Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Federal Department of Economic Affairs,
Education and Research EAER

Agroscope

Variance estimation of change for partially overlapping data

Swetlana Renner

1st October 2018

www.agroscope.ch I good food, healthy environment

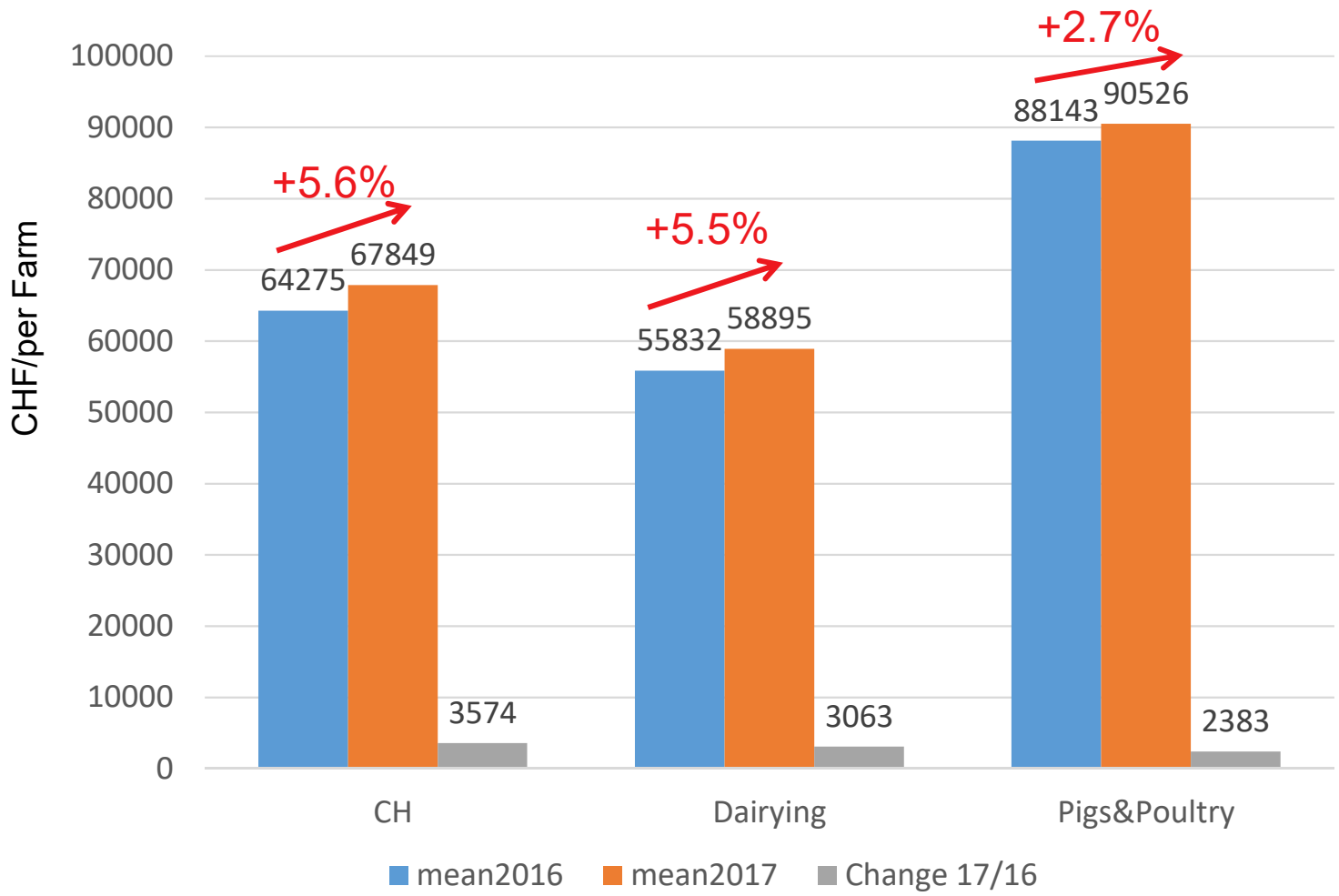
Agroscope

Motivation

- Monitoring changes (or trends) in agricultural performance indicators is of key importance for agricultural policy makers
- All reported indicators based on FADN data are sample estimates (associated with uncertainty)
- Possibility to estimate confidence intervals for change
- Is the estimated difference statistically significant? Real or random?
- We present a possible approach to estimate the variance of change for complex survey, taking into account the overlap between annual data

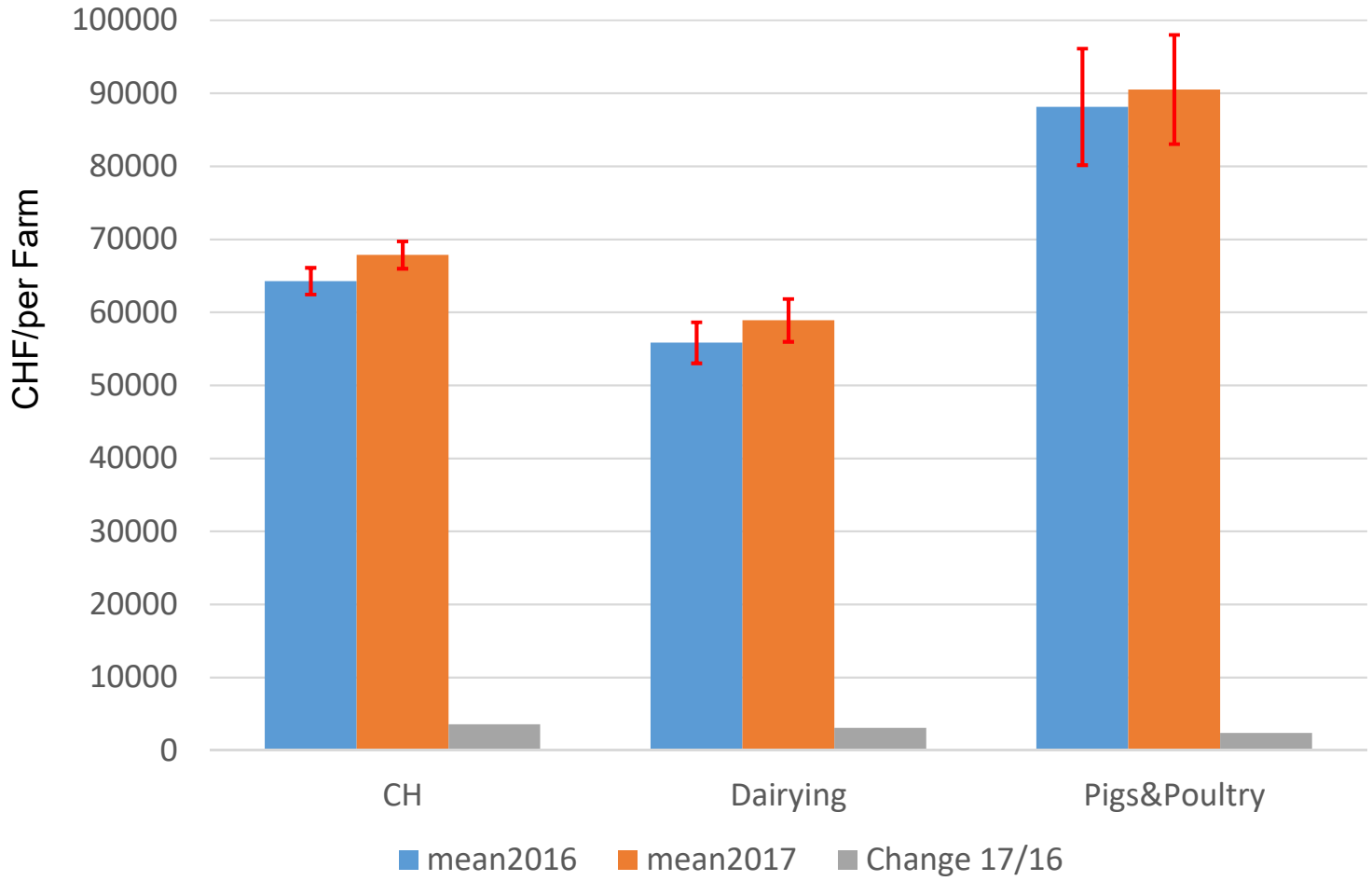
Agroscope

Change in agricultural income

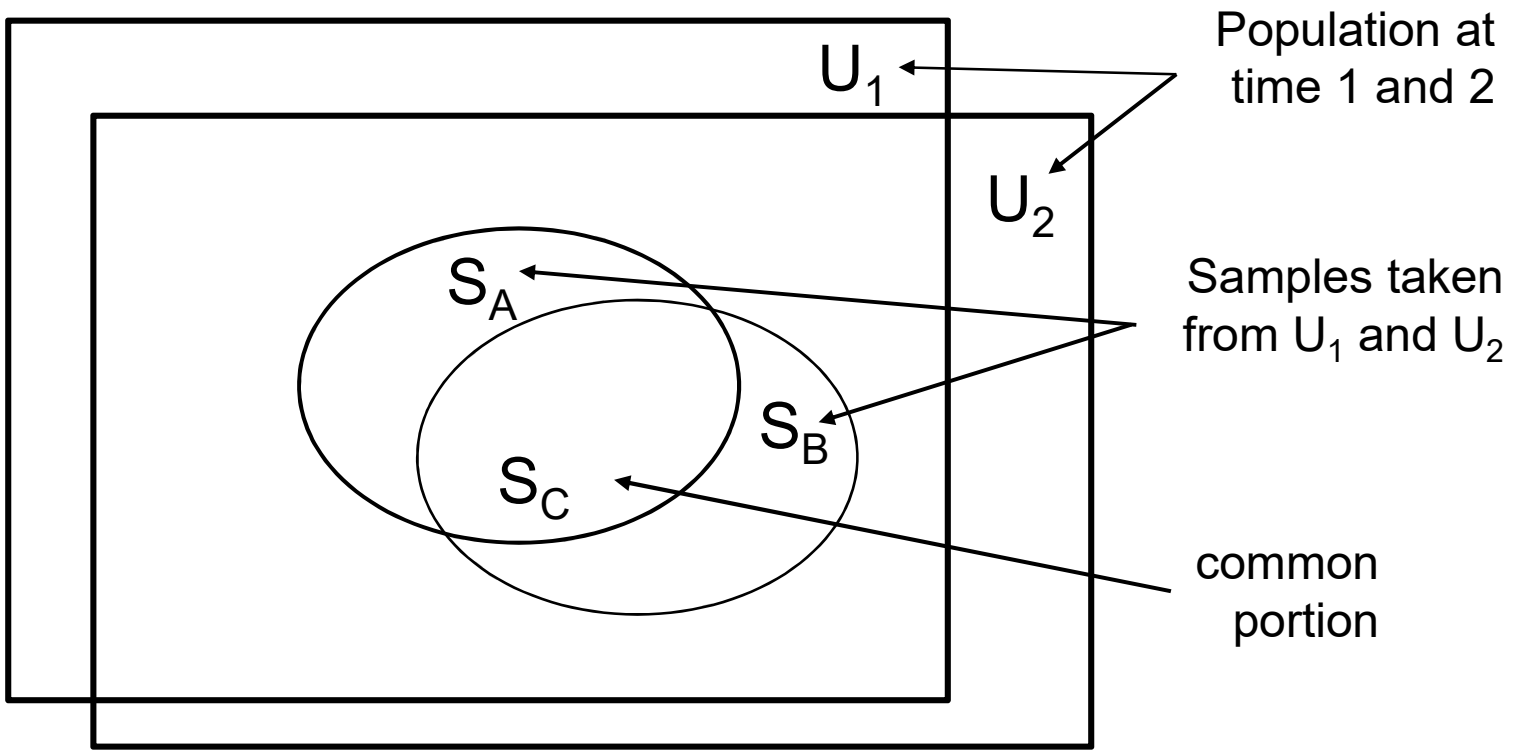


Agroscope

Change in agricultural income



Overlapping samples



Three cases:

1) Independent samples
2) Dependent (or paired) observations
3) Partially overlapping samples

Estimation of change

1) Estimation of change using data of full samples S_A and S_B:

$$\widehat{\Delta\bar{y}} = \hat{\bar{y}}_2^B - \hat{\bar{y}}_1^A$$

2) Estimation of change using common portion S_C:

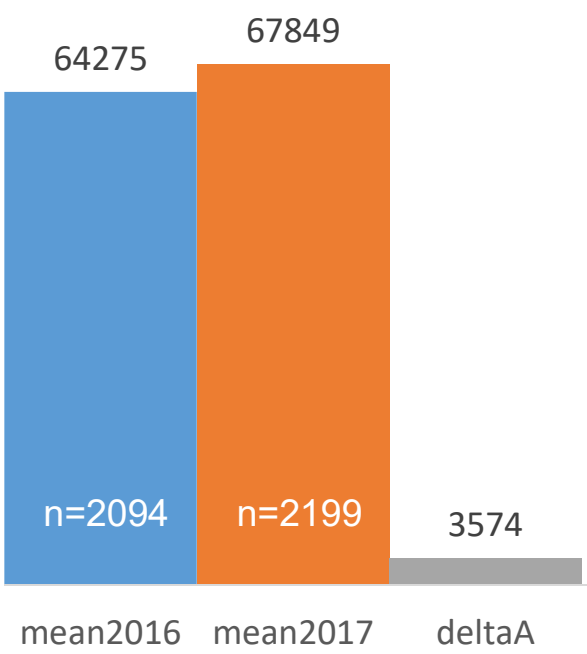
$$\widehat{\Delta\bar{y}}^C = \hat{\bar{y}}_2^C - \hat{\bar{y}}_1^C$$

$\hat{\bar{y}}_1^A$ - estimator of mean for period 1 using full sample S_A
$\hat{\bar{y}}_2^B$ - estimator of mean for period 2 using full sample S_B
$\widehat{\Delta\bar{y}}$ - estimator of change based on full samples S_A and S_B
$\widehat{\Delta\bar{y}}^C$ - estimator of change based on common portion S_C
ρ – Coefficient of correlation for *y* between period 1 and 2
$q = \frac{2n_C}{n_A+n_B}$ – overlap rate, or the fraction of the common sample portion

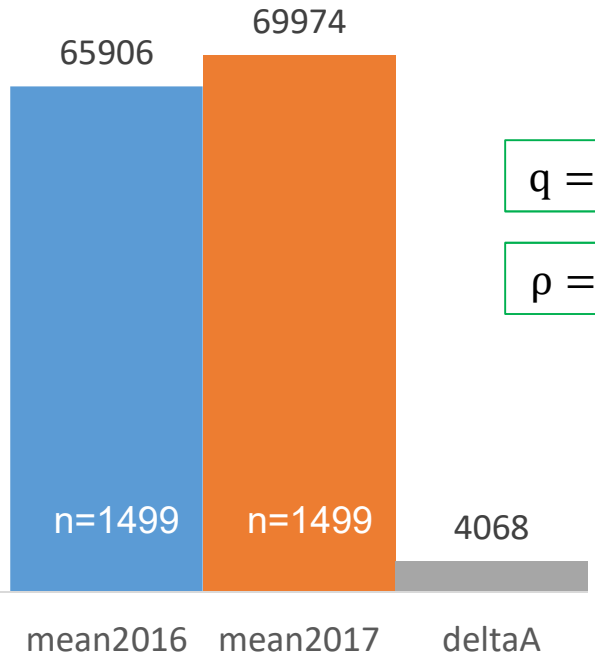
if $\rho \geq \frac{1}{1+q}$, than $\widehat{\Delta\bar{y}}^C$ is more precise (Qualite und Tille, 2008)

Estimation of change

Based on full samples :



Based on common portion:



$q = 0.70$

$\rho = 0.85$

Indep. samples t-test:
(with Welch's df-approximation)

SE: 1884 CHF
t.value: 1.9
p.value: 0.06

Paired samples t-test:
(difference test):

SE: 808 CHF
t.value: 5.0
p.value: 0.00

Agroscope

Competing methods for comparison of two partially overlapping samples

1) Using independent t-test for two full samples
   - assuming no correlation, Standard Error (SE) is overestimated
2) Using paired t-test for paired observations only
   - ignoring independent observations -> loss in df
   - estimated mean and CI are different from reporting values
3) Combining the first two tests (e.g. by weighting) based on independent and paired data (e.g. Samawi and Vogel, 2013)
4) Corrected z-test by Looney and Jones (2003) or extension for unequal variances by Derrick et al (2017)

More general approach: variance estimation of change

More general approach:

t-Test statistic : $$t = \frac{\widehat{\Delta\bar{y}}}{\boldsymbol{SE}(\widehat{\Delta\bar{y}})}$$

Confidence Interval: $CI(\widehat{\Delta\bar{y}}) = \widehat{\Delta\bar{y}} \pm 1.96 \cdot \boldsymbol{SE}(\widehat{\Delta\bar{y}})$

$$SE(\widehat{\Delta\bar{y}}) = \sqrt{\widehat{Var}(\widehat{\Delta\bar{y}})}$$

Approach for variance estimation of change for repeated surveys proposed by Qualite and Tille (2008) and Berger and Priam (2016) is more general and flexible (consistent with sampling design and weighting procedures)

Variance estimation of change

$$\widehat{Var}(\widehat{\Delta y}) = \widehat{Var}(\hat{\bar{y}}_2^B) + \widehat{Var}(\hat{\bar{y}}_1^A) - 2\widehat{Cov}(\hat{\bar{y}}_2\ , \hat{\bar{y}}_1\ )$$

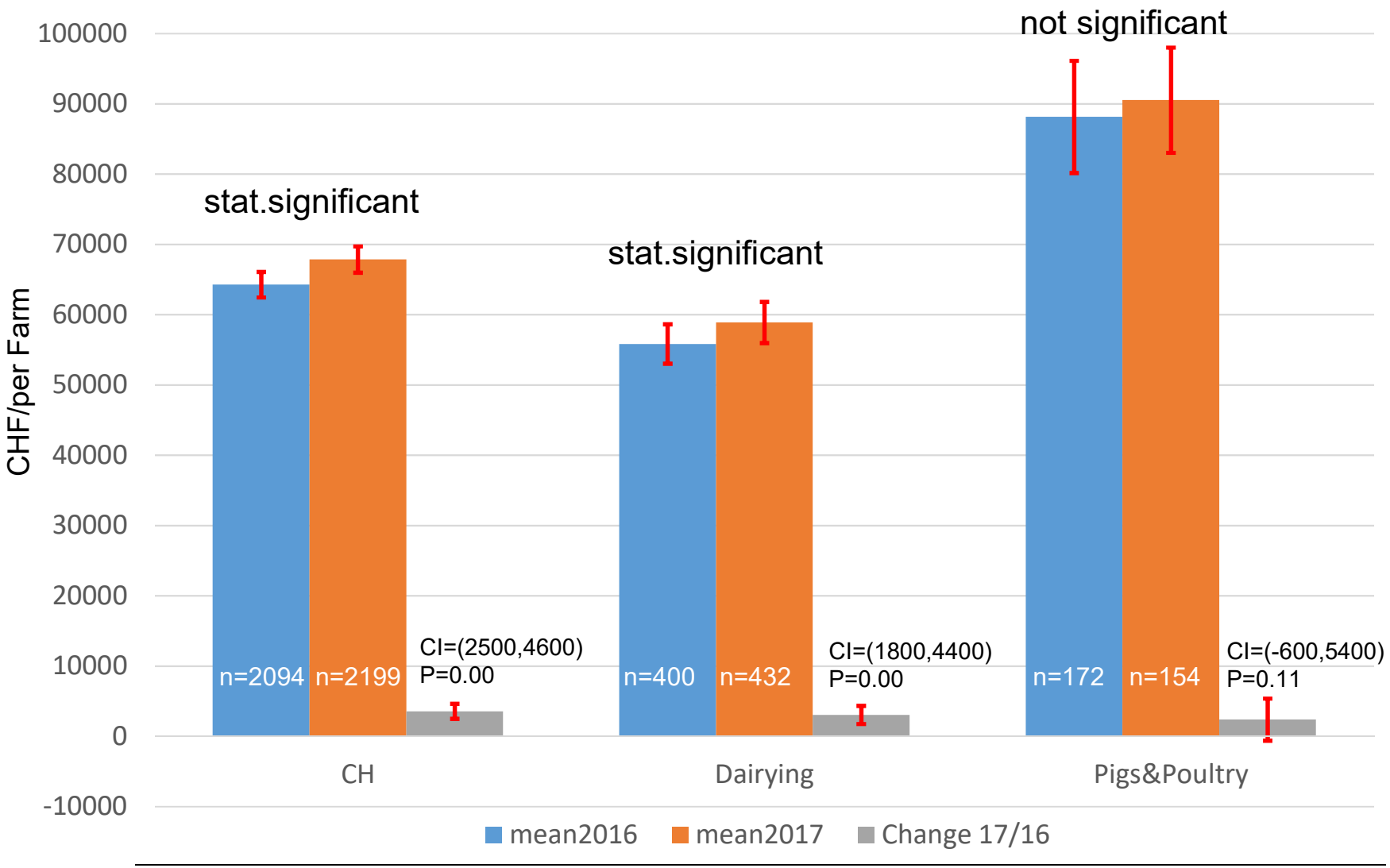
$\widehat{Var}(\hat{\bar{y}}_t\ )$- depending on sampling design and weighting method. We use residual method (Deville and Särndal,1992) for calibrated mean

$$\widehat{Cov}(\hat{\bar{y}}_2\ , \hat{\bar{y}}_1\ ) = \rho_{12} \cdot \sqrt{\widehat{Var}(\hat{\bar{y}}_2^B) \cdot \widehat{Var}(\hat{\bar{y}}_1^A)}$$

ρ_{12} - Coefficient of correlation based on sampling design:

1) Based on common portion S_C (Qualite, 2009)
2) Based on covariance of the residuals of a multivariate regression model (Berger and Priam, 2016)

Change of agricultural income



Agroscope

Conclusions and limitations

- Agroscope reports annual changes in key performance indicators using the difference of the estimated means of full samples. Results are compared with the difference of means based on common portion.
- The variance estimator of change is used in order to judge whether or not the observed difference is statistically significant
- The variance estimator is consistent with sampling design and can be used to calculate the confidence intervals and for testing the differences of overlapping samples
- Limitations: Problem of small samples, violations of normality assumption (e.g. presence of extreme values)
  - ➢ Corrected degrees of freedom
  - ➢ Non-parametric tests, e.g. Wilcoxon test
    (Considering weights? Confidence intervals?)

Literature

Berger, Y.G. and Priam, R., 2016. A simple variance estimator of change for rotating repeated surveys: an application to the European Union Statistics on Income and Living Conditions household surveys. Journal of the Royal Statistical Society: Series A (Statistics in Society), 179(1), pp.251-272.

Derrick, B., Russ, B., Toher, D., & White, P., 2017. Test statistics for the comparison of means for two samples that include both paired and independent observations. Journal of Modern Applied Statistical Methods, 16(1), pp.137-157.

Deville, J.C. and Särndal, C.E., 1992. Calibration estimators in survey sampling. Journal of the American statistical Association, 87(418), pp.376-382.

Qualité, L. and Tillé, Y., 2008. Variance estimation of changes in repeated surveys and its application to the Swiss survey of value added. Survey Methodology, 34(2), pp.173-181.

Qualité, L., 2009. Unequal probability sampling and repeated surveys (Doctoral dissertation, Université de Neuchâtel).

Samawi, H. M. and Vogel, R., 2014. Notes on two sample tests for partially correlated (paired) data, Journal of Applied Statistics, 41:1, 109-117.

Agroscope



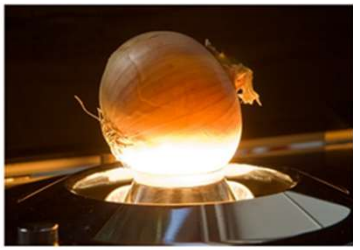
















Thank you for your attention

Swetlana Renner
swetlana.renner@agroscope.admin.ch

Agroscope good food, healthy environment
www.agroscope.admin.ch

Agroscope