- Non-telephone households / respondents
- Mobile only or landline only households /respondents
- Duplication
- Non-household target population
- Out-of-date

# Sampling frames used in the DG ECFIN Consumer Survey

	Number of countries
Census-based sampling frames	9
Telephone-based sampling frames	17
No Sampling Frame	3
Polling Stations Territory	2



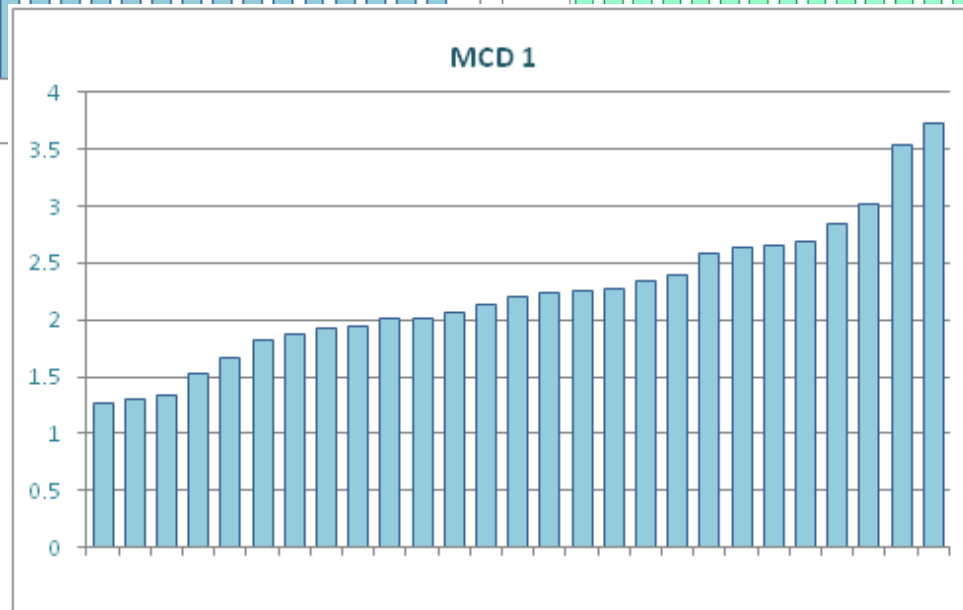
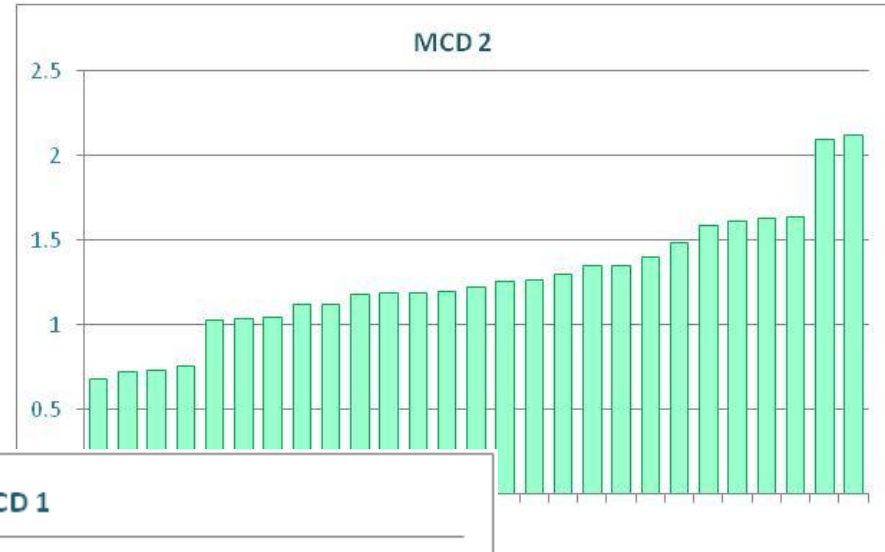
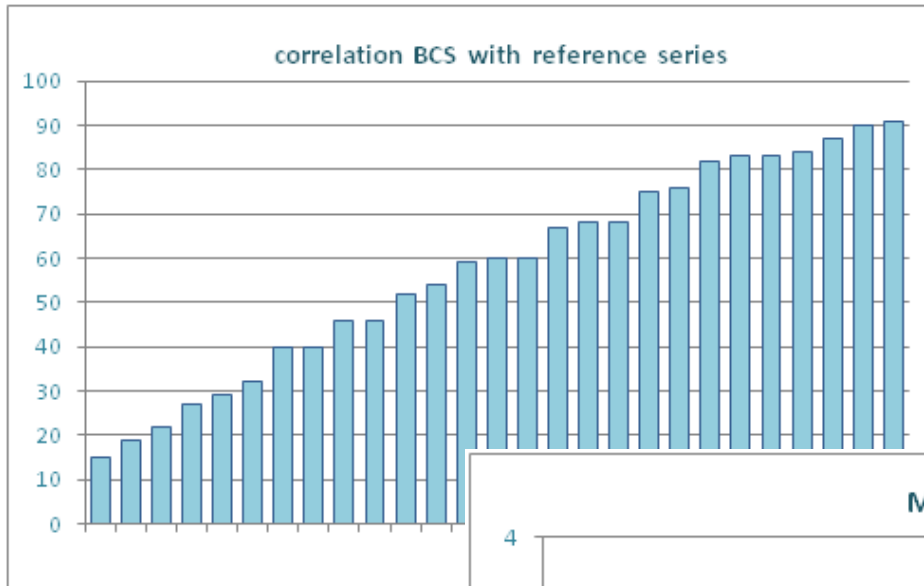
## Sampling frames used in the DG ECFIN Consumer Survey

Mobile and landline telephone numbers used in **realization** of the Consumer Survey

	Frequency
fix only	3
mobile & fix	8
Unspecified, telephone-based SF	6
Unspecified, nontelephone-based SF	6
Doesn't apply	8
<b>Total</b>	<b>31</b>

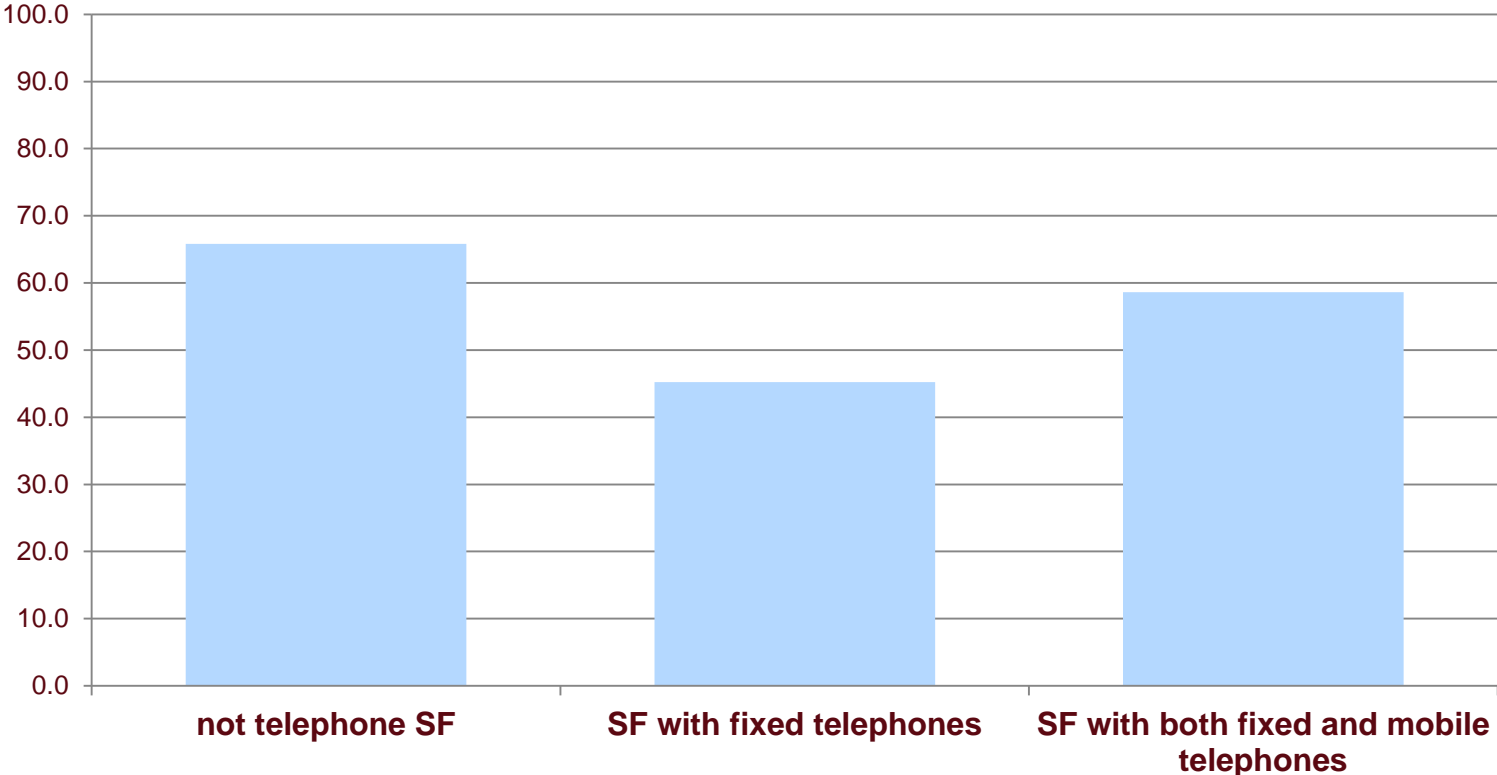
## 2. Quality measures vs. sampling frames

# Quality measures

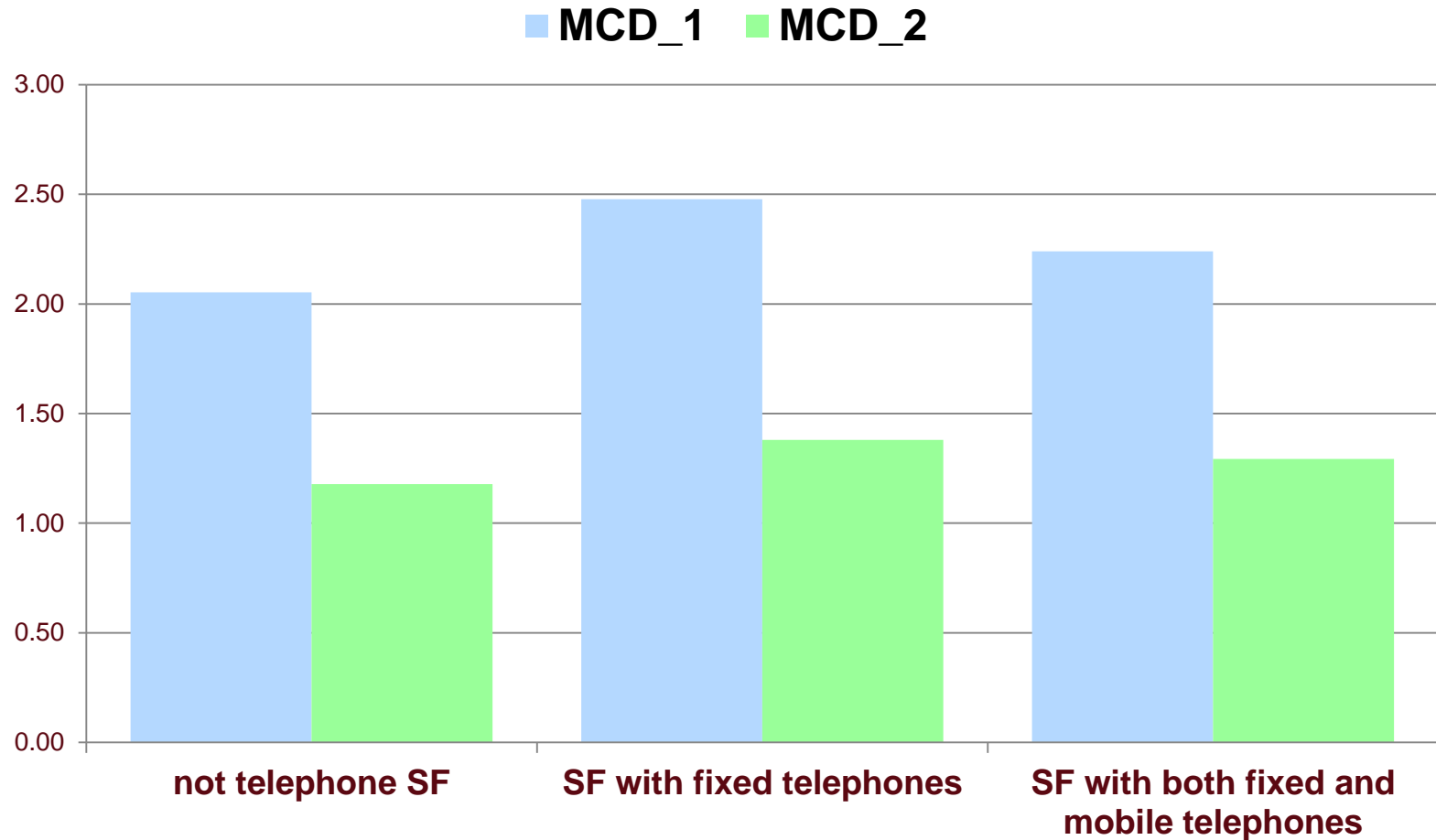


# Type of units in sampling frames - analysis of impact

**correlation BCS with reference series**

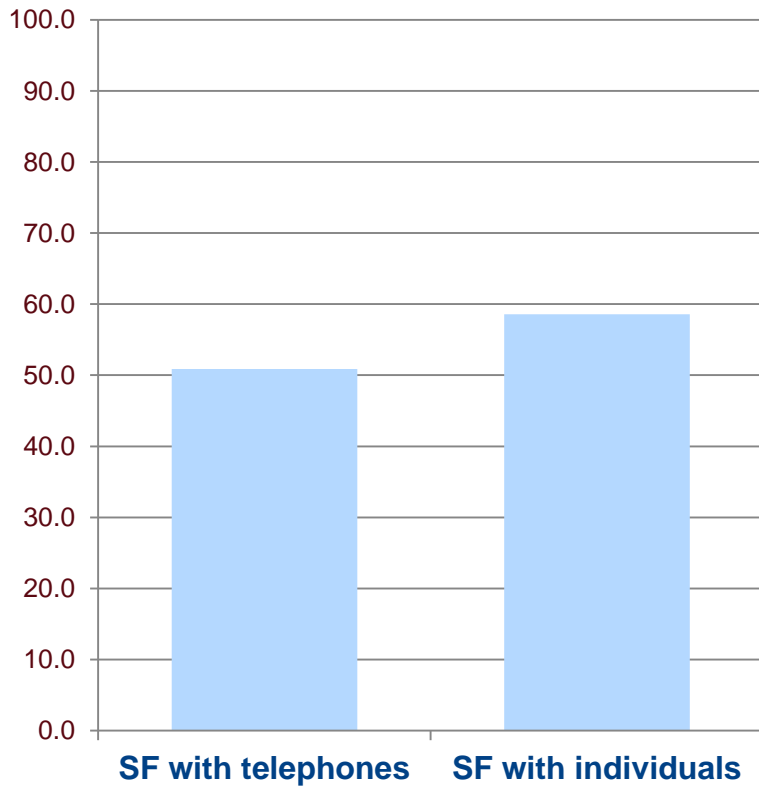


# Type of units in sampling frames - analysis of impact

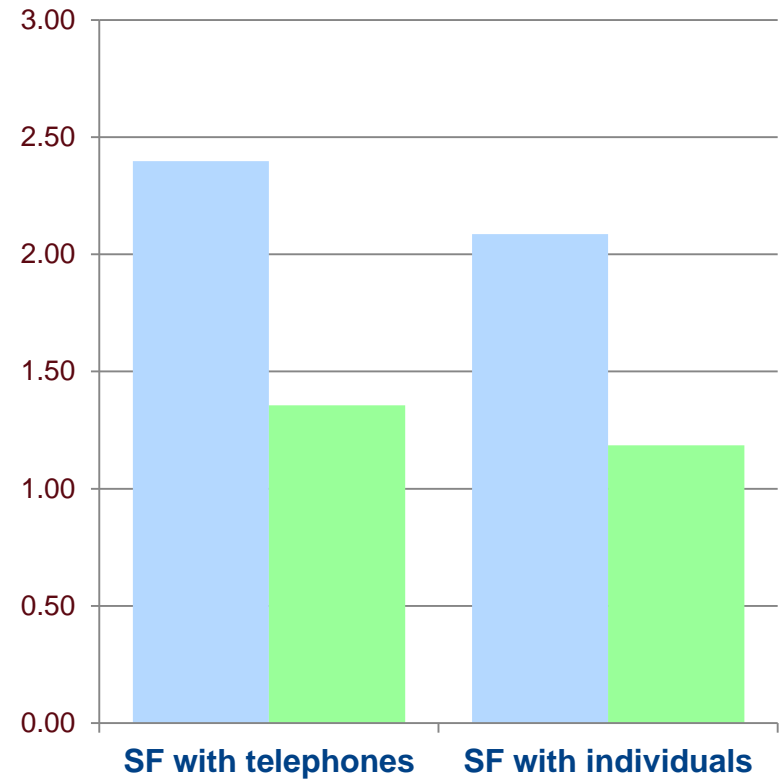


# Type of units in sampling frames - analysis of impact

correlation BCS with reference series



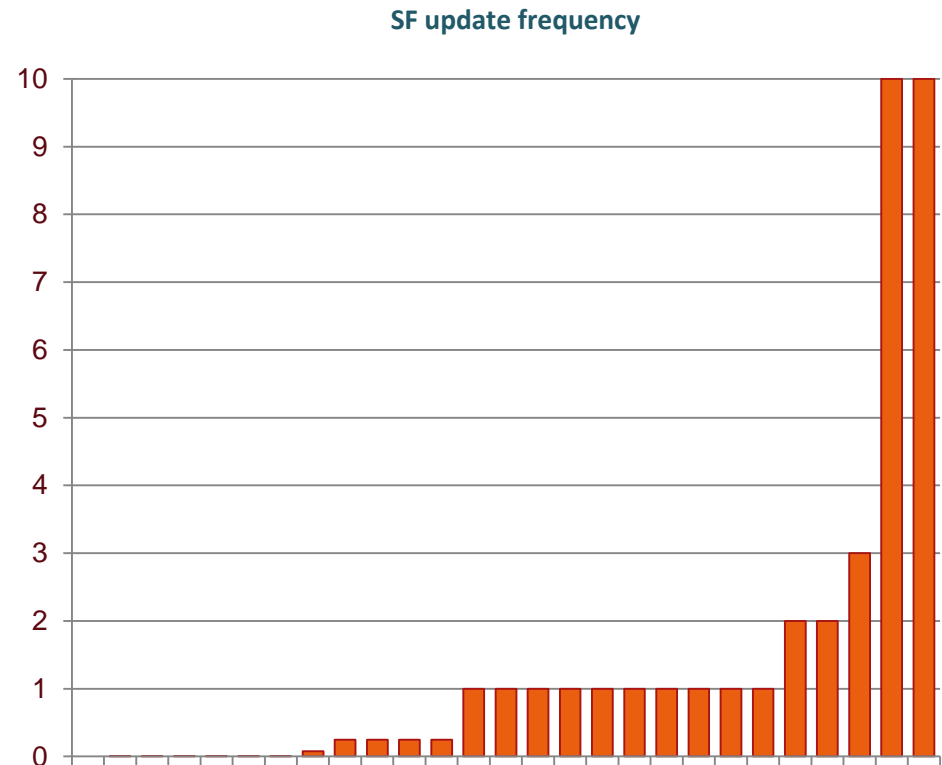
MCD\_1 MCD\_2



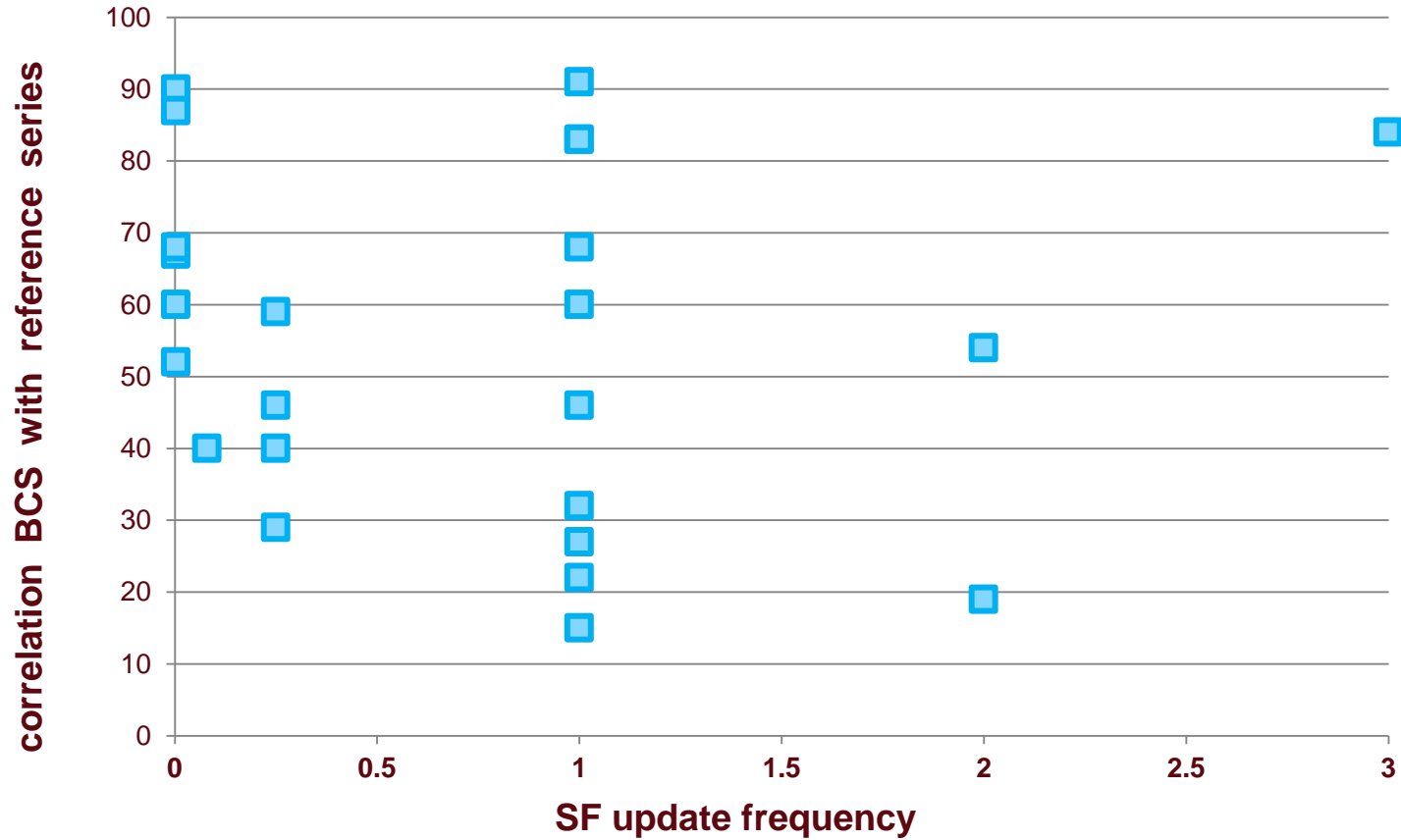
# Update frequency of sampling frames - analysis of impact

## Update frequency categories

- 0.003 – daily/ continuous update
- 0.080 – monthly update
- 0.250 – quarterly update
- 1.000 – annual update
- 2.000 – update every two years
- 3.000 – update every three years
- 10.000 – update every ten years

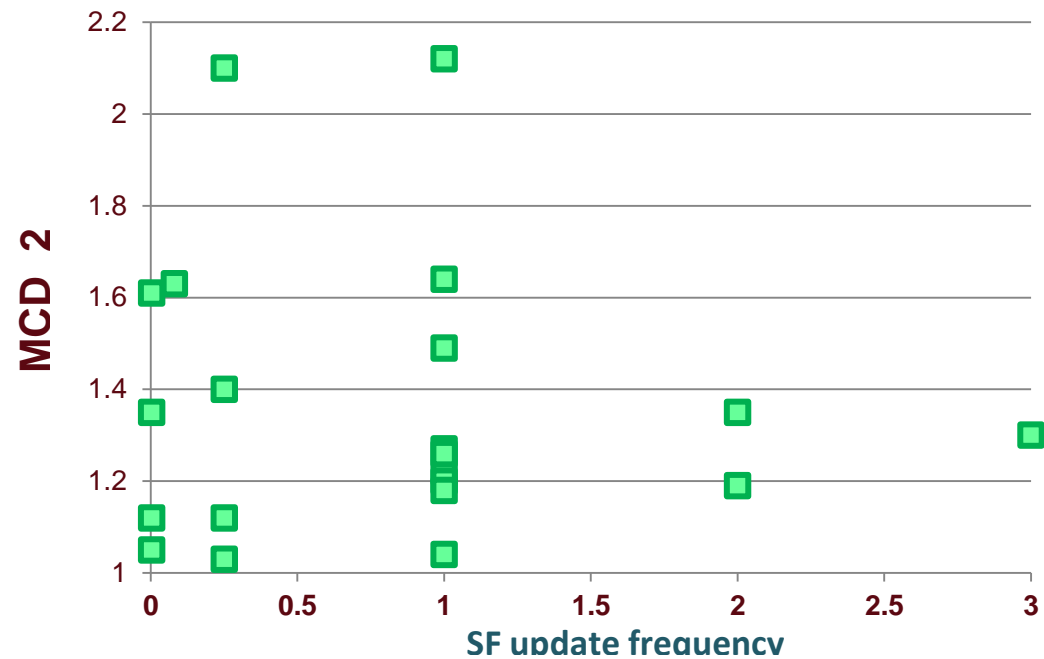
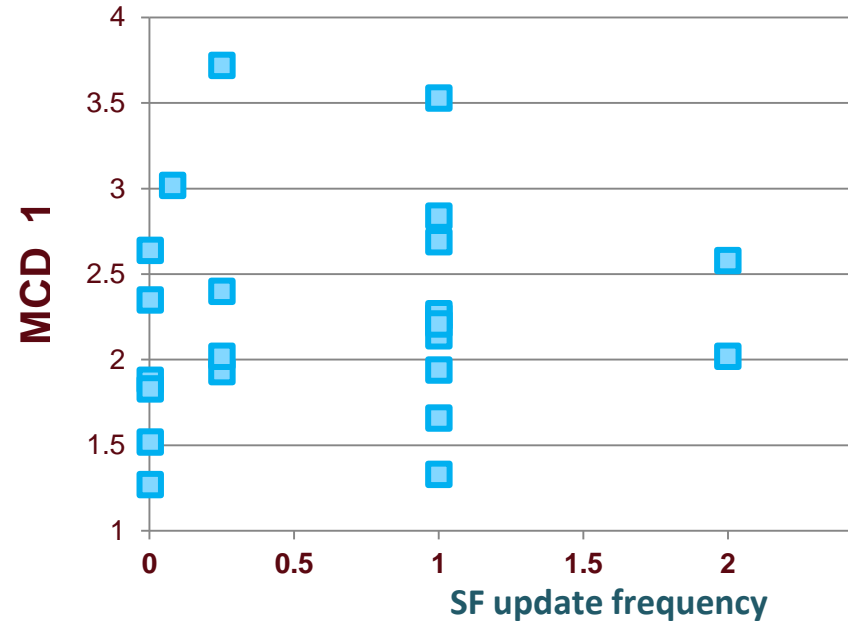


# Update frequency of sampling frames - analysis of impact

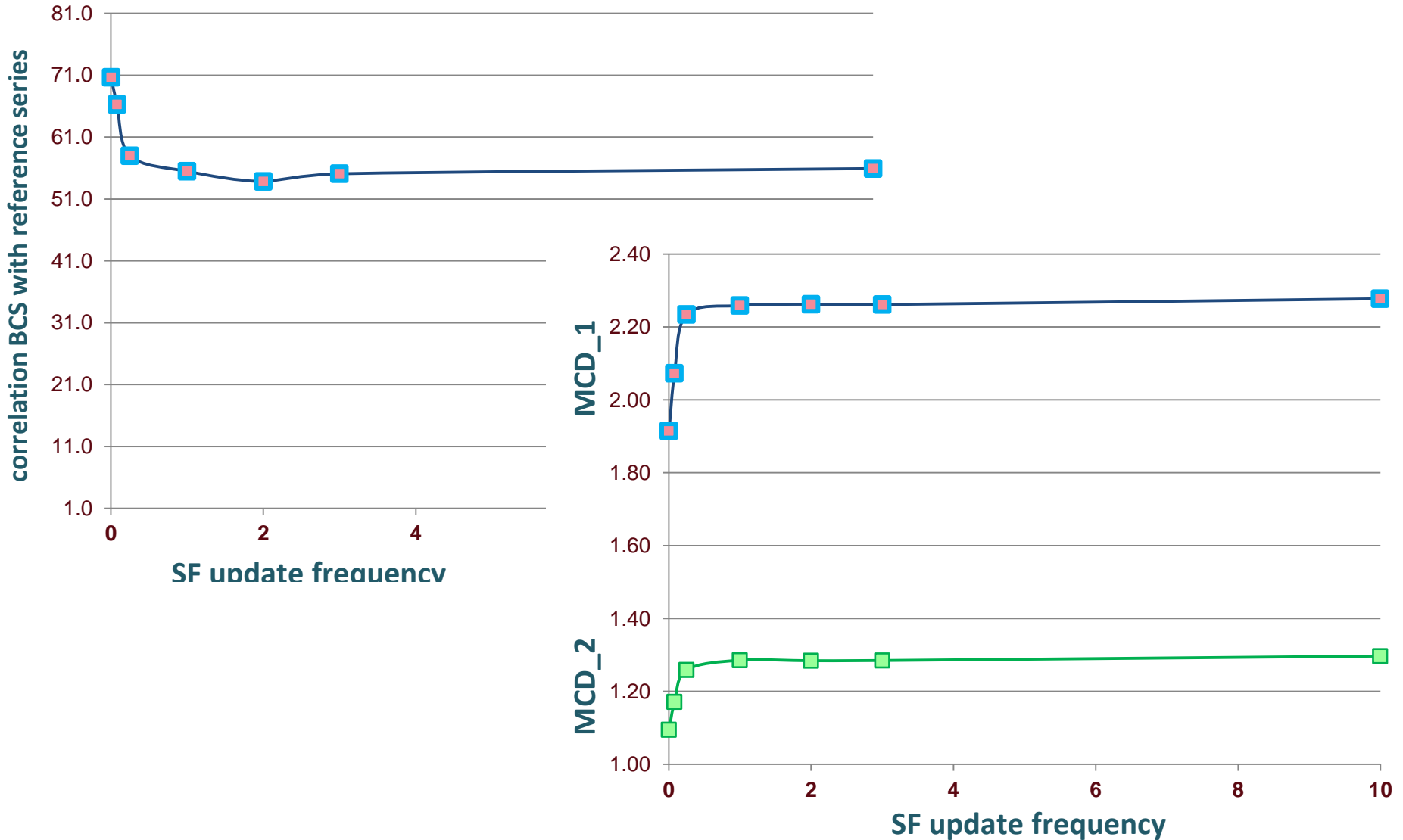




# Update frequency of sampling frames - analysis of impact

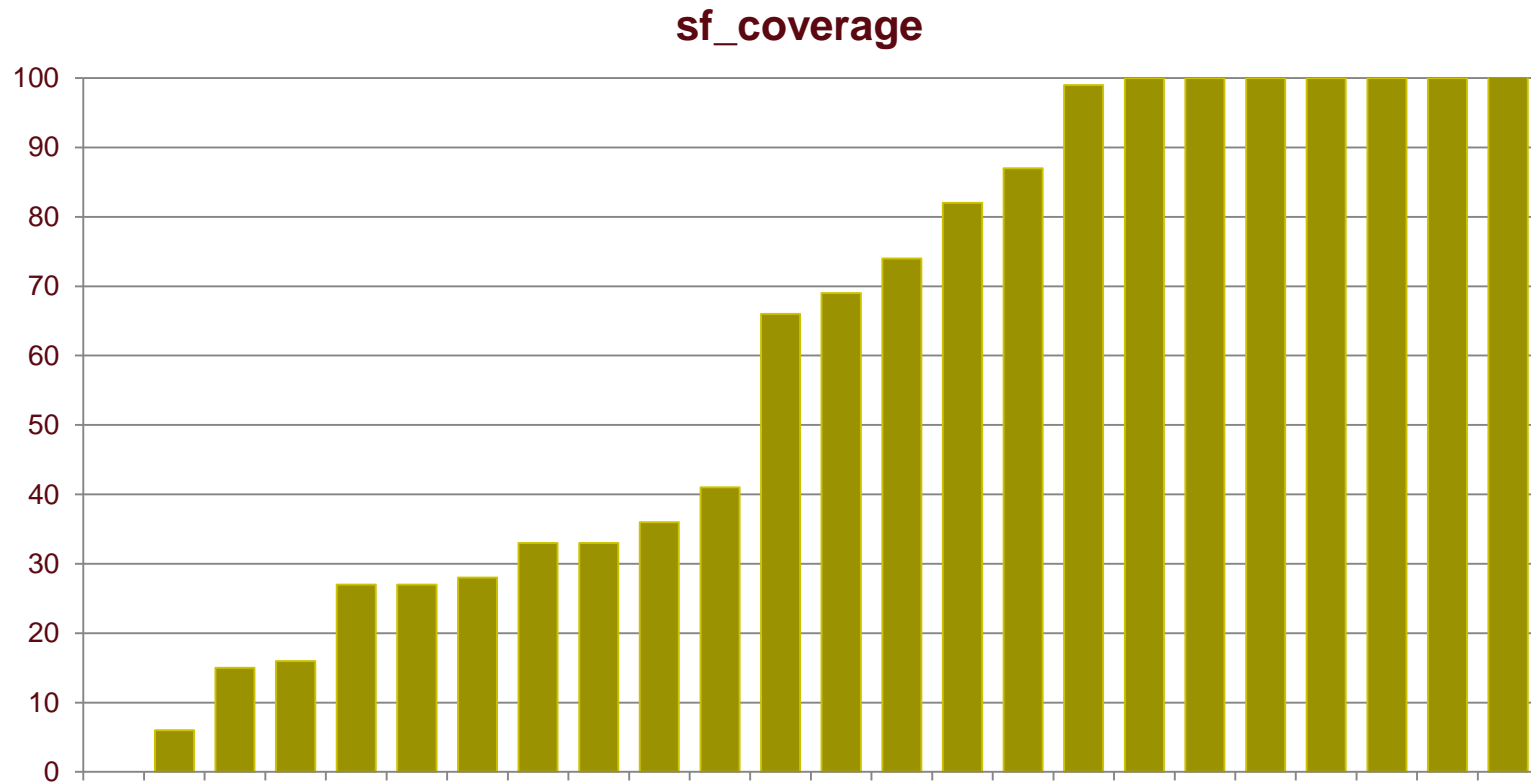


# Update frequency of sampling frames - analysis of impact

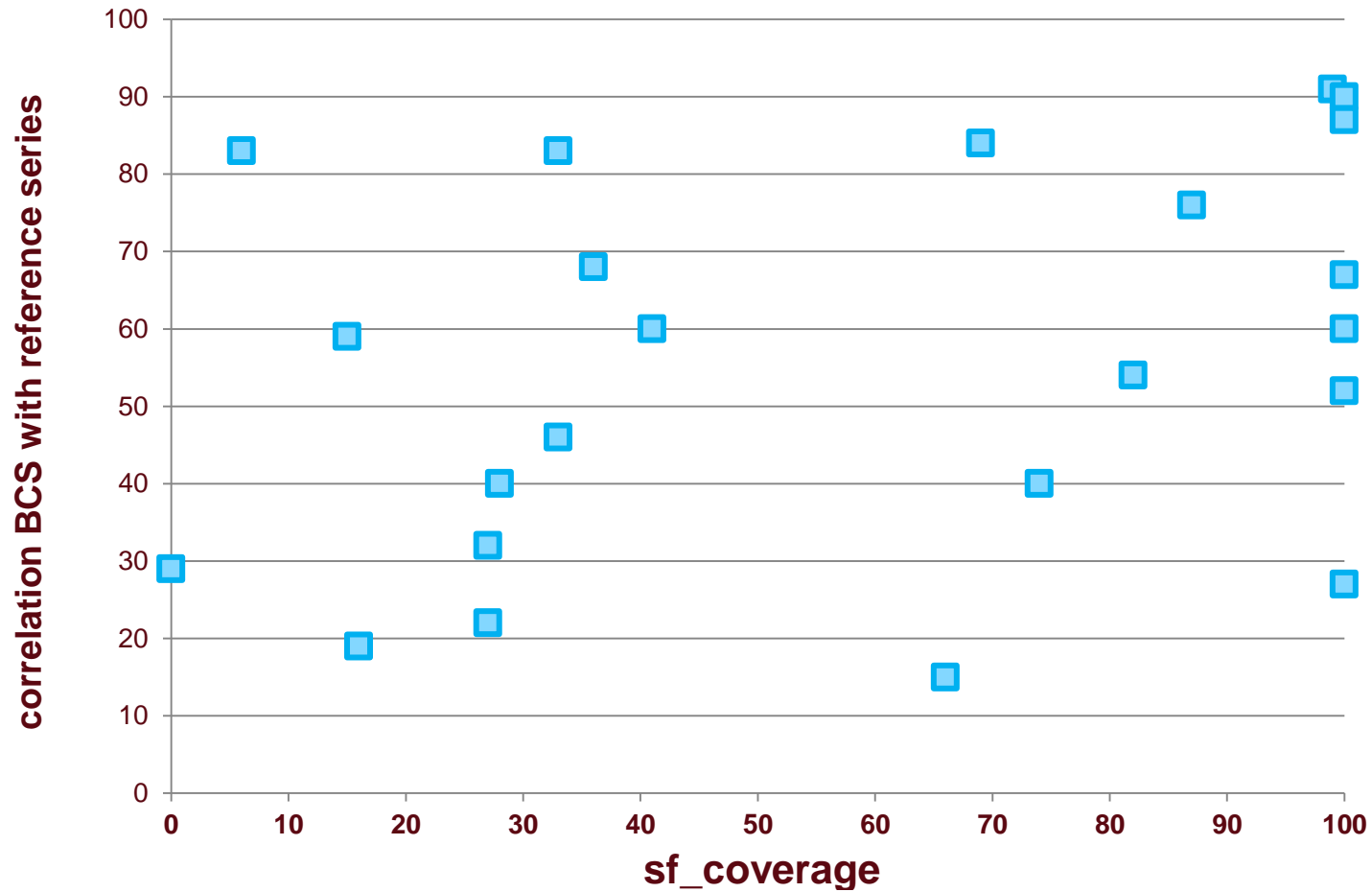


# Sampling frames coverage - analysis of impact

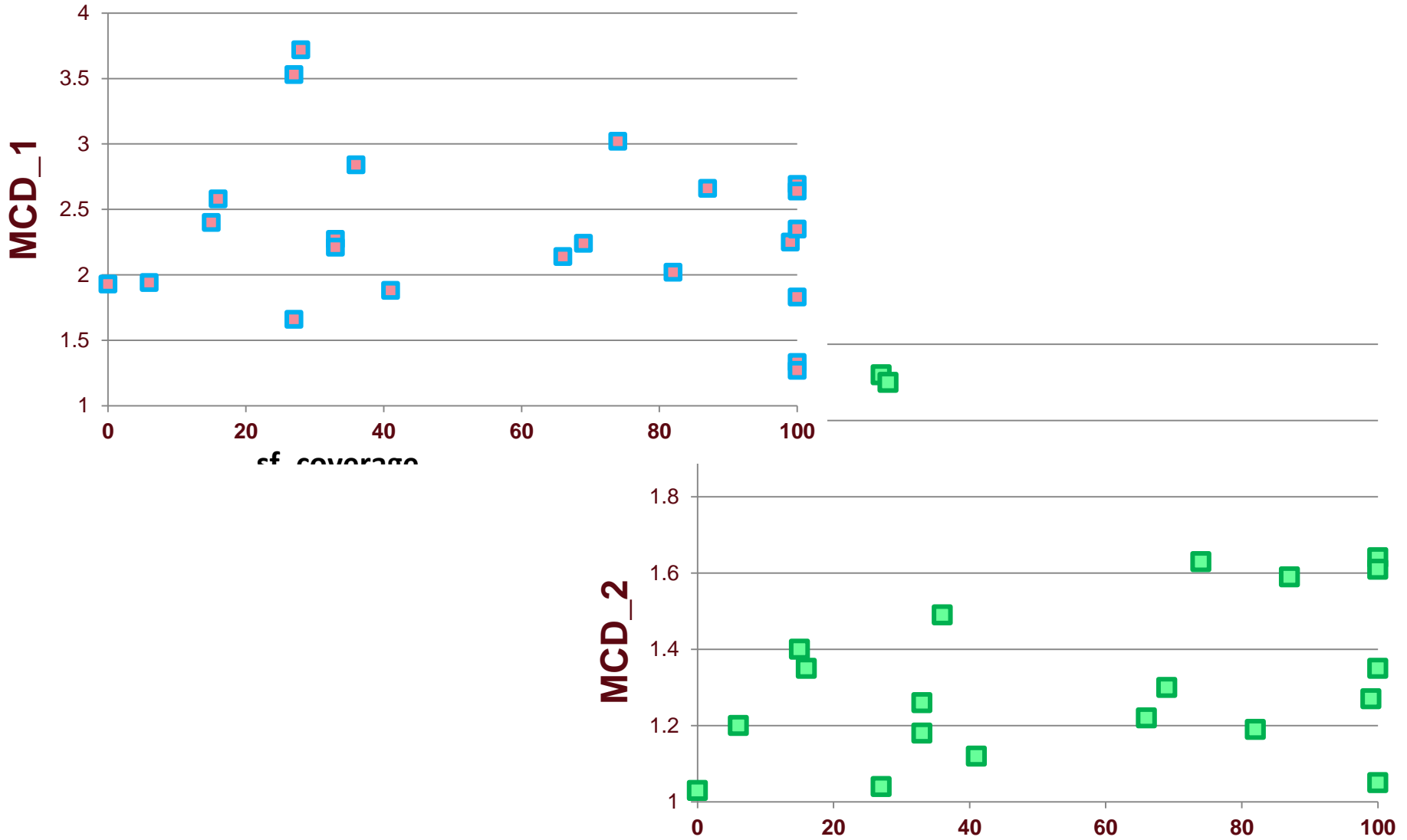
**SF coverage = 100 \* Size of the actual frame list / Population size**



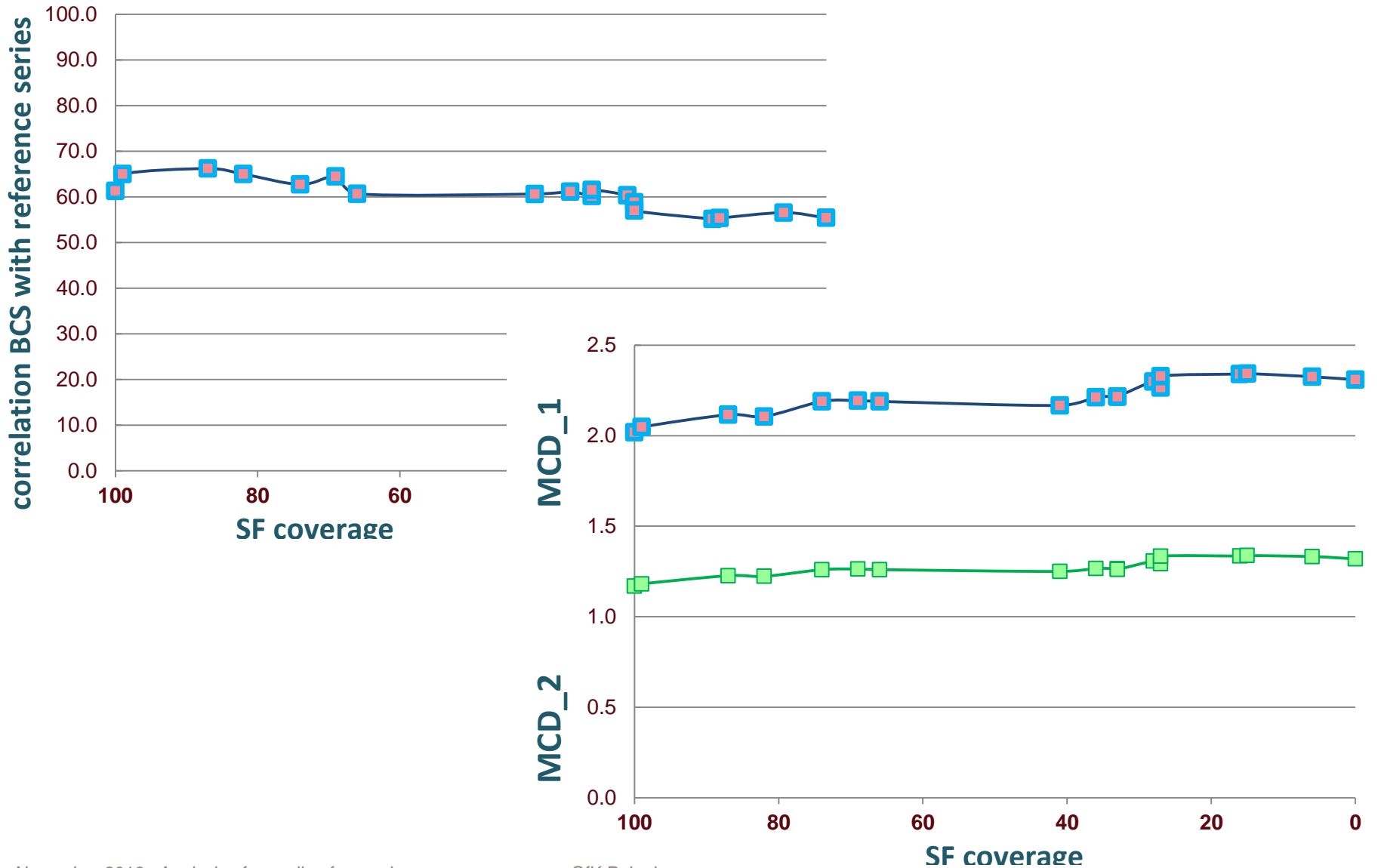
# Sampling frames coverage - analysis of impact



# Sampling frames coverage - analysis of impact



# Sampling frames coverage - analysis of impact



## Additional analysis

mean correlation BCS with reference series for respective level of SF update frequency and category of SF coverage

		correlation BCS with reference series			
		categorized SF coverage			
		0 - 50	51 - 90	91 - 100	Total
SF update frequency	,003	60.0		74.0	71.2
	,080		40.0		40.0
	,250	42.7		46.0	43.5
	1,000	55.7	15.0	59.3	52.7
	2,000	19.0	54.0		36.5
	3,000		84.0		84.0
	10,000		76.0		76.0
	Total	49.2	53.8	65.0	55.4

# Additional analysis

mean MCD\_1 for respective level of SF update frequency and category of SF coverage

		MCD_1			
		categorized SF coverage			
		0 - 50	51 - 90	91 - 100	Total
SF update frequency	,003	1.88		2.02	1.99
	,080		3.02		3.02
	,250	2.68		2.02	2.52
	1,000	2.41	2.14	2.09	2.29
	2,000	2.58	2.02		2.30
	3,000		2.24		2.24
	10,000		2.66		2.66
	Total	2.45	2.42	2.05	2.31



## Additional analysis

mean MCD\_2 for respective level of SF update frequency and category of SF coverage

		MCD_2			
		categorized SF coverage			
		0 - 50	51 - 90	91 - 100	Total
SF update frequency	,003	1.12		1.17	1.16
	,080		1.63		1.63
	,250	1.51		1.12	1.41
	1,000	1.38	1.22	1.21	1.32
	2,000	1.35	1.19		1.27
	3,000		1.30		1.30
	10,000		1.59		1.59
	<b>Total</b>	1.39	1.39	1.18	<b>1.32</b>

# 3. Summary and conclusions

## Main findings

### Better quality measures are likely connected with:

- Individuals are sampling frame units
- If telephone numbers are sampling frame units – fixed together with mobile rather than fixed alone
- The higher sampling frame coverage and more frequent update of sampling frame list
- But the higher sampling frame coverage is definitely more important for the quality of results than frequent update

## Recommendations

- Verify the quality of sampling frame
- Revise an access to other sampling frames:
  - if possible – to census based SF
- if CATI is the only option – consider using fix / mobile mix

# 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. Turner, Anthony G. *Sampling frames and master samples*. United Nations Secretariat. Retrieved 12/11/2012.
4. Data source: *Metadata\_checked\_by\_partners - Consumers.xlsx*
5. Cochran, W. G. (1963). *Sampling Techniques*, 2nd Ed., New York: John Wiley and Sons, Inc.
6. Kish, Leslie. (1965). *Survey Sampling*. New York: John Wiley and Sons, Inc.



THANK YOU FOR YOUR ATTENTION