Economics
Agroscope Science | No. 68 / 2019

Survey system of the Swiss Farm Accountancy Data Network with two samples:

Income Situation sample and Farm Management sample

Authors:
Swetlana Renner, Pierrick Jan, Daniel Hoop, Dierk Schmid, Dunja Dux, Andreas Weber, Markus Lips

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössisches Departement für Wirtschaft, Bildung und Forschung WBF
Agroscope

Masthead

Published by:	Agroscope Tänikon 1 8356 Ettenhausen https://www.agroscope.ch
Information:	Swetlana Renner swetlana.renner@agroscope.admin.ch
Design:	Jacqueline Gabriel
Cover photo:	123rf.com/Gabriela Brändle, Agroscope
Copyright:	© Agroscope 2018
ISSN:	2296-729X
ISBN:	978-3-906804-79-8

For the sake of easier readability, we have dispensed with gender-neutral formulations where their use would be cumbersome or excessively long.

Table of Contents

List of Tables

List of Figures

Abbreviations

FSS	Farm Structure Survey (FSS), conducted annually within the Agricultural Information System (AGIS Database)
AHV/IV/EO	Social security contributions
FSO	Swiss Federal Statistical Office
AY	Accounting year
FOAG	Federal Office for Agriculture
CRM	Customer relationship management
DB	Database
FTP	File transfer protocol
LU	Livestock unit(s)
FADN	Farm Accountancy Data Network
SME	Small and medium enterprise(s)
EAA	Economic Accounts for Agriculture
UAA	Utilised agricultural area
OR	Swiss Code of Obligations
OTRS	Open ticket requesting system
PS	Post-stratification
RRC	Cumulative response rate (Response rate, cumulative)
RRT	Retention rate
RRU	Unweighted response rate (Response rate, unweighted)
RRW	Weighted response rate (Response rate, weighted)
SGM	Standard gross margin
SO	Standard output
SCL	Systematic collection of laws
TVD	Swiss Stock Movement Database («Tierverkehrsdatenbank”)
SWISS FADN	Swiss Farm Accountancy Data Network
ZA2015	Reform of the FADN (introduction of the new system was originally scheduled for 2015)

Symbols

Symbols are listed in the order of their appearance in the report.

$U = \{1, \dots, N\}$	Set of all farms in the target population ('universe')
N	Number of farms in the target population
$S = \{1, \dots, n^S\}$	Set of all activated farms with valid contact data (gross sample, 'selected')
n^S	Number of activated farms with valid contact data
$R = \{1, \dots, n^R\}$	Set of all farms supplying data with complete, plausible and evaluable datasets (net sample, 'responded')
n^R	Number of farms supplying data with complete, plausible and evaluable datasets
h	Index for stratum according to the selection plan
q	Power allocation coefficient
CV	Coefficient of variation (standard deviation or standard error divided by the mean)
Pr_k	Selection probability, i.e. the likelihood of Farm *k* being selected for the gross sample
ρ_k	Participation probability, i.e. the likelihood that Farm *k* (which has been selected for the gross sample) will take part in the sample by providing complete and plausible data
d_k	Sample-design weight for Farm *k*: reciprocal of the selection probability
w_k	Total weight for farm *k* of the net sample
π_k	Inclusion probability, i.e. the total probability of Farm *k* ending up in the net sample
π_{kl}	Pairwise inclusion probability of Farm *k* and Farm *l*, $k \neq l$
r_k	A dummy variable assuming the value of 1 with the participation of Farm *k* and the value of 0 with its non-participation
g_k	Ratio between the calibration weight w_k and initial weight d_k
y_k	Value of the variable of interest for Farm *k*
θ	Ratio between two estimated variables of interest, y_1 and y_2 (y_1/y_2)
$\boldsymbol{x_k} = (\boldsymbol{x_{k1}}, \dots \boldsymbol{x_{kJ}})$	Vector of *J* auxiliary variables for Farm *k*, which are known for both the farms of the sample and for the total population
t	Multiplier of a Student's t-distribution
e_k	Residuals from the calibration model
B	Parameters to be estimated from the calibration model
Var	Variance of an estimator
se	Standard error
CI	Confidence interval of an estimator
Cov	Covariance
Cor	Correlation coefficient

Summary

Located at Agroscope in Tänikon, the Farm Accountancy Data Network (FADN) analyses the economic situation of Swiss farms and farming families on behalf of the Swiss Legislator. To this end, the FADN has been surveying bookkeeping data as well as characteristics of farm structure and farming households from a sample of selected farms for over 40 years. These data are weighted, analysed and made available to the public.
Owing to a number of shortcomings in the previous survey system of the 'reference farms', a fundamental reform of the Farm Accountancy Data Network took place in the period between 2007 and 2016. The new survey system is based on two samples. For income monitoring, the new Income Situation sample, which is based on a random selection of farms and is meant to representatively illustrate the Swiss agricultural sector, was introduced. Originating from the previous 'reference farms', the newly established Farm Management sample is based on a detailed financial accounting, supplemented with variable direct costing, and limited to common and relatively simple-to-recruit farm types. These two samples fulfil different aims, and are surveyed and evaluated in accordance with the specific requirements for data quality and scope.

The changeover to the new survey system is associated with several methodological innovations that are relevant for both samples and have an impact on income estimation. Along with the reform, a new farm typology was introduced that takes account of the official dairy-cow population recording system. The income calculation method was adjusted to ensure comparability with the reference salary of secondary and tertiary-sector employees. The presentation of the cash-flow statement was adjusted to the Swiss Accounting Standards (Swiss GAAP FER) whilst taking account of the peculiarities of account-keeping for a farm in the Swiss context and the data available from the financial accounting. In addition, the sampling criteria for the exclusion of small farms were changed. The statistical-methodological adjustments concern sampling design, selection process and the evaluation methodology.

The present report gives an overview of the new survey system consisting of two samples (Chapter 1), and describes the survey design and the evaluation process of the two samples (Chapters 2 and 3). The figures presented in the report serve to illustrate the methodology, and refer to the data collected for the 2016 accounting year.

Key Data: Income Situation Sample

The figures shown concern the data collected for the **2016 accounting year**.

Basic population	Farms included in the Farm Structure Survey (FSS), conducted annually within the Agricultural Information System
Target or sampled population	Individual enterprises and group farming businesses (natural persons and unregistered partnerships) that lie above the minimum standard output[1] threshold of a region, and thus cover 95 % of the total standard output of the plain, hill and mountain regions
Survey area	The whole of Switzerland
Selection process	Stratified random sample
Form of recruitment	By telephone, with prior written notification by letter
Data collection method	Online survey tool
Recruitment period	November 2016 to April 2017
Data delivery period	January 2017 to August 2017
Requirements for participation	The keeping of financial accounts, data delivery up to August 2017, agreement to the linking of data with information from the FSS Database
Average length of stay on panel (since 2014 accounting year)	1.9 years
Maximum length of stay on panel	Not stipulated
Target sample size	2300
Number of activated[2] farms in 2016 accounting year (gross sample size)	6289
Number of supplied, plausible and evaluable datasets in the 2016 accounting year (net sample size)	2094
Average response rate for new recruits (2016 accounting year)	14 %
Retention rate for farms on the panel (2016 accounting year)	64 % (wave 1), 86 % (wave 2), 81 % (wave 3)
Cumulative response rate (2014 accounting year – 2016 accounting year)	13 %
Contracting authority	Federal Office for Agriculture
Main publications	Media Release, Main Report, Basic Report (www.grundlagenbericht.ch)

[1] Standard output is calculated by the Swiss Federal Statistical Office for each farm, and depicts the average monetary value of agricultural production at producer prices (without direct payments).

[2] Farms randomly selected from the target population that received the announcement letters.

Key Data: Farm Management Sample

The figures shown concern the data collected for the **2016 accounting year**.

Target population or sampled population	Individual enterprises and group farming businesses (natural persons and unregistered partnerships) with a utilised agricultural area of at least 10ha, or with a livestock numbers of at least 8 livestock units from the strata defined in the selection plan
Survey area	The whole of Switzerland in the selected strata
Data collection method	Desktop-based collection tool with an interface for the transfer of data from accounting software
Maximum length of stay on panel	Not stipulated
Data delivery period	March 2017 to August 2017
Requirements for participation	The keeping of financial accounts with variable direct costing, data delivery up to August 2017
Target sample size	2100
Number of supplied, plausible and evaluable datasets in the 2016 accounting year	1535
Contracting authority	Federal Office for Agriculture
Main publications	Production Branches Report (www.grundlagenbericht.ch)

1 Background and Overview of the ZA2015 Reform

1.1 Initial situation

The Swiss Farm Accountancy Data Network (SWISS FADN) forms part of the **agricultural monitoring system** under the authority of the Swiss Legislator[3] and has a dual remit. Firstly, it is tasked with determining key variables such as agricultural income or working income, i.e. the compensation of a full-time family labour unit on the farm; and secondly, it provides a host of business performance indicators for the attention of practitioners, advisory services, researchers and administrators. Since 1976, these tasks have been performed by **Agroscope**, the Swiss Confederation's departmental research institute for agriculture, at its Tänikon site. The collection and evaluation systems are regularly overhauled due to e.g. changing technical possibilities or new legal requirements. In 2003 a revision was concluded, with the change from the test to the reference farms taking place (Meier, 2000; Hausheer Schnider et al., 2004). As before, the 2003 reform also stuck to the principle of fulfilling the two above-mentioned tasks with a single sampling.

In January 2007 the Methodology Service of the Swiss Federal Statistical Office (FSO) criticised the previous quota sample and called for a random selection of farms, in keeping with the Charter of Swiss Official Statistics (FSO, 2002). A second challenge – one that had existed for quite some time – was posed by the insufficient representation of some parts of the target population, such as e.g. the lack of farms from the canton of Ticino, or the meagre representation of farms focusing on special crops, i.e. fruit, grapes or vegetables. Thirdly, more farms were delivered and compensated in certain individual strata than envisaged in the selection plan. This excessive delivery constituted a less-than-efficient use of resources. The ZA2015 Reform was initiated in order to tackle these challenges. Moreover, during the reform the new accounting law came into effect (Articles 957 to 963b of the Swiss Code of Obligations[4]), which had to be taken into account for small and medium enterprises (SMEs) from the 2015 accounting year. This law is also binding on farms[5], which led to the development of the 'Agricultural SME Chart of Accounts' (AGRO-TWIN AG, 2014). Because of existing inconsistencies in the comparison of working income per family labour unit with the reference salary of employees in the secondary and tertiary economic sectors, income calculation in the SWISS FADN was overhauled. Furthermore, over the course of the reform process it emerged that income estimation was influenced by the fluctuation caused by farms entering and exiting the sample. Consequently, this sample effect was reported in the Main Report and Basic Report for the most important variables from the 2009 accounting year onwards (Schmid and Roesch, 2010).

Before the ZA2015 Reform, agricultural-monitoring evaluations were based on farm accounting data supplied by the **reference farms**. So-called 'farm accounting' is a detailed bookkeeping method with double-entry accounting and variable direct costing, carried out according to the requirements of the FADN, which was a precondition for participation in the Reference Farm sample. The owners of the software 'Agro-Twin' – at the time, the only accounting software used for farms supplying data to the FADN – began designing the successor product in 2007.

[3] Pursuant to the Ordinance of 7 December 1998 relating to the Assessment of Sustainability in Agriculture (SR 919.118), the economic situation is measured *inter alia* by means of reference farms (Articles 2 and 4 to 6). In addition, the FADN is contained in the Ordinance of 30 June 1993 relating to the Conducting of Statistical Surveys by the Swiss Conferation (SR 431.012.1).

[4] SR 220, Federal Law supplementing the Swiss Civil Code (Part Five: Code of Obligations) of 30 March 1911 (status as at 1 April 2017).

[5] For the smallest enterprises with a turnover of less than CHF 500 000.-, regardless of their legal form, a simplified statement of income and expenditure plus a statement of financial position is sufficient. SMEs benefit from a limited audit requirement and are not obliged to submit a cash-flow statement. Moreover, sole proprietorships and partnerships are released from the obligation to supply the annexe.

In connection with this, it was essential to clarify whether there was still a need for a second balance sheet, i.e. the simultaneous preparation of financial accounts for the tax authorities and farm accounts for the FADN. The issue was debated by a working group in 2008 and 2009, with the result that in future, the FADN will determine income on the basis of financial accounting (Lips et al., 2009).

1.2 Two-sample concept

With the ZA2015 reform, a **two-sample concept** was therefore designed (Lips et al., 2009 and 2011) in order to optimally meet the requirements for agricultural monitoring. In the I**ncome Situation** sample, the focus is on those variables that are absolutely necessary for calculating agricultural income, working income and cash flow. Compared to the previous system, this means a significantly lower survey effort for farm managers and accounting offices. This creates an important precondition for recruiting farms from strata which to date have scarcely been represented, or for acquiring farms for the evaluation from previously unrepresented or only sparsely represented regions. By contrast, the **Farm Management** sample requires an accounting method with variable direct costing, and focuses on widespread and relatively easily recruited farm types. Originally, the plan was to draw both samples by random selection. Provided that neither sample is biased, it would therefore have been possible to use the farms from both samples to estimate agricultural income (Roesch and Lips, 2013). Over the course of the reform, it became obvious that although random recruitment is feasible for the Income Situation sample, it is nevertheless highly challenging. Because of the major effort involved, we dispensed with random selection for the Farm Management sample, and instead transferred the reference farms already in the survey into the new Farm Management sample. More-detailed information on both samples can be found in Chapters 2 (Income Situation Sample) and 3 (Farm Management Sample) of this report.

1.3 Methodological innovations

The introduction of the new survey system has been accompanied by several methodological innovations which are relevant for both samples and have an effect on the income estimation:

- Whilst farms participating in the previous system were obliged to keep detailed double-entry accounts with variable direct costing according to the requirements of the FADN (‘farm accounts’), in the new system **financial accounting** according to the new accounting law provides the informational basis for both samples. This change has resulted in a significant simplification of data acquisition for farm managers and accounting offices. However, differences in accounting practice (e.g. nonlinear depreciation, the distinction between business and private) mean that the earlier income estimate is not completely comparable with the new one (break in the time series).
- **Income calculation** in the FADN was overhauled because of existing inconsistencies in the comparison of working income per family labour unit with the reference salary of employees in the secondary and tertiary economic sectors. This relates to the operational offsetting of the employer’s contributions for the first and second pillar of the pension scheme of the farm manager and his or her partner, and the calculation of the costs of their residence.
- The representation of the **cash-flow statement** was adjusted to the Swiss Accounting Standards (Swiss GAAP FER) whilst taking account of the peculiarities of account-keeping for a farm in the Swiss context and the data available from the financial accounting. An important change concerns the definition of the fund. The cash-flow statement of the two new samples is based on the ‘liquid assets’ fund, instead of on the so-called ‘net current monetary assets' used previously.
- A further methodological change concerns the slight adaptation of the **farm typology**. Whereas dairy farmers were previously differentiated indirectly based on milk utilisation (differentiation of the cows according to commercial and non-commercial milk production, as well as suckler cows), the new farm typology distinguishes dairy farmers according to the percentage of dairy cows out of the entire cattle population.

- The **sampling criteria** that define the field of observation (i.e. determine which farms can take part in the survey, and which Swiss farms are represented by the sample) for the two samples have also changed.

  Unlike in the previous system of reference farms, in the two new samples the group farming businesses are taken into account for the agricultural income estimate, in addition to the individual farms. The minimum thresholds for the exclusion of small farms were also adjusted. For the Farm Management sample, a simplified delimitation with just two instead of eleven physical performance indicators was introduced. The exclusion of small farms in the Income Situation sample is made on the basis of minimum economic size, with the result that 5 % of the standardised value of gross production (also referred to as standard output) of the region in question is delimited (or put differently, 95% of the standard output is represented by the sample).
- The **statistical adaptations** concern sampling design, selection process and evaluation methodology. Farms in the Income Situation sample are selected by means of a stratified random sample according to the selection plan, which is optimised in terms of accuracy of income estimate. The weighting methodology for extrapolating the results to the total population was also adjusted (calibration instead of post-stratification). Farms in the Farm Management sample will continue to be selected by their accounting offices and passed on to the FADN, although the financial incentive system has been slightly adjusted and the selection plan limited to the most important Swiss farm types.

The methodological adjustments carried out over the course of the changeover to the new system (income calculation based on financial accounting, new presentation of cash-flow statement and adapted farm typology) are explained in detail (in German) in Agroscope Science Report No. 68 (Renner et al., 2018). The new sampling criteria and statistical adaptations of the two samples ‘Income Situation’ and ‘Farm Management’ are described in Chapters 2 and 3 of this report.

Chronology of the reform process

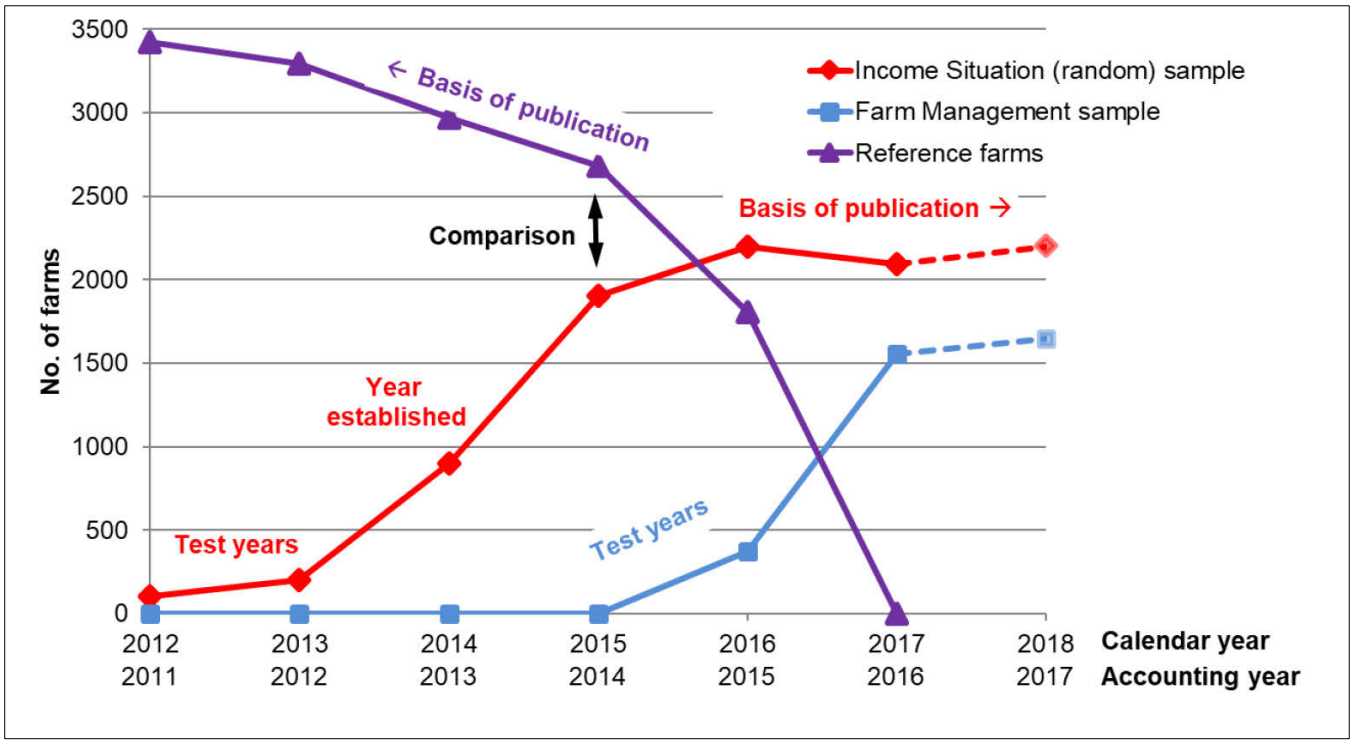
The chronology of the reform process is given in Figure 1, with the number of farms of the various samples being shown on the vertical axis.

In the 2012 calendar year (2011 accounting year), the Reference Farm sample encompassed 3425 farms. During the reform process, this number was continually whittled down. The new random-selection-based Income Situation sample was extensively tested in the three calendar years 2011, 2012 and 2013. The Income Situation sample was established in calendar years 2014 (919 farms in total) and 2015 (1907 farms in total).

For calendar year 2015 (accounting year 2014), a comparison of the old and new systems was possible. This revealed a bias in the Reference Farm sample, described in detail in Chapter 2.8.

The changeover from the previous Reference Farm sample to the Income Situation sample took place in 2016, i.e. with the data from the 2015 accounting year. On the occasion of the Agricultural Economics Conference held in Tänikon on 15 September 2016, and based on the 2014 and 2015 accounting years, for the first time it was possible to present the change in income on the basis of the Income Situation random sample – a milestone in the history of the FADN. Since then, the Income Situation sample has constituted the basis for publishing agricultural-income figures.

The importing of some of the reference farms into the Farm Management sample occurred from the 2016 calendar year onwards. Farm production-branch results from this sample were first published in the 2017 calendar year.



Basis of publication:
Reference farms

Basis of publication:
Income Situation Sample

Figure 1: Chronology of the reform process.

2 Income Situtation Sample

2.1 Objectives

Pursuant to the mandate of the Swiss legislator, Agroscope carries out income monitoring for the Swiss agricultural sector. The Income Situation sample forms the basis for the estimation of agricultural income and working income for the agricultural sector at the level of Switzerland as a whole, as well as at regional (plain, hill and mountain) level. Accordingly, it covers the totality of farm types, regions and area categories (so-called 'strata') in Switzerland. The farms are randomly selected from the target population, so that statistically reliable results can be published. The participating farms and their accountants are recruited by a recruiting agency, with the latter being informed only of the relevant address and telephone number, but having no access to the accounting data submitted. The accounting data are submitted directly to Agroscope via an online questionnaire and then pseudonymised, i.e. stored and analysed without names and addresses. This procedure ensures the anonymity of the participating farms.

In order to achieve as high a level of willingness to participate as possible, and to keep those farms that are willing to participate on the panel as long as possible[6], the collection effort for those supplying data must be kept in check. For this reason, an online questionnaire was developed which enabled data submission independently of the accounting software or a specific accounting framework. Whole-farm key figures from the financial accounting supplemented with details from the tax declaration as well as the income situation of the household are collected (e.g. non-agricultural income and private consumption). The non-monetary data are available from the farm-structure survey and do not need to be collected, as they are linked with the collected data by means of a pseudonymised ID.
The subchapters below give a detailed explanation of the statistical and methodological approach to income monitoring on the basis of the 2015 and 2016 accounting years.

2.2 Definition of the sampled population

At the outset, we must define the basic population of farms represented by the Income Situation sample, and the farms from which the required random samples are drawn. The **basic population** of the farms consists of all farms[7] in Switzerland that are captured in the annual Farm Structure Surveys of the Agricultural Policy Information System (FSS database). The **target population** only includes commercial sole proprietorships and group farming businesses from a particular size onwards. These farms are to be represented by the Income Situation sample. This distinction also means that only certain farm types with specific legal forms, and which exceed a minimum economic size, have a chance of being selected for the random sample. Moreover, the so-called **sampled population** contains only those farms included in the FSS database at the time of the drawing of the random sample, or at the time of data analysis. The complete list of these farms is made available to the Swiss FADN by the Federal Office for Agriculture (FOAG), and constitutes the **sampling frame** for the drawing of the random sample.

[6] A maximum retention period of eight years was originally planned in order to ensure the accurate portrayal of structural change using the new recruits. Because of the high panel attrition, however, there is no need to specify the maximum retention period, since every year a high percentage of new farms must be recruited in order to make up for the loss of farms from the panel.

[7] The FSO definition of agricultural holding (*landwirtschaftlicher Betrieb*) applies (FSO, 2016). A business is considered to be an agricultural holding if it engages in plant production and/or livestock husbandry year-round; consists of one or more production facilities; is legally, economically, organisationally and financially autonomous, and independent of other farms; reports an operating result of its own, and is farmed throughout the year. In addition, the farm must fulfil at least one of the following conditions: possess at least 1 hectare utilised agricultural area **or** 0.3 hectares special crops **or** 0.1 hectare in a protective environment; **or** 8 sows **or** 80 fattening pigs **or** 80 fattening-pig places; **or** 300 head of poultry.

In addition to structural data (area under cultivation, livestock units, workforce, etc.), the FSS database also contains further information on the farm (operating and legal form, farming system, production zone, etc.), as well as additional attributes calculated by the Swiss Federal Statistical Office (FSO) (farm type, standard output, standard gross margin). These key figures are used to define the sampled population, create strata (Chapter 2.3.1), develop the selection plan (Chapter 2.3.2), and calculate the weights (Chapter 2.6.2). In addition, and with the consent of the participating farmers, these data are linked with the collected accounting data, in order to reduce the data-collection effort for those supplying data.

The **definition of the sampled population** is effected firstly on the basis of **operating and legal form.** Non-commercial farms, foundations, associations and public bodies are all excluded, since these differ greatly from the actual farms in terms of structure, field of activity and objectives. Individual family farms constitute the majority of agricultural holdings in Switzerland. The existing system does not allow the comparison of family farms with legal entities; for this reason, neither public limited companies nor private limited companies (GmbH) are currently the focus of our investigation.

The second distinction is made according to the **economic size of the farm**. For the analysis of the income situation, only those farms exceeding specific minimum thresholds are taken into account, with the result that the target population contains only those farms that make a substantial contribution to total agricultural production.[8] In addition, the definition according to economic size is also important for practical reasons, viz., in order to ensure the quality of the sample, only those farms which keep financial accounts can take part in the survey. Many small agricultural holdings do not keep financial accounts, however, and therefore are excluded from participation in the survey. According to an additional survey conducted in 2010 by the FSO (2012), around one-third of all farms in Switzerland carry out manual bookkeeping only. This is actually the case for half of all farms with a utilised agricultural area of under 10 ha. By differentiating small farms, a large percentage of farms without financial bookkeeping are excluded before recruitment. This reduces recruitment costs whilst increasing the probability of participation of the contacted farms.

Standard output (SO) is used for the definition of farms according to economic size. This key figure is calculated for each agricultural holding by the FSO, and illustrates the average monetary value of agricultural production at producer prices (without direct payments).[9] The threshold value is determined such that 95 per cent of the total standard output of the plain, hill and mountain regions is covered (Roesch, 2013).[10] The thresholds for the three regions are set separately instead of using a single threshold, so that mountain farms are not excluded disproportionately from the sampled population owing to their significantly below-average SO.
For a closer analysis of the influence of the height of the SO threshold on several important structural variables, we recommend the article by Roesch (2015). The standard-output thresholds calculated for the three regions (plain, hill and mountain) on the basis of the FSS data from 2015 and 2016 can be found in Table 1.

[8] A similar definition of the field of observation is used by the European Commission for the comparable data survey conducted by the Farm Accountancy Data Network. This system includes only those farms whose economic size exceeds the threshold established for this country (EU, 2015; EU, 2014).

[9] Standard output is used in the European FADN system both in the differentiation of the sampled population and in the definition of the farm types (EU, 2011).

[10] Here, the ranked standard outputs of all farms are added up (beginning with the farm with the highest standard output) until 95 % of the total standard output of a region is reached.

Table 1: Thresholds for minimum economic farm size (standard output)

Region	SO Thresholds	
	2015	2016
Plain	CHF 84 401	CHF 84 771
Hill	CHF 53 856	CHF 53 852
Mountain	CHF 36 343	CHF 37 190

The **results of the definition of the sampled population** according to both criteria for 2015 and 2016 are shown in Table 2. Out of a total of around 51 000 to 52 000 agricultural holdings covered in the FSS database, approx. 2100 (4 %) are excluded from the outset owing to their unsuitable operating and legal form. Through the differentiation of the small farms whose SO lies below the threshold values, a further 13 500 farms are eliminated. These farms account for only around 5 % of the total SO, although in terms of numbers they account for approximately one-quarter of all farms in the basic population. All in all, the sampled population encompasses 70 % of the farms in the basic population, and 90 % of total standard output in 2015 and 2016.

Table 2: Result of the definition of the sampled population

Definition Criteria	No. of Farms		Percentage of Farms (%)		Percentage of SO (%)	
	2015	2016	2015	2016	2015	2016
Basic population	51 979	51 185	100	100	100	100
Definition through farm type and legal form	49 884	49 135	96.0	96.0	94.4	94.4
Additional definition through standard output, farm type and legal form **(sampled population)**	36 414	35 713	70.1	69.8	89.7	89.7

For the drawing of the random sample for AY 2016, which was carried out in September 2016, the FADN had access to the latest available structural survey data from 2015 (i.e. the sample frame was one year out of date). According to the definition, the sample population contained a total of 36 414 farms from all three regions of Switzerland. The random sample was drawn on the basis of the previous year's list of farms. At the time of the evaluation of the accounting data from AY 2016 in August 2017, the structural data of the surveyed year (2016) were available. This more-up-to-date dataset consisting of 35 713 farms was used to extrapolate the results. The fact that the sampling is based on a data source that is one year out of date means that some of the farms of the sampled population of the surveyed year (here, 2016) cannot be covered by the sample. This applies e.g. to those farms newly created through re-founding, division or reorganisation, or which, owing to the new SO definition, are now part of the field of observation, even though they were still too small at the time the selection plan was drawn up. On the other hand, farms which actually no longer formed part of the field of observation for the surveyed year can wind up in the sample.[11] Such so-called 'coverage errors' are corrected during the weighting/ extrapolation process (for more on this subject, see Chapter 2.6).

Table 3 shows the percentages of the basic population of all agricultural holdings represented by the sampled population for some selected areas or selected livestock numbers.

[11] If the farms no longer belonging to the current sampled population nevertheless take part in the sample, they will still be used for the evaluations.

Table 3: Percentage of the basic population of all agricultural holdings represented by the sampled population

FSS Key Figure	Whole of Switzerland	Plain Region	Hill Region	Mountain Region
Utilised agricultural area	86.8 %	84.3 %	87.7 %	90.1 %
Open arable land	87.0 %	85.8 %	93.0 %	96.2 %
Grassland	86.9 %	83.1 %	86.9 %	90.1 %
Permanent crops	82.4 %	83.9 %	78.8 %	74.3 %
Total livestock numbers	89.9 %	88.5 %	90.8 %	91.5 %
Total cattle numbers	92.9 %	91.9 %	92.7 %	94.8 %
Total pig numbers	87.6 %	85.5 %	91.9 %	90.2 %
Total poultry numbers	95.1 %	94.2 %	97.2 %	95.2 %

2.3 Sampling plan and selection process

The selection plan determines the criteria according to which the agricultural holdings are selected for participation in the survey, in order to illustrate the sampled population as representatively as possible. The survey system is based on a single-stage **stratified random sample**, and encompasses all farm types and regions. After being defined (see Chapter 2.2), the sampled population is divided into smaller subsets (strata) using specific stratification characteristics, so that random samples can then be drawn from each stratum. Stratification increases the accuracy of the estimate. The size of the sample is determined so as to allow certain requirements concerning the accuracy of the estimation results to be met (i.e. lowest possible standard errors for the main variables investigated), with the number of farms being limited by the budget for recruitment and compensation.
In the first step, a selection plan is calculated for the envisaged net sample. This **net selection plan** is prepared in order to determine the optimal (theoretical) number and distribution of the farms supplying data. Taking into account (stratum-specific and linguistic-region-dependent) response rates, and based on the net selection plan, the number of farms to be activated (gross selection plan) is determined. This **gross selection plan** serves as a template for the drawing of the random sample.

2.3.1 Attributes and criteria for the stratification

When selecting the stratification variables, the following aspects were taken into account:

- The representativeness of the sample with regard to the three regions (plain, hill and mountain) and the eleven farm types (as per ZA2015-typology S3, see Appendix 1) is to be ensured;
- Stratification attributes are to be known for all farms of the sampled population, and correlated with the most important study variables (i.a. agricultural income);
- Strata of maximum heterogeneity are to be created, with a smaller variance within and a larger variance between the strata, in order to further increase accuracy at the level of Switzerland as a whole;
- Specific requirements concerning accuracy both at the level of Switzerland as a whole and for certain lower aggregation levels (e.g. regions, farm types) are to be met.

Bearing in mind these aspects, stratification is performed according to the following three attributes:

- Region (plain, hill, mountain)
- Farm type (eleven types as per ZA2015 S3 typology, see Appendix 1)

- Farm size, defined by the standard gross margin (SGM). Division of the stratification level 'region x farm type' into one to three size categories with flexible boundaries.

In the first step of stratification, all farms are split into groups based on the combination of region and farm type (region x farm type). Depending on how many farms in the sampled population belong to a corresponding category, they are divided in a further step into two or three size categories, in order to increase the accuracy of the estimate. If too few farms are available, this allocation does not take place.[12] Since AY 2015, SGM has been used for farm size, since this variable takes into account both the number of animals and the available land, as well as having a high correlation with agricultural income.[13]

The **boundaries for allocation to size categories** are optimally chosen for each combination of region and farm type, in order to increase the precision of the estimate. The optimum SGM boundaries are calculated by means of the 'cumulative root frequency' methodology of Dalenius and Hodges (1959). Based on the predetermined sample size (2300), the optimal SGM boundaries are calculated such that the variance is minimised on all region and farm-type levels. The optimal size-category boundaries are listed according to region and type in Table 4.

Table 4: Size-category boundaries as measured by standard gross margin (CHF) according to farm type *

Type / Region	1511	1512	1521	1522	1523	1531	1541	1551	1552	1553	1554
Plain	102 868 177 409	170 016 390 355	105 949 173 047	72 295	112 784	113 780	140 825 288 124	141 670 242 577	89 509 151 020	169 053 323 146	115 908 208 787
Hill	103 018	77 891 157 376	75 798 123 978	41 584 68 028	70 119 130 210	80 746	115 562 229 457	144 976	73 938	111 733 210 936	88 046 162 050
Mountain	-	95 314	59 479 97 311	30 853 49 296	50 002 89 901	58 988	173 391	-	-	82 204 153 204	47 130 82 082

* Farm type according to S3 farm typology (see Appendix 1). Two boundaries mean that three size categories were created; one boundary stands for two size categories; and where no boundary is indicated, there was no allocation according to size.

2.3.2 Sample size and calculation of the selection plan

The **optimal sample size**, i.e. the total number of farms in the sample, was set once and for all before the start of the survey by Roesch and Lips (2013) at 2300 farms.[14] This figure takes account of the accuracy requirements and the core parameter values (mean and standard deviation) of the target variable (standard gross margin) as well as the likely costs and the available budget. Stratification was performed according to region, farm type, and for two size categories (under 20 ha and over 20 ha) in each case.[15]
With the original stratification as well as an accuracy requirement of ± 1.95 per cent (= half the length of the confidence interval relative to the mean) for the standard gross margin, a sample size of 2317 farms was calculated using the 2014 structural data.

Since then, this target size of 2300 farms has been retained in order to ensure continuity over the years in the preparation of the selection plan, and owing to longer-term contractual agreements with the recruiting agency.

[12] The rules for combining size categories are explained in Chapter 2.3.3.

[13] In the testing and development years of AY 2010 to AY 2014, there was a fixed allocation of the farms to two size categories – under 20 ha, and over 20 ha – for all combinations of farm type and region.

[14] The sample size of 2300 farms was defined according to formula 3.14 by Roesch and Lips (2013).

[15] The original approach of two fixed size categories based on the UAA was replaced by three variable size categories based on the SGM.

The **optimal allocation** of farms to the individual strata is determined on the basis of the stipulated sample size of 2300 farms. A critical factor here is the variance of the standard gross margin, with this key figure being used as a proxy for the most important study variable – agricultural income – owing to the high correlation of these two figures.
The following criteria are borne in mind when determining the number of farms in the strata: (i) the relative dispersion around the stratum mean, (ii) the estimated accuracy in the strata, and (iii) the practical feasibility of the random sample.[16]
The optimal allocation of the sample to the individual strata n_h is performed according to the power allocation formula of Bankier (1988):

$$n_h = n \cdot \frac{N_h^q \cdot CV_h}{\sum_{h=1}^{L} N_h^q \cdot CV_h} \quad , \tag{1}$$

where *n* represents the total target (net) sample size, *L* is the number of strata, N_h is the number of farms in the sampled population of the stratum *h*, and CV_h is the coefficient of variation[17] (standard deviation divided by the mean) of the standard gross margin in the stratum *h*.
The coefficient *q* ('power allocation coefficient') was set at the level *q=0.5*. This was done with the aim of achieving a compromise between the optimal allocation according to Neyman (minimum coefficient of variation of the standard gross margin for the whole of Switzerland at *q=1*) and the equal distribution of the coefficient of variation between the strata (at *q=0*).
The more farms there are in the sampled population from a particular stratum and the greater the variability (dispersion) of the standard gross margin in this stratum, the more farms there should be in that stratum of the sample.[18]
Some strata (size categories) are **merged** or defined as **'take-all strata**' if the following requirements cannot be met:

1) The minimum accuracy requirement for strata (the relative error as measured by half of the 95 % confidence interval is limited to the range of ± 25 % of the estimated mean of the standard gross margin);19
2) The sampling fraction of a stratum must not be higher than 30 % (bearing in mind thelow response rates);
3) At least ten farms per stratum.

If the accuracy requirements cannot be met owing to a low number of farms or the high variability of the SGM in a stratum, then the size categories are merged. For some strata – especially small ones – the expected response rate cannot ensure that enough farms will be recruited if the sampling fraction is over 30 per cent of the sampled population. In this case, all farms of the sampled population of this stratum are contacted ('take-all stratum'), with the aim of recruiting as many farms as possible.

[16] 'Practical feasability' means that the number of farms supplying data must be realistic, bearing in mind the number of farms in the sampled population as well as the response rate.

[17] When calculating the coefficient of variation, only those farms in the range between the 2nd and 98th percentile are taken into account (excluding extreme values).

[18] Variability can be defined either in absolute terms as a standard deviation or in relative terms as a coefficient of variation. Up until AY 2014, the standard deviation (according to Formula 3.13 of Roesch and Lips, 2013) was used. From AY 2015 onwards, after consultation with the FSO, it was decided to use the coefficient of variation (according to Formula 2.2 of Bankier, 1988). Partly for this reason, but chiefly because of the introduction of stratum-specific size categories, for the same size of sample, it was not only possible to reduce the margin of error in the estimation of accuracy for the sample as a whole to ± 1.3 %, but it was also possible to improve the accuracy in numerous strata. This applies in particular to arable farms in the plain region, farms specialising in pigs and poultry, and all combined farms. Furthermore, it results in significantly fewer farms of type 1512 ('special crops') being needed.

[19] Here, the issue is the accuracy requirement for the strata, and not for the entire sample. This variable is defined according to formula 3.7 by Roesch and Lips (2013).

In the selection plan for AY 2016, this applies to farm types 1512 and 1541 in the mountain region, as well as farm types 1511 and 1531 in the hill region.[20]
The AY 2016 selection plan for the **net sample** Income Situation, calculated with the structural data from 2015, is shown in Table 5.

Table 5: Number of farms in the net sample according to the selection plan for AY 2016

	Farm type*											
Region	1511	1512	1521	1522	1523	1531	1541	1551	1552	1553	1554	Total
Size category 1												
Plain	30	78	30	11	21	17	33	40	19	60	55	394
Hill	12	27	46	11	37	22	34	19	17	42	42	309
Mountain	0	45	60	22	55	39	30	0	0	21	21	293
Total:	42	150	136	44	113	78	97	59	36	123	118	996
Size category 2												
Plain	20	34	21	11	13	13	12	22	11	30	30	217
Hill	29	12	31	9	16	59	14	15	15	23	19	242
Mountain	0	87	33	14	29	31	58	0	0	12	10	274
Total:	49	133	85	34	58	103	84	37	26	65	59	733
Size category 3												
Plain	31	48	31	0	0	0	11	28	21	35	35	240
Hill	0	15	47	12	20	0	15	0	0	32	25	166
Mountain	0	0	51	47	36	0	0	0	0	14	15	163
Total:	31	63	129	59	56	0	26	28	21	81	75	569

* Farm type according to S3 farm typology (see Appendix 1).

Since participation in the Income Situation survey is voluntary, only some of the contacted farm managers take part in it. This so-called **non-response problem** (see chapters 2.5.1 and 2.6.1) is counteracted by contacting a larger number of farms in order to reach the effective target of 2300 data-supplying farms as per the selection plan. The response rates (or willingness to participate) of previous recruitment campaigns, differentiated by strata, serve as the bases for these estimates. In strata which we know from previous experience to have low response rates, the number of farms to be contacted (gross sample) is increased to a larger extent than in strata with higher response rates (Figure 7 in Chapter 2.5.3). If, for example, only half of the farms contacted in a stratum provided their data in the previous year, then the number of farms now to be contacted is doubled, in order to achieve the target number of data-supplying farms. In the case of very low response rates, it may be necessary to contact all of the farms of a stratum ('take-all stratum'). In addition, the differences in willingness to participate depending on linguistic region are taken into consideration when the gross sample is determined. Since the Swiss-German farms have an above-average response rate compared to those of the French- and Italian-speaking regions of Switzerland, they would be over-represented in the net sample if this fact were not taken into account. Furthermore, the percentage of farms that have already taken part in the survey in previous years also plays a role (Figure 7 and Table 10 in Chapter 2.5.3). The higher the percentage of such farms in a stratum, the fewer the new farms that will need to be contacted in order to reach the sampling target.

[20] A number of strata were completely deleted from the AY 2014 selection plan owing to unrealistically high sampling fractions. This was the case for farm-type 1541 in the mountain region.

For the above-mentioned reasons, both the stratum- and language-specific response rates are taken into account when calculating the **gross sample** for the farms on the panel and for the new activations.
The results of the calculation of the number of farms to be activated in the gross sample of the 2016 accounting year are given in Table 6, and show that for AY 2016, there were a total of 6289 farms to be activated across all strata:

Table 6: Number of farms to be activated in the gross sample for AY 2016

	Farm Type*											
Region	1511	1512	1521	1522	1523	1531	1541	1551	1552	1553	1554	Total
												Size Category 1
Plain	94	180	84	17	72	71	83	68	33	144	164	1 010
Hil	37	134	99	19	136	75	103	47	39	65	120	874
Mountain	0	232	155	57	108	233	75	0	0	45	59	964
Total:	131	546	338	93	316	379	261	115	72	254	343	2 848
												Size Category 2
Plain	47	104	43	18	32	46	33	44	22	97	82	568
Hill	29	78	63	15	51	59	32	30	54	57	44	512
Mountain	0	87	60	24	43	150	58	0	0	27	29	478
Total:	76	269	166	57	126	255	123	74	76	181	155	1 558
												Size Category 3
Plain	67	295	70	0	0	0	31	120	62	102	114	861
Hil	0	58	122	31	89	0	42	0	0	76	100	518
Mountain	0	0	209	100	92	0	0	0	0	43	60	504
Total:	67	353	401	131	181	0	73	120	62	221	274	1 883

* Farm type according to S3 farm typology (see Appendix 1).

2.3.3 Selection process / Drawing the random sample

Based on the calculated number of farms to be activated in the gross sample (Table 6), a list of farms is drawn up which is used by the recruiting agency for getting in touch. All of the farms in the sampled population from the 2015 structural data survey (see Chapter 2.2) serve as the **sampling frame** for the random sample of the 2016 accounting year. Firstly, all farms that belong to the **panel** and took part in the survey the previous year (AY 2015) are selected without fail ('Old Members sample', abbreviated as 'S-Old). Secondly, those farms which took part in the sampling in the years preceding AY 2015 and which agreed to take part again, but which, for various reasons (e.g. because no financial bookkeeping was available by the submission deadline) did not supply any data for the 2015 accounting year, also form part of the panel. Such one-year absences are permitted for S-Old-members due to their relatively high willingness to participate. Thirdly, when drawing the random sample, priority is given to those farms not contacted the previous year.[21] Should there not be sufficient numbers of such farms in the sampled population of a stratum to reach the total number set by the gross selection plan, then, lastly, a drawing is also made from those farms which were asked to participate in previous years, but which either explicitly or implicitly refused (the most common grounds for refusal are listed in Chapter 2.5.2).

[21] This procedure, which will only be implemented from AY 2017 onwards, is also statistically justified. For one thing, when calculating the selection probability, it is important to distinguish between the farms that have never yet been selected, and those that were selected in previous years. If the farms are selected two years in a row, the probability of selection is cumulative. Moreover, we can assume that such farms have a lower willingness to participate.
Presumably, the behaviour or circumstances of those farms which would not/ counld not take part the first time, will not change greatly the following year.

This approach reduces the likelihood of the same farms being contacted two years in a row. In very small strata, however (e.g. arable farming in the hill region, with seventy farms), or in the strata with high sampling fractions (horse/sheep/goats and pigs/poultry), it may be necessary for all farms in the sampled population to be activated, to ensure that as many farms as possible are included in the sample.
In addition to the number of farms which, according to Table 6, are to be activated, so-called '**reserve packets**' are prepared for each stratum. Farms from these are only activated if it is noted during the course of recruitment that it will otherwise not be possible to reach the desired number of farms.
A simple random sample is drawn from each stratum while ensuring that the composition of German or French/Italian farms supplying the data matches as closely as possible that of the sampled population, bearing in mind the language-specific response rates. Since the response rate for the French- and Italian-speaking farms is usually lower, they are normally over-represented in the gross (i.e. to-be-contacted) sample.[22]
This selection process yields a **list of selected farms and** their corresponding identification numbers for the 2016 accounting year. In order to ensure anonymity, these lists are sent to the FOAG, where the identification numbers are supplemented with the relevant addresses and passed on to the recruiting agency. These lists are used by the recruiting agency to contact and/or to recruit the farms. Because of this separation, the FADN receives the accounting data in pseudo-anonymised form, i.e. without any personal details (e.g. name, address). For their part, the FOAG and the recruiting agency know only the names and addresses, and have no access to the supplied accounting data.

2.4 Recruitment and data collection

2.4.1 Requirements for participation and financial compensation

To be able to participate in the data survey, the selected farms must meet two important **requirements**. Firstly, they must keep **financial accounts**, including a profit-and-loss account and balance sheet. Farms that only keep a simple list of revenue, expenditure and asset components (also know as 'manual balancing' or 'cash accounting') may not take part in the survey. A second requirement is that the data be entered in an online questionnaire by the farmer or his accounting office no later than the latest possible **delivery deadline** in August.[23] If this deadline cannot be met, the farm cannot take part in the survey.

The data suppliers (farmers and accounting offices) are compensated for their participation in the data survey and the associated effort. CHF 260 in **compensation**[24] are available per completely plausible dataset supplied. If the farmer works with an accounting office, the monetary data must be recorded by the accountant in the online questionnaire. In this case, the accountant receives CHF 200 as compensation for the effort associated with compiling the accounting data. The farmer receives CHF 60 for the provision of his data and the recording of the non-monetary data (e.g. workforce or livestock numbers). If the farmer does his accounts completely on his own (i.e. individual entries and annual financial statement), he is responsible for the complete recording of his farm's data in the online questionnaire. In this case, he receives the total compensation sum provided for per farm. Starting with the 2019 accounting year, and regardless of whether they supply their data themselves or via an accounting office, farmers will receive, from the second year of supplying data onwards, an additional bonus (a so-called 'continuity supplement') of CHF 40 for their long-term participation on the panel. This compensation is only paid out if the questionnaire is filled in completely, and all data are plausible.

[22] If the number of Swiss French- and Italian-speaking farms in the sampled population of a stratum is exhausted, the sample is topped up with farms in the German-speaking area of Switzerland.

[23] The Federal Office of Agriculture's stipulation that the results on the income situation in the agricultural sector for year t must be published no later than the beginning of October of the year t+1 was decisive for the setting of this deadline.

[24] The compensation amount given here per data-supplying farm applies for accounting years 2017 and 2018, and may be revised at any time before the start of an Agroscope survey campaign in consultation with the Federal Office for Agriculture.

2.4.2 Data-collection process

This chapter provides a general synoptic overview of the recruitment[25], data acquisition and data-validation process, as well as the associated division of labour.
To ensure cost-efficient data acquisition, part of the work associated with the data collection process was outsourced. Annual recruitment of the participating data suppliers and the associated contacting of over 6000 farms are carried out by an external professional **recruiting agency** which is independent of agricultural interests and operates countrywide. This recruiting agency is administered by a survey and market research institute with an appropriate survey infrastructure.[26]. The recruitment design was developed and continually optimised during three test stages in 2011, 2012 and 2013 and in the set-up phase of the sample (2014 and 2015) in collaboration with the then-contractor of the recruiting agency (LINK Institute for Market and Social Research). To ensure the anonymity of the data survey, data acquisition and analysis are conducted separately from recruitment. Data acquisition, in which the recruiting agency is not directly involved, is performed with a **Web-based questionnaire**, structured and programmed in close collaboration with an accounting expert (AWeber GmbH, Utzenstorf, Switzerland) and an IT service provider (ELCA Informatik AG, Zurich). Validation of the supplied data is of crucial importance for safeguarding its quality, and is performed by an accounting office[27], the so-called ***fachliche Hotline und Plausibilisierungsstelle*** ('Expert Hotline and Validation Agency'). This agency also supports data suppliers in the completion of the online questionnaire.

Figure 2 provides a simplified representation of the data-collection process, using the example of the **Survey Campaign 2017** (2016 accounting year). In October of year t, the **sample is drawn by Agroscope** on the basis of the structural data made available by the FOAG and as per the prepared selection plan (see Chapter 2.3). From November of year t and until the end of March of year t+1, **the recruiting agency recruits** the farmers (and, if applicable, their accounting offices) according to the selection list previously drawn up by Agroscope. At the time of their recruitment, the data suppliers can choose from among three delivery deadlines: end of April/beginning of May; end of May, beginning of June; around 10 July. The online questionnaire may be completed from the first half of January of year t+1, i.e. only after the surveyed accounting year has finished.

From the start of recruitment to the conclusion of the data-collection campaign, the data suppliers have access to two **hotlines**. The recruiting agency's hotline (first-level hotline) answers all general or organisational questions about data collection asked by the data suppliers. The Expert Hotline and Validation Agency (second-level hotline) is available to the data suppliers for questions concerning the collected data or the online questionnaire. It supports farmers and their accountants with data collection and with solving any problems arising in this connection.

There are four stages of dataset quality control (for details, see the quality assurance scheme in Chapter 2.4.5). The first two automated stages are built into the online questionnaire. Stage one occurs directly when data is entered in the forms, and stage two before completion of data collection. The third checking stage is performed by the Expert Hotline and Validation Agency. The last quality-assurance stage is carried out by the Swiss FADN itself on the complete dataset imported into the FADN database. Only once all four stages are successfully completed is a dataset considered completely plausible.

[25] 'Recruitment' in this context means the acquisition of the farm and its accountant for participation in the data survey.

[26] For the time period 2018 to 2022, the recruitment contract is administered by the LINK Institute for Market and Social Research (LINK Marketing Services AG, Lucerne). The recruitment contract was awarded in a public WTO tendering procedure.

[27] TSM Treuhand GmbH (Bern) is the contractor for the Expert Hotline and Validation Agency for the period 2018 to 2022. The hotline contract was awarded in a public WTO-tendering procedure.

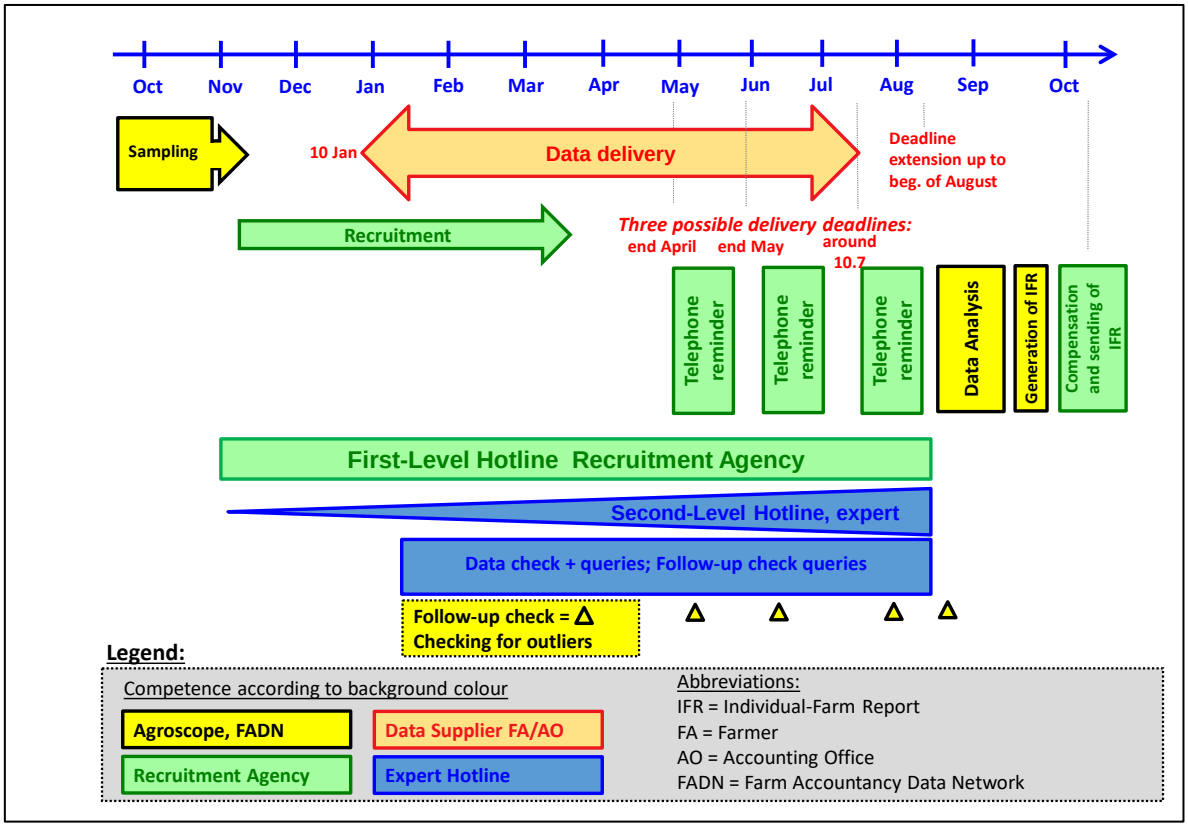


Figure 2: Simplified illustration of the entire data-collection process, using the example of the 2017 survey campaign.

If the data are not delivered by the deadline agreed with the data supplier, the recruiting agency reminds the data supplier of the pending data delivery in up to two so-called '**reminder interviews**', and a new delivery deadline is jointly set.

Once the last regular delivery deadline (around 10 July) has passed, the recruiting agency contacts any data suppliers that are still in arrears (so-called '**final reminder interview**') and grants them a deadline extension of three weeks (until the evening of the first working day in August) for the data collection.
After expiry of the extension, the data undergo in-depth analysis by Agroscope, and each participating farm receives by post an '**individual-farm report**' (see Chap. 2.7.2), together with the financial compensation by postal cheque. The results of the **analysis** are published by Agroscope in early October.

Interface coordination and real-time data-collection monitoring

As can be seen from the process overview, several institutions are involved in the data collection process. A special tool – the **CRM-OTRS** – was developed to coordinate the interfaces of recruiting agency, Agroscope (Farm Accountancy Data Netwok), Expert Hotline and Validation Agency, and online questionnaire. All three partners have access to this tool. This tool consists in essence of a web-based open ticket requesting system (OTRS), which was – through the incorporation of a customer database into it – supplemented by a customer relationship management (CRM) module. The OTRS is a system used to manage enquiries and contact (whether by email or telephone) with the data suppliers. It performs a crucial function in the data collection process. All email correspondence with the data suppliers takes place via the OTRS, and all telephone contacts with the data suppliers are documented there. Through the integration of the CRM module, the automatic sending of emails from the questionnaire to the CRM, and the linking of the CRM database with the recruiting agency's recruitment database, the CRM-OTRS enables a complete real-time monitoring of the survey process for each activated farm, whilst respecting the data protection and anonymity regulations.

The CRM-OTRS is used in particular by the Expert Hotline and Validation Agency for data control and correction queries. Each ticket is assigned a customer number (farm ID) either automatically or manually. This allows the link to be established with the CRM database, which contains all customer information. The CRM database is in turn automatically or manually updated or added to on the basis of the opened and closed tickets. The main advantage of this system is its ability to handle very high volumes of calls and emails whilst allowing the employees involved in the whole data collection process to keep track of everything. At the same time, this system ensures complete traceability. Moreover, current information on data delivery, be it information on a specific data supplier or reportings on the status of the data survey, can be retrieved from the CRM database at any time. The accounting information itself is not captured in this tool.

2.4.3 Recruitment process

The entire recruitment process can be broken down into two stages:

- In the **recruitment stage**, the work of the recruiting agency consists chiefly in contacting the farmers (and in the 'accountant' case, their accounting offices as well) according to the selection plan predefined by Agroscope, and in persuading them to take part in the survey.
- During the **survey stage**, the recruiting agency reminds the data suppliers to record their data in the online questionnaire if they have not done this by the agreed deadline.

Once the survey phase is over, participants are compensated at the same time as the individual-farm reports are sent to the participating farms and accounting offices. Although these two tasks are not part of the recruitment process in the strict sense, they are performed by the recruiting agency for reasons of data protection.[28]
The main individual steps of the recruitment process are described in a very simplified way below.[29]

Step 1: Sending the annoucement letters

In the lead-up to phone contact, the recruiting agency informs the randomly-selected farmers by letter about the data survey to be conducted.

Step 2: Recruiting the farmers

The recruiting agency makes contact with the farm manager, informs him about the aims, object and procedure of data collection, and motivates him to participate in the survey.
If consent is obtained, the agency ascertains whether the farmer keeps his own financial accounts ('farmer' case), or whether he has an accounting office draw up his accounts ('accountant' case). If the farmer keeps his own accounts and takes part in the survey, a deadline for data delivery is agreed. If the farmer refuses to take part in the survey, the reason for refusal is recorded.

Step 3: Recruiting the accounting office ('accountant' case)

If an accounting office is responsible for the end-of-year accounts of the farm, it is contacted by the recruiting agency, informed about the aims, object and procedure of data collection, and motivated to take part. If financial accounts are available for the farm, a data-delivery deadline is agreed. If the accounting office refuses to take part, the reason for refusal is recorded.

[28] For reasons of data protection, Agroscope does not know the contact data of the participating farms. Accordingly, Agroscope can neither compensate the participating farms directly, nor send the individual-farm report to them directly.

[29] This description does not claim to be comprehensive. A detailed, in-depth description can be found in the requirements specification of the WTO invitation to tender for the recruiting agency (Agroscope, 2017a).

Step 4: Sending the access data for the online questionnaire and, if applicable, the farm-manager's questionnaire in paper form

The recruiting agency sends the person (farmer or accounting-office staff) responsible for entering the accounting data into the online questionnaire an email with the necessary access data. In the 'accountant' case, the farm manager receives an email message with the access data for completing the farm-manager's questionnaire (part one of the online questionnaire, see also Chapter 2.4.4). Should the farm manager be unable or unwilling to complete the farm-manager's questionnaire online, he will receive a paper version of the latter by post from the recruiting agency.
The online questionnaire is generally available for data collection from the beginning of January onwards.

Step 5: email/letter reminders for the farm-manager's questionnaire (in the 'accountant' case)

The farm managers are reminded by letter or by email to complete the farm-manager's questionnaire online or on paper and to return it to their accounting office, if they have not yet done so by the time of the reminder.

Step 6: Reminder interviews before each delivery deadline

If no data has been delivered two weeks before the agreed delivery deadline (or three weeks before the final delivery deadline), the recruiting agency sends a reminder email to the data supplier. Once the agreed deadline has passed, the recruiting agency gets in touch by telephone.

Step 7: Final reminder interview

Once the final regular delivery deadline (around 10 July) has passed, the recruiting agency contacts by telephone all data suppliers that have not filled-out the questionnaire by this deadline, and grants them a final deadline extension of three weeks (until the evening of the first working day in August).

Step 8: Sending of refusal email

Once the final delivery deadline of beginning of August has passed, the recruiting agency notifies by email those farmers who completed the farm-manager's questionnaire (themselves) online, but whose accounting office has not completed the relevant online questionnaire, that they will not receive any compensation.

Step 9: Financial compensation of the farmers and accounting offices

Around two months after completion of data collection, the recruiting agency sends each participating farm a letter of compensation containing a postal cheque for the financial compensation and the individual-farm report. The financial compensation is also transferred to the participating accounting offices, and a detailed billing is sent by post.

2.4.4 Questionnaire design and data collection

The online questionnaire used for the data collection was developed from scratch in the three test phases, and is continually adapted and improved for data-quality purposes.
The data required in order to assess the economic situation of the Swiss agricultural sector cover various areas of an agricultural enterprise and household. The questionnaire compiles non-monetary and monetary indicators for the following four areas:

- **Basic data** on the farm (e.g. year taken over, form of cooperation). Some of the basic data is needed for navigation through the different questionnaire forms and to ensure data quality (e.g. sole proprietorship or group farming business/generational community, ownership of business properties as private or business assets).

- **Personal details** on family members, employees and their labour input.
- Whole-farm data from the **financial accounting** (balance sheet, income statement and other revenue/ expenditure within the financial accounting).
- For individual enterprises only: supplementary data from the **tax return**[30] concerning income and private consumption outside of the financial accounting (e.g. non-agricultural self-employment income or wage income, social security and pension contributions).

Additional non-monetary farm information such as land use (area of the different crops) is available from the Farm Structure Survey database, and does not need to be collected in the questionnaire. In order for the use of these data to be legally possible, the farm manager must consent to the matching of the questionnaire- and FSS-data in a statement included in the questionnaire. Without this consent, participation is not possible.

Although detailed livestock figures are available in the FSS database, the average animal population there is based on the average livestock numbers for the survey year t-1 from the TVD (Swiss Stock Movement Database). For this reason, aggregated information on the animal populations is additionally collected in the online questionnaire. If livestock figures in the FSS database differ too greatly from the data recorded in the questionnaire and if the data supplier confirms the accuracy of the latter, then, exceptionally, the livestock figures in the questionnaire are used for this farm.

In addition to the data collection forms, the questionnaire contains the following forms**:**

- Data control/plausibility checks forms for **ensuring data quality** (Chapter 2.4.5)
- Contact details for queries to the extent that this is permitted by data protection, and details for paying the accountant via IBAN transfer
- Feedback on the questionnaire

The non-monetary farm data to be recorded in the questionnaire as well as the declaration of consent (Figure 3) must be provided by the farm manager. If the farm manager has no access to the online questionnaire, he can complete a paper version of this part of the questionnaire[31] and send it to his accountant, who then enters the data in the online-questionnaire. The monetary data both within and outside of the financial accounting are generally supplied by the farm's accountant. An exception to this are farms that record the business transactions in the financial accounting records and prepare the financial statement themselves; these farms complete the entire questionnaire themselves online (Figure 4). These three cases are taken into account in the design of the questionnaire, with different login accesses and corresponding branchings in the questionnaire.

[30] The tax assessment decision from the Tax Administration need not be available at the time of data entry. Provisional amounts may also be entered.

[31] This part of the questionnaire is also called the 'farm manager's questionnaire'.

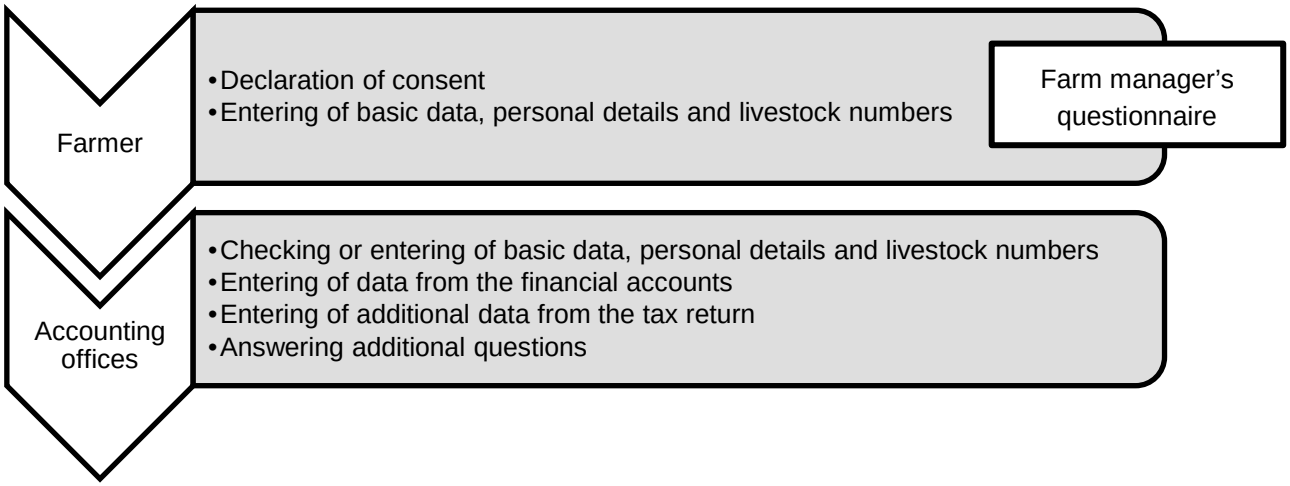


Figure 3: 'Normal data-input' scenario: Farm manager completes farm manager's questionnaire; accountant provides data from the financial accounting and tax return, and completes the questionnaire.

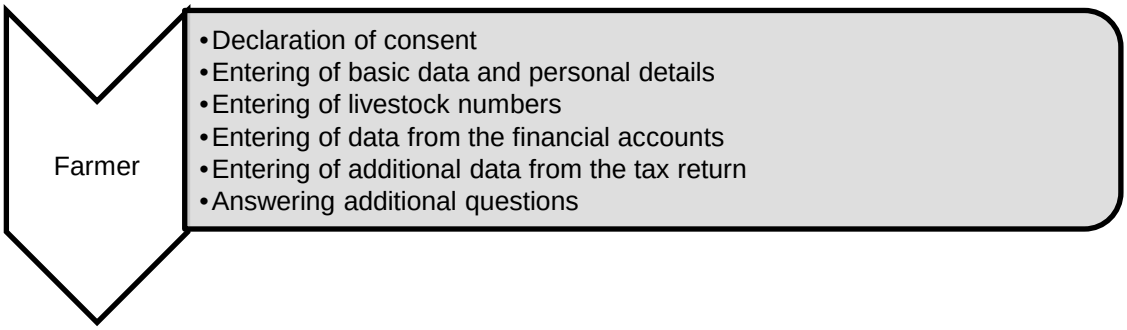


Figure 4: 'Farmer records business transactions in financial accounting records and prepares annual financial statements himself' scenario: Farm manager fills in the entire questionnaire himself, and completes data entry.

Various functionalities in the questionnaire provide the data supplier with data-entry support. Descriptions and information on the collected data are incorporated directly in the questionnaire, both as introductory text in the individual forms and as information texts right beside the input field.[32]

The entry form for the profit and loss account has a **help function** which allows the data supplier to trace the profit and loss account figures entered in the questionnaire. This function is particularly helpful with troubleshooting when one's own profit and loss account or chart of accounts differs from that shown in the questionnaire. This function is used to record, for the individual figures, the account category of one's own profit and loss account from which the entered figure originates. Here, a figure can originate from a single or several account categories of one's own profit and loss account. By using this function, the entries can be displayed as a profit and loss account structured according to one's own chart of account in a data-control sub-form and compared with one's own year-end closing.

During the data-supply period of January to August, the online questionnaire may be viewed via the officially accessible, fictitious model farm.[33]

[32] Model financial statements showing which bookkeeping figures are to be entered in which fields of the questionnaire are also available for some bookkeeping programs.

[33] www.einkommenssituation.ch → "Datenlieferung" (Data supply) → "Online-Erhebungsbogen" (Online questionnaire) access data: Username: LBeispiel, Password: 111

Merging the datasets from the online questionnaire and the farm structure survey (FSS)

After each data delivery deadline, the completed datasets – if necessary, checked by the Expert Hotline and Validation Agency – are exported from the online questionnaire, then merged with the appropriate Farm Structure Survey FSS datasets. For the first two deadlines, only FSS data from survey year t-1 are available. From July onwards, the newly available datasets as well as those completed previously are merged with the FSS data of survey year t and definitively imported into the database. A joint farm identification number is available for merging the two datasets.

2.4.5 Quality assurance through plausibility checks and data cleansing

Data quality assurance for the Income Situation sample takes place at various stages (Figure 5). The Expert Hotline offers support to data suppliers, should they have any problems or questions. To do so, it uses the CRM-OTRS tool (Chapter 2.4.2) for contact with the data suppliers as well as for assigning statuses regarding data delivery and plausibility checks, financial compensation, and receipt of an individual-farm report (IFR).

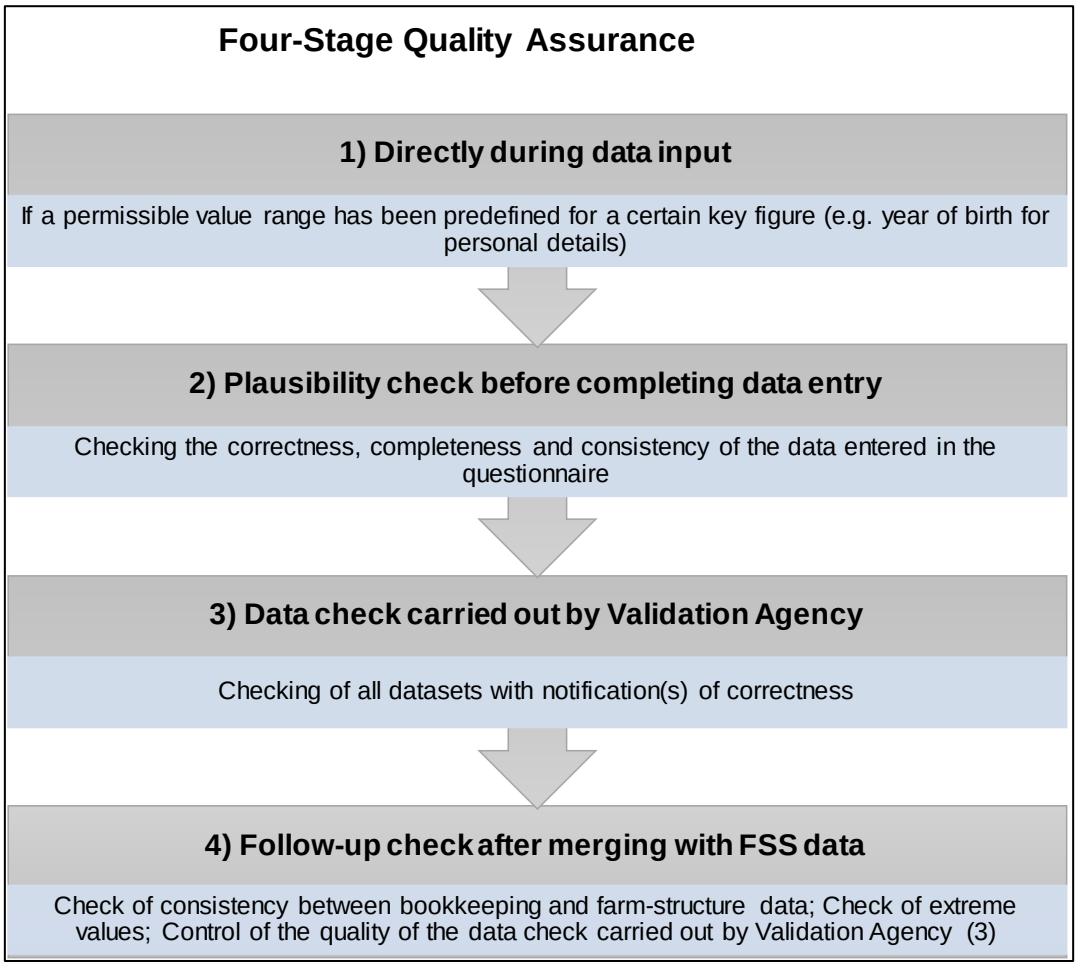


Figure 5: Illustration of the four-stage quality assurance process during data collection for the Income Situation sample.

The first two quality-assurance stages occur in the online questionnaire directly **when entering the data** and as **plausibility checks** after all data have been entered and before closing the questionnaire. In the questionnaire, data are automatically checked when the user switches to the 'plausibility checks' form.

The plausibility checks verify the correctness and completeness of the data entered, and encompass the following issues:

- Consistency between balance sheet and profit and loss account
- Consistency between opening and closing balance sheets and inventory change
- Verifying the completeness of financial accounting items whose entry is obligatory
- Consistency of dependent data (e.g. presence of employed labour force vs. staff costs, working days for salaried activities vs. income from salaried activities, livestock numbers vs. revenue from animal production, etc.)
- Large year-to-year change in annual profit
- Large year-to-year change in labour-force numbers
- Large year-to-year change in livestock numbers

The results of the data plausibility check are displayed to the data supplier in a separate form of the web-based questionnaire. If the plausibility checks are not passed, entries must be checked and corrected in order to complete data collection. In the case of a failed plausibility check, if the data supplier is sure that the data entered are correct and consistent, he can provide his explanation in the form of a so-called 'notification of correctness'. This enables him to complete data collection despite the failed plausibility check. The third stage – **data checking** – is administered by the Expert Hotline and Validation Agency in a timely manner once the data has been supplied. All datasets containing notifications of correctness are checked. In the case of non-plausible notifications of correctness, the Validation Agency contacts the data supplier and asks him to correct his data. Only once all plausibility checks have been passed or all notifications of correctness are plausible does the supplier's dataset count as delivered and eligible for financial compensation for participation in the survey.

The fourth stage – the **follow-up check** – is carried out by Agroscope's Farm Accountancy Data Network after each data delivery deadline, i.e. four times per survey year. After the datasets from the online questionnaire are merged with the farm structure survey data and the key bookkeeping figures are calculated, the consistency of the two merged datasets (FADN questionnaire and FSS data) is tested, and farms with extreme values for important key figures are checked.

The checks cover the following issues:

- Extreme values for agricultural income or working income
- Extreme values in the balance sheet (liquid assets, equity)
- Non-family and family labour input compared to farm size, wages and social security contributions
- Inconsistencies between the recorded revenues in animal and plant production and the corresponding livestock numbers and crop areas.

If outliers or inconsistencies are detected at this stage, the FADN asks the Expert Hotline and Validation Agency to get in touch with the data supplier and to clarify with him the causes of the problems observed. Where necessary, the Expert Hotline and Validation Agency corrects the information provided by the data supplier in the online questionnaire.

The FADN then decides whether the corrections made or the explanations given can be accepted, and whether the dataset is usable for analysis. If the Expert Hotline and Validation Agency is unable to contact the data supplier owing to time constraints or for other reasons, non-plausible datasets are excluded from the analysis.

2.5 Response rate and representativeness

The Income Situation sample is a survey based on voluntary participation. Data for this survey is collected using a fairly complex and comprehensive questionnaire, and captures information pertaining to sensitive topics, including the income situation of farming households. Willingness to participate in such surveys is often relatively low (Moore and Welniak, 2000). Despite the great efforts made and improvements in the recruitment and data-collection processes, only around one-third of all contacted farms (including those that contributed their data in previous years) take part in this survey. Because low response rates can lead to numerous problems (poor precision, estimation bias, lack of representativeness of certain groups of farms), they deserve closer analysis.

2.5.1 Supplied farms

Below, the representativeness of the survey and the influencing factors are explained, using the 2016 accounting year as an example. During the entire survey period for AY 2016, which includes a post-recruitment of reserve packages in some strata (see Chapter 2.3.3), a total of 6293 farms were **activated**. Among the activated farms were 2796 farmers who had already taken part in the survey in previous years (Old Members sample, abbreviated as S-Old) and 3497 'new' farms (New Members sample, or S-New) that had never supplied their data to the sample.[34] Some of the farms could not be reached owing to invalid addresses or missing or wrong telephone numbers, whilst a further portion of contacted persons state that they do not run a farm business. This applies to a total of 520 farms (8 % of the activated addresses), and is treated as a **neutral non-response** (Table 7).

Despite several attempts to contact them by telephone, some of the farmers with valid contact addresses (around 5 %) **could not be reached**.
These instances of non-response cannot be treated as neutral, since the contactability of the farms may be correlated with one or more variables of interest.

Although the smallest farms were excluded from the sampled population, the percentage of **farms with no financial accounting** – representing on average 7 % of all farms with valid contact addresses – is relatively high. This applies mainly to the new recruits (where the percentage of farms with no financial accounting stands at 14 %). In particular, 'special crops' farms and farms specialising in horses, sheep and goats often just keep simple records with no balance sheet or profit and loss account, and for this reason are not permitted to take part in the survey (see Chapter 2.4.1).

All in all, 51 % of the activated farms with valid contact details **refused** to take part in the survey, with 29 % declining to take part from the outset, 3 % of the farms dropping out during the reminder process, and 18 % of the farms never completing data collection, despite having agreed to participate at the outset ('tacit' refusal).

The accounting data provided by 2115 of the 2117 supplied farms could be linked with the structural variables of the FSS database. The difference in this case of two farms can be attributed e.g. to farm mergers from one year to the next.

[34] The New Members sample also includes those farms which were selected and contacted in previous years but have not yet provided any data, or those that took part in the survey more than two years ago and have supplied no further information since then.

Table 7: Number of farms according to recruitment status (AY 2016)

Category	S-Old[1]	S-New[2]	Total	Percentage[3]
Activated farms	2 796	3 497	6 293	100 %
Invalid contact details and other reasons for neutral non-response	116	404	520	8 %
Activated farms with valid contact details	2 680	3 093	5 773	100 %
Farms with no financial accounting	14	418	432	7 %
Non-response owing to refusal	932	1 987	2 919	51 %
of which declined to take part during recruitment	407	1 244	1 651	29 %
of which declined to take part when reminder sent	57	143	200	3 %
of which tacit refusal	468	600	1 068	18 %
Farms not reached	38	267	305	5 %
Farms supplying data	1 696	421	2 117	37 %
Evaluable datasets	1 686	408	2 094	36 %

[1] S-Old ('Old Members sample') are the farms which have taken part in the survey in previous years.

[2] S-New ('New Members sample') are the farms recently selected for the sample

[3] The percentage of farms with invalid contact details refers to all of the activated farms. All other percentages relate to the farms with valid contact details (valid sampled population).

In the course of the plausibility check or follow-up check, it may be necessary to exclude additional farms taking part in the survey from the evaluation. In AY 2016, 21 farms with extreme or non-plausible data were excluded from the subsequent evaluations, with the result that 2094 **evaluable farm operating results** were available. This represents 36 % of the farms with valid addresses, and 33 % of all activated farms.

As can be seen from Table 7, 19 % of the effective sample consists of **newly recruited farms** (S-New, 408 farms). An overwhelming percentage of the farms (81 %) have been supplying their data to the FADN for at least two years, and belong to the **panel** (S-Old). Such farms are markedly more willing to participate than those that have never taken part in the survey. Over 64 % of the S-Old farms approached took part in the survey again. This is due on the one hand to the lower non-response rate owing to invalid contact details or a lack of financial accounting, and on the other to the far lower frequency of refusals during recruitment for S-Old farms than for S-New farms (15 % vs. 40 % of the valid contact details, respectively). Although willingness to participate increases along with time served on the panel, the continual **farm dropout from the panel** means new farms must be recruited in order to maintain the desired sample size of 2200–2300 farms. Moreover, the new recruits ensure that the sample accurately reflects the current structure of the sampled population in terms of stratification characteristics. This adjustment is particularly important, given the ongoing structural change in the agricultural sector.

Because of the high non-response rate among the activated farms that have never taken part in the sample, four to ten addresses had to be activated for each dataset received, depending on the stratum. Out of the 3497 newly activated farms in total, 408 evaluable datasets were included in the sample. The most common reasons for non-response among the new recruits are given in the following chapter.

2.5.2 Most-common reasons for non-response among new recruits

For the 2016 accounting year, 14 % of the newly activated farms with valid contact data throughout Switzerland provided complete and plausible datasets. In AY 2015 the figure was still 20 %, and in AY 2014, 23 % across all strata. The decrease is largely due to the strata that are difficult to recruit. These strata have a large year-on-year non-response rate.

In order to offset this, a great many new farms must be activated. If the number of farms that have never previously been selected for the sample is insufficient, then the farms which declined to participate in previous years are reapproached. Such farms generally have an even lower willingness to participate than farms being activated for the first time.

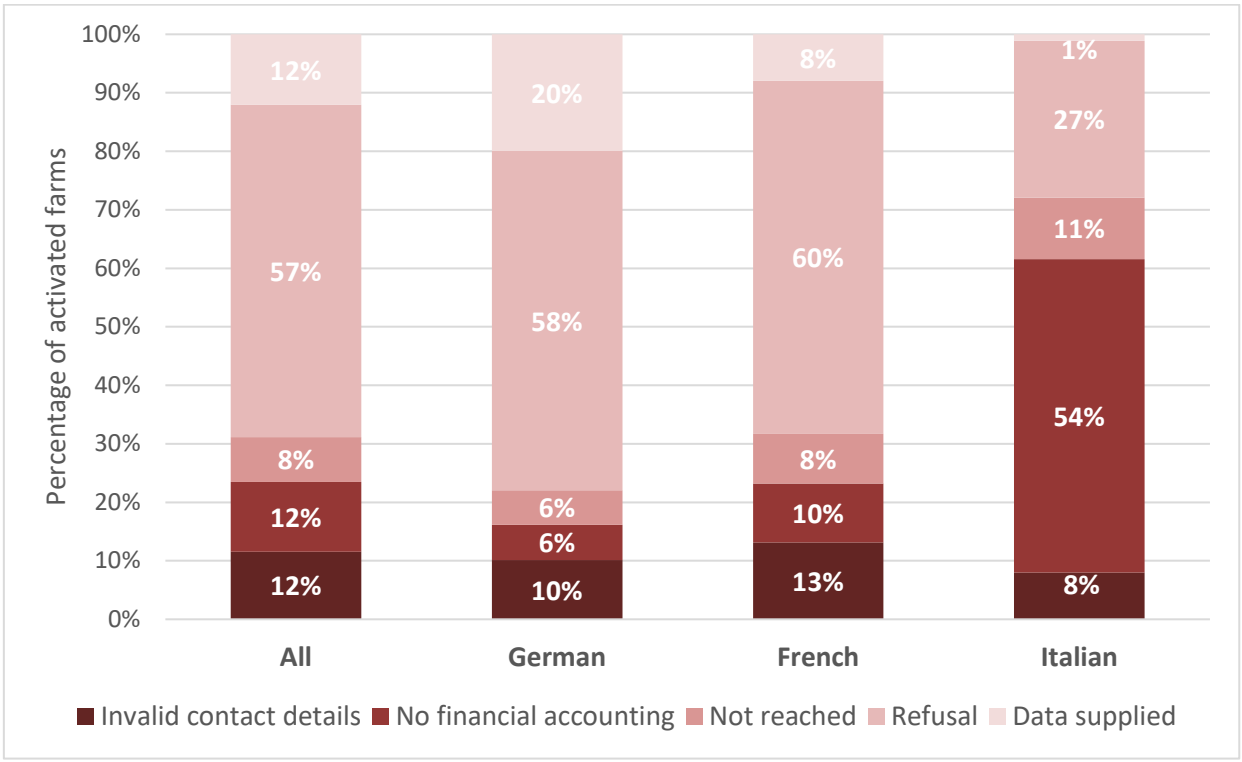


Figure 6: Reasons for non-response among new recruits according to linguistic region (AY 2016).

Response rate varies significantly according to **linguistic region**. At 22 %, it is highest in German-speaking Switzerland (where it was 24 % in AY 2015). French- and Italian-speaking Switzerland record very low response rates of 9 % and 1% respectively. There are major differences between the three linguistic regions, not just in terms of the non-response rates, but also as regards the relative importance of the various reasons for non-response (see Figure 6). In the French-speaking regions, there are major problems with the quality of the contact details of the farms listed in the FSS database. Addresses and/or telephone numbers are unavailable or invalid for 13 % of the newly activated farms in French-speaking Switzerland. In addition, 8 % of the French-speaking farms could not be reached despite several attempts to contact them by telephone. At 10 %, the percentage of farms excluded from the survey due to the absence of financial accounting is higher in French-speaking Switzerland than in German-speaking Switzerland, where the figure is just 6 %. This problem is even greater in the case of the Italian-speaking farms, 54 % of which carry out no financial accounting, which is why they are unable to supply any data to the FADN.

Besides the already-mentioned problems of invalid contact details (12 % of newly activated farms) and an absence of financial accounting (12 % of the newly activated farms, or 14 % of the farms with valid contact details), **refusal to participate** (on the farmer's or accountant's part) represents a most common reason for non-response among the new recruits. Over 64 % of the farms with valid contact details and 57 % of all activated addresses drop out of the sample owing to a refusal to participate.

Table 8: Statistics on the reasons for refusal (active, explicit refusal) in the case of new recruits (AY 2016)

Reasons for Refusal	Number	Percentage
Deadline for supplying data too early	285	21 %
No time / No interest	260	19 %
Not part of the field of observation	134	10 %
Protest against FOAG / agricultural policy / Agroscope	121	9 %
Personal reasons	110	8 %
Too heavy a workload	106	8 %
Does not think that farm figures are meaningful	77	6 %
Accountant recommended refusal	68	5 %
Data-protection concerns	46	3 %
Low compensation / high effort	25	2 %
Questionnaire too complex	30	2 %
Other reasons	125	9 %
Total refusals (at time of recruitment and reminder interviews):	**1 387**	**100 %**

We can distinguish between an **explicit** and an **implicit refusal**. In the first case, the data supplier (farm manager or accounting office) explicitly declines to participate in the survey during the recruitment interview or during a reminder interview, with the recruiting agency noting the reasons for refusal according to the predefined categories. In the case of implicit (tacit) refusal, although the recruiting agency receives a positive response from the data supplier (at both the recruitment and reminder interviews), the latter fails to supply complete and plausible data by the last possible delivery deadline in August. Around one-third of all refusals (600 cases) fall into the 'implicit' ('tacit') category.

The frequencies of the **reasons given for refusal** in the case of active refusals are summarised in Table 8. Most (21 %) of the farms declining to participate cannot take part in the survey because the last (official) data-supply deadline in July is too early for them, and their financial accounts are not available at this time. A lack of time or lack of interest in the survey is mentioned as the second most common reason for refusal (19 %). During the interview, one in ten farms was found not to belong to the field of observation (because, for example, the farm was a legal entity, or because the farm manager did not view his business as an agricultural holding). Around 9 % of the refusals can be put down to a protest against the FOAG or against agricultural policy. Refusals for personal reasons and owing to a heavy workload are also common. Too-low compensation for participation in the survey, the high level of complexity of the questionnaire or concerns about data protection were seldom mentioned as the reason for refusal.

At 86 %, the percentage of newly activated farms with valid contact details which for some reason do not participate in the survey is very high, and leads to very low response rates. Chapter 2.5.3 below deals with the calculation of response rates, which is done differently depending on the issue and purpose in each case.

2.5.3 Response rates and retention rates

The response rate is considered to be an important indicator of the quality of the survey, since low response rates can lead to biases in the estimation of the variables of interest.
The **simple (unweighted) response rate** ('Response Rate Unweighted', **RRU**) is defined as the ratio of the number of supplied (evaluable) datasets to all activated and valid contact addresses[35]:

[35] Alternatively, the number of all activated farms could be used as the basis, which yields a lower response rate. The missing or invalid contact details are the result of the quality of the personal details in the FSS database, and may be treated as a neutral non-response, assuming that these details are not correlated with willingness to participate, or with the variable of interest.

$$RRU = \frac{\text{No.of supplied (evaluable) datasets}}{\text{No.of activated farms with valid contact details}} \quad (2)$$

To evaluate **willingness to participate or recruitment success,** we use the RRU1, which is calculated according to the following formula for all farms supplying data:

$$RRU1 = \frac{D}{D+IR+ER+NF+NC} \quad (3)$$

To take account additionally of the **quality of the collected data**, the *RRU2* is calculated with evaluable datasets (supplied datasets minus subsequently excluded non-plausible data):

$$RRU2 = \frac{D-NP}{D+IR+ER+NF+NC} \quad (4)$$

The meanings of the abbreviations used in the formulas are given below:

D – Number of delivered, complete and plausible datasets (**D**elivered)
NP – Number of delivered datasets subsequently classified as non-evaluable and excluded from the evaluation (**N**ot **P**lausible)
IR – Number of recruited farms that have not supplied any data (**I**mplicit **R**efusal)
ER – Number of farms where the farm manager or his accounting offices have refused to participate during the recruitment or reminder interview (**E**xplicit **R**efusal)
NF – Non-response owing to lack of financial accounting (**N**o **F**inancial Accounting)
NC – Number of farms with valid contact details which could not be reached by telephone despite several attempts at contact (**N**o **C**ontact)
Table 9 presents the two total **unweighted response rates** (RRU1 and RRU2) of the newly activated farms that delivered complete (or evaluable) datasets for AY 2014–2016.

Table 9: Total response rates of the S-New farms

Farms newly activated in the accounting year in question (S-New)	**AY 2014**	**AY 2015**	**AY 2016**
a) No. of activated S-New with valid contact details:	5 891	4 039	3 019
b) No. of S-New delivering complete and plausible data:	1 375	826	421
c) No. of S-New delivering evaluable data:	1 348	817	408
RRU1 (b/a)	23.3 %	20.5 %	13.9 %
RRU2 (c/a)	22.9 %	20.2 %	13.5 %

The response rates vary greatly, depending on farm type, region and farm size. From Figure 7, we can see that the probability of participation is much lower than the average in the case of Special Crops farms (type 1512) and farms specialising in horses, sheep or goats (type 1531). At 2.8 %, the response rate for Special Crops farms in the mountain region is extremely low. On the other hand, Commercial Milk farms (type 1521), Pigs and Poultry farms (type 1541) and Combined Pigs and Poultry farms (type 1553) are characterised by a higher willingness to participate.

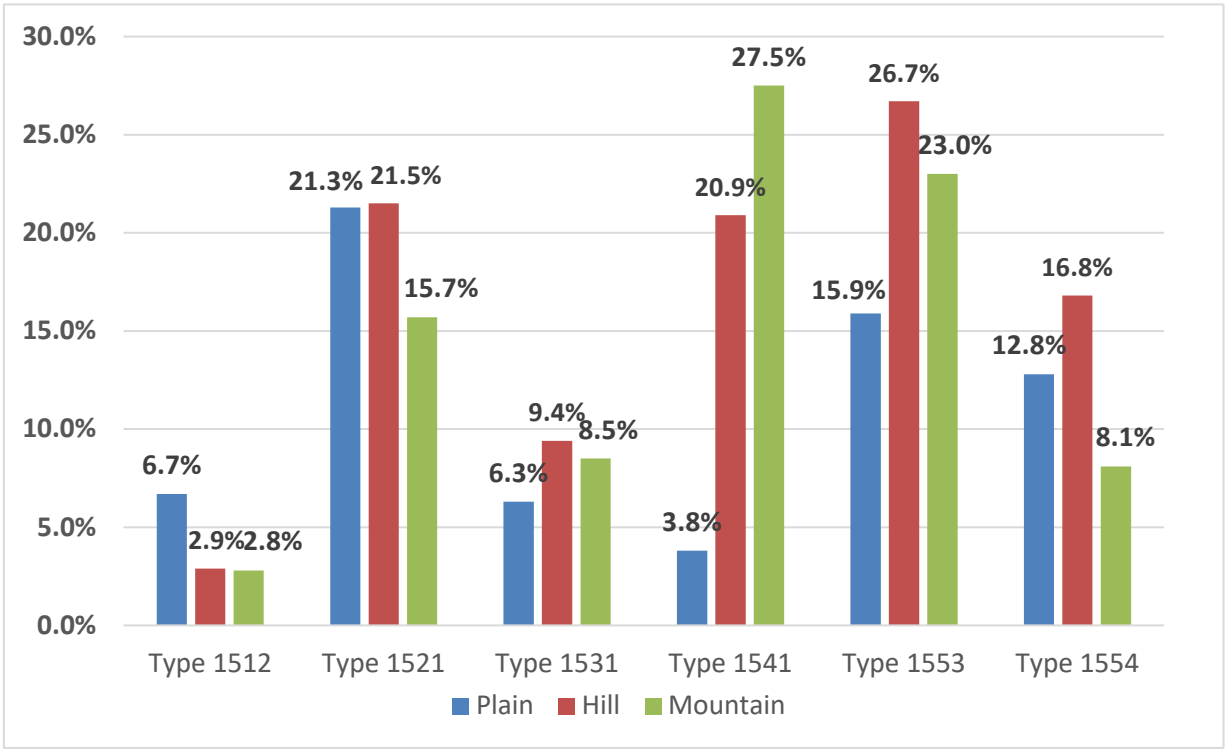


Figure. 7: Unweighted response rates (RRU2) of the S-New farms for selected farm types (see Appendix 1) and regions.

The **weighted response rate** is used to evaluate the success of the data survey in terms of reflecting the sampled population via the sample. Depending on farm type, region and size category, the farms have different probabilities of being selected for the stratified random sample. If the farms are allocated non-proportionally in the sample, these different sampling fractions or selection weights must be considered when calculating response rates.[36]

For this, the following formula is used for the weighted response rate (RRW[37]):

$$RRW = \frac{\sum d_i D_i}{\sum d_i (D_i + IR_i + ER_i + NF_i + NC_i)} \quad (5)$$

d_i – Sample-design weight (reciprocal of the sampling fraction of the gross sample)
D_i = 1 if Farm i has supplied the data and 0 otherwise
IR_i = 1 if farm recruited, but no data delivered (implicit refusal) and 0 otherwise
ER_i = 1 if Farm i refused to supply data (explicit refusal) and 0 otherwise
NF_i = 1 if Farm i has no financial accounting and 0 otherwise
NC_i = 1 if Farm i not reached by telephone and 0 otherwise

Comparing the weighted response rate RRW with the unweighted response rate RRU1 (see Figure 8), it becomes clear that the composition of the farms according to the selection plan plays an important role in the level of the average response rate. The fact that the weighted response rates turn out to be higher than the unweighted ones suggests that farms from difficult-to-recruit strata (e.g. Special Crops farms, type 1512, and Horse/sheep/goats, type 1531) are given a lower weighting owing to the higher sampling fraction.

[36] The exact procedure for calculating the weights is explained in Chapter 3.6.2.

[37] As with RRU1 and RRU2, a distinction can also be drawn here between RRW1 and RRW2, depending on whether all of the delivered and complete datasets, or just the evaluable datasets, are taken into account. Admittedly, these variants differ only minimally from each other, which is why, for simplicity's sake, we focus on just the first variant.

Although only 14 % of the farms selected for the gross sample of AY 2016 took part in the survey, the net sample represents 17 % of the sampled population. Moreover, when calculating the weighted response rate, further auxiliary variables (e.g. standard output) can be taken into account, in order to gauge the influence of the non-response on the quality of the estimate of the variable of interest.

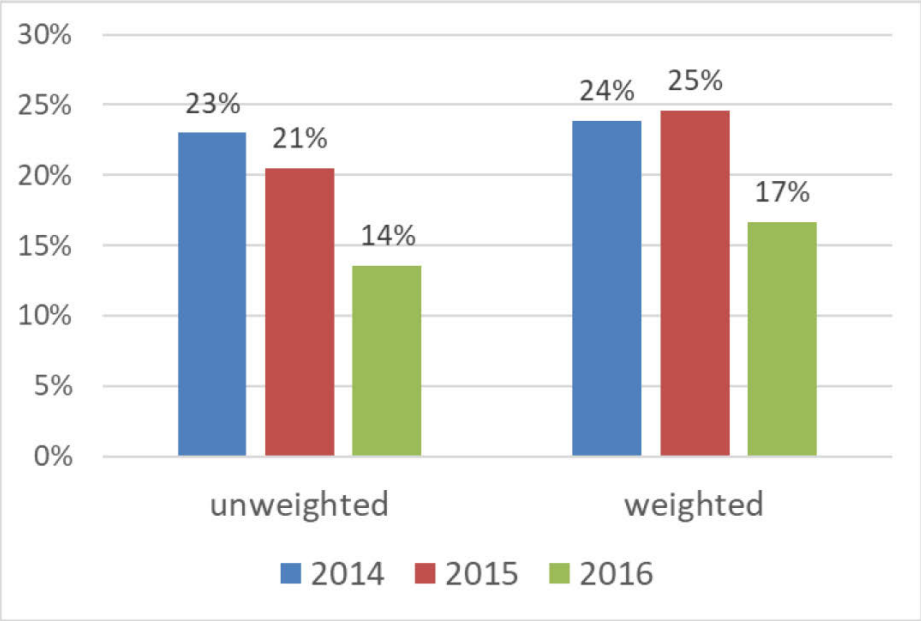


Figure 8: S-New farms response rates for AY 2014–AY 2016 (unweighted vs. weighted).

For various reasons, some of the S-Old farms that took part in the survey in previous years, or in the previous year, dropped out of the sample. This phenomenon is described as **panel attrition**. As a rule, panel attrition is higher between the first two waves than in subsequent waves. Farms that have delivered complete and plausible data for two consecutive years are highly likely to remain on the panel.
For the S-Old farms, the **retention rates (RRTs)** between two consecutive survey waves (*t* and *t-1)* are calculated for each initial recruitment year according to the number of delivered farms for the corresponding accounting years (*j* and *j-1*) as follows:

$$RRT_{t,t-1}^{j,j-1} = \frac{No.of\ delivered\ farms\ \ in\ wave\ t\ for\ \text{AY}\ j}{No.of\ delivered\ farms\ \ in\ wave\ (t-1)\ \text{for}\ AY\ (j-1)} \quad (6)$$

The retention rate shows firstly what percentage of the farms remain on the panel, and thus serves as an indicator of the effort required to maintain the desired sample size. Secondly, it is an indicator of the quality of the estimator of the change in the variables of interest between the years, since both the farms that drop out and the new farms can strongly influence the estimate.[38]
The **cumulative response rate *(RRC)*** can be calculated for each initial recruitment year as the ratio of the number of delivered datasets of the last survey wave *T* for AY *j* to the number of activated farms in AY *j-T+1* (in the initial recruitment):

$$RRC \quad = \frac{No.\ of\ delivered\ farms\ in\ last\ wave\ T\ of\ AY\ j}{No.of\ activated\ farms\ in\ first\ recruitment\ in\ AY\ (j-T+1)} \quad (7)$$

The cumulative response rate *(RRC)* can also be calculated as the product of the unweighted response rate *RRU* in the first year of recruitment (in the year of the first wave) and the retention rates of all subsequent waves:

[38] Basically, an accurate estimate of the change is only possible with panel farms for which data are available for both years.

$$RRC^{j} = RRU_{t=1}^{j} \times \prod_{t=1}^{T} RRT_{t,t-1}^{j,j-1} \quad (8)$$

Table 18 gives an overview of (i) the number of newly activated farms in the previous years, (ii) the number of delivered farms in the initial recruitment year in each case and in the subsequent survey years, and (iii) the calculated retention rates or the cumulative response rate of the S-Old farms. The farms from test years 2010 to 2012 are not shown separately, but are considered to form part of the initial activation of AY 2013, since we are dealing here with a small number of farms from just a few strata.[39] Using the example of the recruitment year AY 2013, Table 10 clearly shows that although the retention rate (RRT) is higher than the response rate (RRU) of the new recruitment (Table 9), it is lower in the second wave (64 %) than in the subsequent wave (93 %). In the fourth wave, the retention rate falls again slightly (81 %), with just 441 (47 %) of the farms of the first wave (919 farms) still taking part in the fourth-wave data survey. In AY 2016, the cumulative response rate (RRC) of the S-Old farms over the entire survey period was 13 %. Owing to the sharp drop in the response rate of the newly recruited farms in AY 2016, a lower cumulative response rate is to be expected the following year.

Table 10: Retention rates and cumulative response rates of the S-Old farms which were activated in the previous years and took part in the AY 2016 survey

	Initial Recruitment Year			Total S-Old
	AY 2013*	AY 2014	AY 2015	
a) Number of activated farms in the year of initial recruitment (with valid addresses)	3 144	5 891	4 039	**13 074**
b13) Data delivered for AY 2013	919	-	-	**919**
b14) Data delivered for AY 2014	586	1348	-	**1 934**
b15) Data delivered for AY 2015	544	837	817	**2 198**
b16) Data delivered for AY 2016	441	719	526	**1 686**
RRT between AY 2014-2013 (b14/b13)**	64 %			
RRT between AY 2015-2014 (b15/b14)	93 %	62 %		
RRT between AY 2016-2015 (b16/b15)	81 %	86 %	64 %	
RRC between AY 2016-2013 (b16/a)	**14 %**	**12 %**	**13 %**	**13 %**

*Including farms from test years

**The RRT between waves 1 and 2, waves 2 and 3, and waves 4 and 3 of the Initial Recruitment year in each case is given in blue, orange and green type, respectively.

2.6 Estimation procedure

In addition to the problem of low willingness to participate described in Chapter 2.5, this chapter deals with the other potential sources of error that can lead to biases in the estimation of the variables of interest.
Two weighting approaches that can be used to correct these biases in the Income Situation sample are introduced. In addition, the formulas for estimating various parameters (mean, quotient, absolute and relative change) as well as the corresponding formulas for estimating variance are presented. The accuracy and reliability (precision) of the means of several selected variables are then determined for the two weighting processes. These results are taken into consideration along with other criteria to select the appropriate weighting method for the Income Situation sample.

[39] This leads to the slight overestimation of the RRC of AY 2013, since the response rate of the first wave is not calculated correctly.

2.6.1 Potential sources of error

The following errors may arise in the Income Situation sample:

1) Sampling error
2) Sampling-frame error
3) Non-response errors
   - Unit non-response
   - Item non-response
4) Measurement error

A **sampling error** arises when we observe the sample percentage of the population rather than using the whole population (complete census). This type of error can arise in any empirical study, depends on the size of the sample, the design of the selection plan, and the chosen selection process, and can be reduced through increased sample size as well as optimal stratification and allocation of the sample. For a random sampling conducted according to the probability sampling theory, the sampling error can be easily controlled and quantified.

A **sampling-frame error** arises when the target population and the list of farms being used as the frame for the drawing of the sample do not match. This error is inevitable in the Income Situation sample, since only the previous year's FSS data are available at the time of the sampling. These data constitute the sampling frame for the survey that will take place one year later. Owing to structural changes, it is possible that certain farms selected for the sample will no longer exist or meet the requirements for participation in a later year, e.g. if their size falls below the defined SO-threshold (as is the case with around 6 % of the target population). This problem is referred to as 'overcoverage', and can be easily solved by subsequently excluding these farms from the sample.[40] On the other hand, new farms arise that were not in the FSS database when the sample was drawn (approx. 4 %). This problem is referred to as 'undercoverage', and must be corrected, as it can lead to bias. It can also happen that some farms which failed to reach the minimum threshold values in the year the sample was drawn owing to their size at the time, actually lie over the threshold in the survey year, due to growth or restructuring. The opposite situation – where the farms no longer belong to the current target population in the survey year, since they now fall short of the minimum threshold – is actually more common. Sampling frame errors are easily corrected.

Non-response errors represent a more serious problem, especially when they are systematic in nature and difficult to identify. A distinction can be made between an item non-response, where a survey participant does not respond to some of the items, and a unit non-response, where all of the information on a sample unit is missing. In principle, only fully completed and plausible datasets are accepted, compensated and evaluated. Incomplete details are accepted only in the part of the questionnaire requesting information about the household income that requires data outside of the financial accounting. Thus, the non-response error concerns only the non-agricultural income and private expenditure of a small percentage of the farms (between 5 and 7 %), and can be estimated or corrected via **imputation**.[41] Unit non-response can occur e.g. owing to a lack of financial accounting, late delivery of data, refusal, or various other reasons (see Chapter 2.5).

[40] Farms which no longer exist, or which have a different legal or operating form, are identified and excluded at the time of recruitment. If, however, the actual farm size is below but near the stipulated SO-threshold in some years, the farms in question are not excluded from the current field of observation.

[41] In standard publications, we dispense with such correction and include a note about the possible bias.

If this is not taken into account in the selection plan (see Chapter 2.3.2), non-response leads to a smaller sample size, and hence to lower estimator precision. In addition, a systematic non-response can lead to the sample no longer being representative in terms of certain characteristics. If these characteristics are correlated with the variables of interest, this leads to a biased estimator, which must be corrected with special methods.

Measurement errors or data entry errors, which arise when wrong responses are given in the questionnaire, are largely identified during the plausibility check and data cleansing (see Chapter 2.4.5 on quality assurance).
The following chapter describes the weighting methodology for correcting potential biases. The calculated weights are used to estimate the means (or the changes over the years).

2.6.2 Weighting methodology

In a complex survey, such as the disproportionately stratified random sample 'income situation', the measurement values from the sample cannot be directly generalised to the target farm population of interest. **Weighting or extrapolation** is absolutely necessary in order to generalise the results from the sample to the target farm population. When the weighting is performed, the data of the delivered farms from the sample are included in the calculation of various parameters (mean, quotient, variance) using a multiplicative factor (weight). Weighting makes it possible to achieve both better precision (reduction in the variance of the estimator) and a correction of the bias arising from various errors (sampling-frame error, non-response error).

Weighting is based on **auxiliary variables** that are correlated with variables of interest and are known for the target population. The auxiliary variables that are relevant for determining the probability of participation can also be used to correct bias.

Usually, farms with different inclusion probabilities are drawn in the sample. With the disproportionate stratified random sample, for instance, for each farm in a stratum, the inverse sampling fractions are used as weights in the estimation of the mean. The aim of this so-called **sampling-design weighting** is to correct the different selection probabilities deliberately created by the selection plan. Without this correction, the estimator of mean would be biased by the sampling error. With a 100 % response rate, i.e. if all farms selected for the sample took part in the survey (no non-response), the reciprocal of the stratum-specific selection probability (or the sampling fraction) could be used as the so-called sampling-design weight.
Weighting can correct not only the sampling bias caused by the different selection probabilities, but also the bias caused by the non-response. In this case, the **total inclusion probability** can be presented as the product of the selection probability and the participation probability. Then, the two random events (a farm being drawn from the sampled population by random sampling, and the selected farm delivering its complete and plausible data to the FADN) are included together in the weighting. Whereas the selection probabilities are predetermined by the selection plan and therefore known, the individual participation probabilities are unknown and must be estimated. Various approaches are available for this (see e.g. Särndal and Lundström, 2005). In the following two sections, we introduce two weighting methods that were considered and tested for their suitability for the Income Situation sample.[42]

Post-stratification (PS)

Post-stratification is a weighting method applied within the data evaluation by stratifying the sample after data collection. The aim of post-stratification is to achieve the greatest possible agreement with the stratification according to sampling design, despite the non-response.

[42] In addition to the post-stratification and calibration weights, the weights calculated with the Propensity Scores model were tested. However, because this methodology was discarded owing to a large number of extremely high weights, it is not presented in this article.

Within the sampling design, in a stratified random sample, the set of all farms of the sampled population $U = \{1, \dots, N\}$ is divided into *L* discrete strata such that $N = \sum_{h=1}^{L} N_h$. A simple random sample is then drawn from each stratum *h.* Of the set of all activated farms with valid contact details in the gross sample $= \{1, \dots, n^S\}$, only some of the farms $R = \{1, \dots, n^R\}$ provide complete, plausible and evaluable datasets (net sample).
If post-stratification is performed with the same stratification variables used for the sampling design, then the distribution of farms in the net sample strata is adjusted to the distribution of the sampled population. The post-stratification weights (PS weights) are thus the same for all farms in the same stratum, and correspond to the number of farms in the sampled population of stratum N_h divided by the number of farms in the net sample *R* of stratum *h* (n_h^R) (delivered, plausible and evaluable datasets). The PS weight for farm *k* in stratum *h* is the same:

$$w_{hk}^{PS} = N_{hk}/n_{hk}^{R} \qquad (9a)$$

Post-stratification takes account of all three causes of bias (sampling-design error, sampling-frame error and non-response error) in a single step. This becomes apparent when the weight is divided into two components:
Sampling-design weight: $d_{hk} = N_{hk}/n_{hk}^{S}$
Probability of participation: $\rho_{hk} = n_{hk}^{R}/n_{hk}^{S}$

The first component is the **sampling-design weight** d_{hk}. This weight corresponds to the reciprocal of the sampling fraction of a stratum, and is used for the non-proportional optimal allocation of the stratified random sample. The number of farms in the sampled population N_{hk} is set in relation to the number of farms randomly selected from the stratum n_{hk}^{S}. When calculuating the post-stratification weights we use the number and the distribution of the farms N_{hk} from the latest available sampling frame, rather than from the sampling frame of the previous year (on which the sampling design was based). This adjusts the sample to all of the sampling-frame changes.

The second component is the **probability of participation**, calculated for each stratum. This is the likelihood that a farm selected from a specific stratum and contacted will take part in the survey. The probability is calculated as the share of delivered farms n_{hk}^{R} out of the activated farms n_{hk}^{S} with valid contact details. This share encompasses all reasons for non-response (refusal, absence of financial accounts, non-contactibility). Using these components, the post-stratification weight can be broken down as follows:

$$w_{hk}^{PS} = \frac{N_{hk}}{n_{hk}^{R}} = \frac{N_{hk}}{n_{hk}^{S}}\frac{n_{hk}^{S}}{n_{hk}^{R}} = \frac{d_{hk}}{\rho_{hk}} = \frac{1}{\pi_{hk}} \qquad (9b)$$

The post-stratification weight therefore corresponds to the sampling-design weight divided by the stratum-specific probability of participation. Since the sampling-design weight corresponds to the reciprocal of the probability of selection, the PS weight can be represented as the reciprocal of the entire inclusion probability π_{hk} (product of the selection- and participation probabilities). Post-stratification is easy to use and convey. An additional advantage of post-stratification is that this methodology does not usually produce any extremely high weights. For the Income Situation sample, the maximum value in AY 2016 was 56 (see Figure 9). The majority of weights range between 8 and 30. Not all relevant information available in the data is used for post-stratification, however. Bias is only corrected on the basis of the distribution of the number of farms in the strata which are defined according to three categorical variables (region/type/size category). The methodology, which can take a greater number of auxiliary variables into account, is presented in the next chapter.

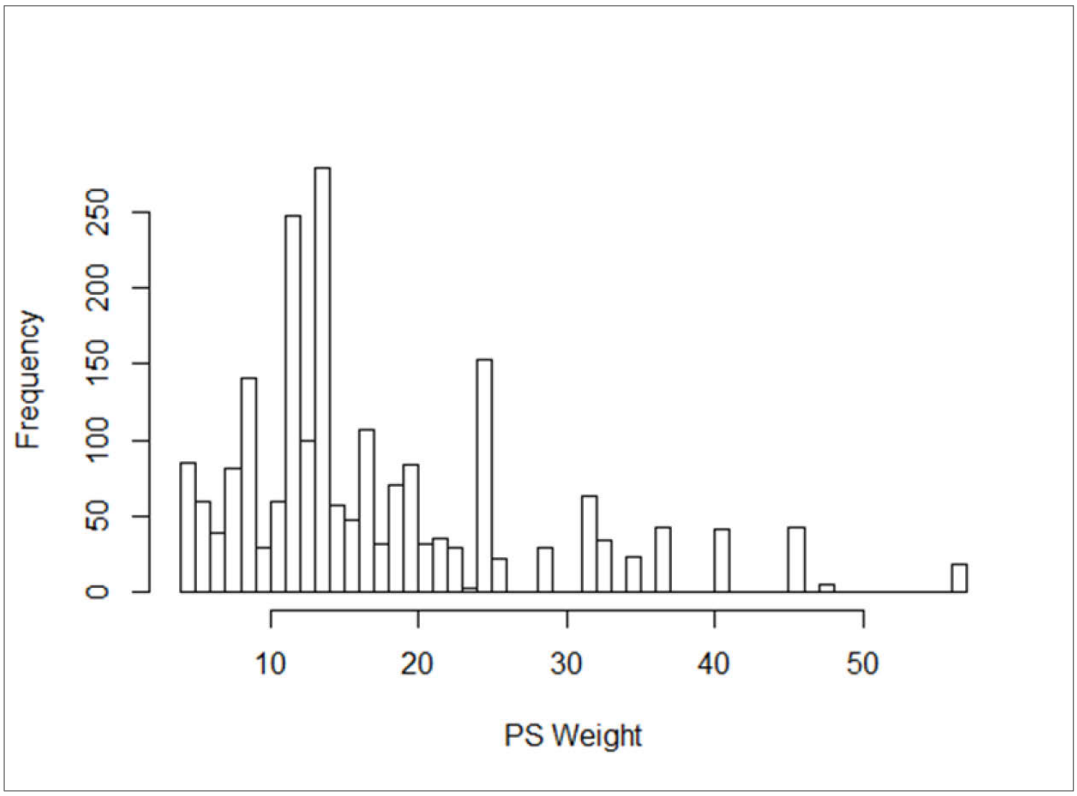


Figure 9: Distribution of post-stratification weights (AY 2016).

Calibration

With the calibration methodology (Deville and Särndal, 1992; Lundström and Särndal, 1999), the information available from the auxiliary variables is used to calculate weights. These weights – being as close as possible to the initial (e.g. sampling-design) weights – should enable an unbiased estimate of certain auxiliary variables. If the auxiliary variables selected for the calibration explain a large part of the non-response and/or are correlated with the variables of interest, then compared to post-stratification, this methodology not only enables greater precision of the estimator (reduction of the confidence interval), but also a better correction of bias.

The calibration weights are determined from a constrained **optimisation problem**. The objective function is formulated via a **distance function** such that the discrepancy between the weight w_k and the initial weight d_k is minimised:

$$\sum_{k \in R} D(w_k; d_k) \quad \underset{w_k}{\rightarrow} \boldsymbol{min} \tag{10}$$

The **constraints** are imposed for the vector $\boldsymbol{x_k} = (\boldsymbol{x_{k1}}, \ldots \boldsymbol{x_{kJ}})$ with J auxiliary variables to ensure that the sample estimates (sum or mean) calculated with the calibration weights w_k are consistent with the known values of the sampled population:

$$\sum_{k \in R} w_k \cdot \boldsymbol{x}_k = \sum_{k \in U} \boldsymbol{x}_k \tag{11}$$

Reciprocals of the stratum-specific selection probabilities N_h/n_h^S according to the selection plan serve as **initial weights** d_k, with N_h being the number of farms in the stratum h, and n_h^S being the number of farms according to the gross selection plan[43].

[43] If sampling-design weight is used as the initial weight, then the calibration aims to determine the total probability of non-response. Since AY2017 we have used the post-stratification weight as an initial weight. As a result, the probability

Calibration is performed at different aggregation levels. This means that the constraint equations are imposed both for Switzerland as a whole and at the level of the individual subgroups: three regions (plain, hill, mountain) and eleven farm types. Calibration variables include not only the number of farms in the strata, but also some additional auxiliary variables that are strongly correlated with the most important variable of interest – agricultural income. This applies to the most important size variables: utilised agricultural area, livestock and standard output[44]. The selection of the calibration variables and the constraint equations of the calibration model vary according to the aggregation level (see Table 11).

Table 11: Selecting the variables for the calibration model

Calibration variables	No. of constraints per level		
	Switzerland as a whole	Region	Farm type
No. of farms	1	3	11
Utilised agricultural area, in ha	1	3	-
Livestock, in livestock units (LU)	1	3	-
Standard output, in CHF	1	3	11

A total of 38 **constraint equations** were imposed in the calibration model for approx. 2100-2200 farms in the sample.[45] In order to avoid extreme weights, we refrain from using any additional calibration variables (e.g. language, age, form of agriculture).[46]

The weights were calibrated with the linear and truncated specification of the distance function (Deville and Särndal, 1992).[47] This variant allows a check of the distribution of the ratio of the weights $g_k = w_k/d_k$. When calculating the weights for AY 2016, the limits of g_k were set to the minimum value of 1 and the maximum value of 10.

Calibration generates individual weights for each farm (or individual inclusion probabilities). Unlike post-stratification, where the same probability of participation is assumed for all farms of the same stratum, calibration allows for greater flexibility. As a result, the calibration weights follow a more continuous distribution than the post-stratification weights (see Figure 10).

of non-response is modelled in the first step during post-stratification, then additionally corrected in a second step during calibration.

[44] We dispense with the calibration at stratum level (region x type x size category), because the number of constraints in this case would be too large, and would produce too many extreme weights. A differentiated use of various auxiliary variables according to farm type (e.g. cattle LU only for dairy farms) was also tested. Compared to calibration with the standard output, however, this variant offered hardly any advantages, and owing to its additional complexity was therefore not implemented.

[45] The desired total sample size of 2300 farms has not been achieved yet in the first years 2015-2016.

[46] No distinction was drawn between S-Old and S-New farms: the total inclusion probability for both farm groups was assumed to be identical, regardless of when initial recruitment was.

[47] The weights were calculated with the statistical software 'R' (R Core Team, 2017) using the 'sampling' package ('calib' function; Tillé and Matei, 2016).

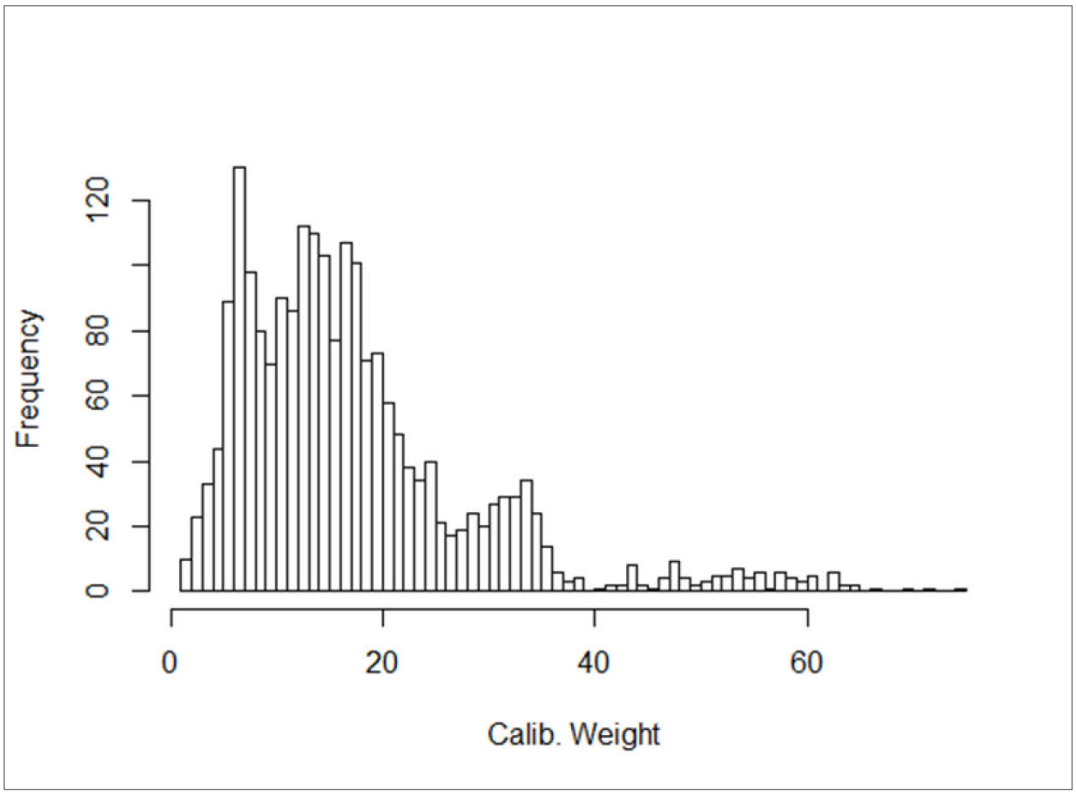


Figure 10: Distribution of the calibration weights (AY 2016).

2.6.3 Parameter estimates (mean, quotient, yearly change rate)

This chapter presents formulas for estimating various parameters (mean, quotient, absolute and relative change).

The focus of the standard publications (e.g. the Basic Report) are the **means** of various variables which are estimated on the basis of the Income Situation sample for various domains (Switzerland as a whole, regions, farm types). Weighting play an important role in the estimation. Using the Horvitz-Thompson estimator (Horvitz and Thompson, 1952), the mean of a variable of interest *y* of the sampled population $\bar{y} = \sum_{k \in U} y_k / N$ is calculated as a weighted mean:

$$\hat{\bar{y}}^{HT} = \frac{\sum_{k \in S} r_k w_k y_k}{\sum_{k \in S} w_k} \qquad (12a)$$

where w_k represents a (total, corrected) weight for farm *k*, which in the case of a missing non-response corresponds to the sampling-design weight of the stratified sample $w_k = d_k = N_{hk}/n_{hk}$ (reciprocal of the sampling fraction of stratum *h*). Where non-response is present, the weight is corrected with the probability of participation ρ_k: $w_k = d_k/\rho_k$. Thus, the (corrected) weight corresponds to the reciprocal of the entire inclusion probability $1/\pi_k$. A dummy variable r_k assumes the value 1 for all farms delivering data and the value 0 for all non-participating farms. Since the sum of the weights is equal to the number of farms of the sampled population *N*, the Horvitz-Thompson estimator for the mean can be represented as follows:

$$\hat{\bar{y}}^{HT} = \frac{1}{N} \sum_{k \in R} \frac{y_k}{\pi_k} \qquad (12b)$$

Depending on the weighting methodology used (see Chapter 2.6.2), either the post-stratification weight or the calibration weight is used when estimating the mean with the formula (12a). For a comparison of the means calculated with the two methods, see Chapter 2.6.5.

If a **ratio** between two variables *y1* und *y2* is to be examined, the quotient of the means is calculated as follows:[48]

$$\hat{\theta}^{HT} = \frac{\widehat{\bar{y}1}^{HT}}{\widehat{\bar{y}2}^{HT}} = \frac{\sum_{k \in S} r_k w_k y_{1k}}{\sum_{k \in S} r_k w_k y_{2k}} \quad (13)$$

For example, the average working income per family labour unit is calculated according to this formula as the estimated mean working income divided by the estimated mean family labour input.
In addition to the means $\hat{\bar{y}}_j$ of the year investigated *j*, the **absolute and relative changes** in the means compared to the previous year *(j-1)* must be estimated:

$$\widehat{\Delta \bar{y}} = \hat{\bar{y}}_j - \hat{\bar{y}}_{j-1} \quad (14)$$

$$\widehat{\Delta \bar{y}}_R = \frac{\left(\hat{\bar{y}}_j - \hat{\bar{y}}_{j-1}\right)}{\hat{\bar{y}}_{j-1}} = \frac{\hat{\bar{y}}_j}{\hat{\bar{y}}_{j-1}} - 1 \quad (15)$$

$\hat{\bar{y}}_j$ and $\hat{\bar{y}}_{j-1}$ are the means estimated from the full sample of the current year *j* and the previous year *(j-1)*, respectively, using the formula (12a).
Because the Income Situation sample is a panel study, the precision of the estimate of change depends on the share of the farms participating in the survey in both years. Since every year a number of farms drop out of the sample whilst others are newly included in it (see Chapter 2.5), the composition of the sample changes from year to year. This can substantially influence the estimation of the annual change for some key figures of interest (Roesch, 2011). To eliminate this undesired effect, the change can be determined on the basis of the constant (overlapping) sample, i.e. considering only those farms that participated in the survey during both of the years under consideration (balanced panel). Qualité and Tillé (2008) show that, given a relatively high overlapping share and a high correlation of the variable of interest between the years, the use of the constant (overlapping) sample is to be preferred over the use of the full sample for the estimation of change (see Chapter 2.6.4).

2.6.4 Estimating variance

The variance of the estimated parameters plays an important role in assessing the precision of a sample-based estimator. In addition, the square root of the variance (i.e. the standard error) can be used to determine the confidence interval that encompasses the true value with a given probability and assuming a normal distribution. The 95% confidence interval is calculated as follows:

$$CI(\hat{\bar{y}}) = \hat{\bar{y}} \pm t \cdot \widehat{se}(\hat{\bar{y}}) = \hat{\bar{y}} \pm 1.96\sqrt{\widehat{Var}(\hat{\bar{y}})} \quad (16)$$

If the confidence interval is to be calculated with just a few observations for a subgroup only (e.g. for a farm type and region) rather than for the entire sample, a corresponding multiplier *t* of a Student's t-distribution is then used in formula (16).[49] With a sufficient number of observations (over 20), the t-multiplier converges to the value 1.96.
Besides the confidence interval, the **coefficient of variation** is calculated in order to assess the precision of the differently scaled key figures. The coefficient of variation is dimensionsless, and is calculated as follows:

$$\widehat{CV}(\hat{\bar{y}}) = \frac{\sqrt{\widehat{Var}(\hat{\bar{y}})}}{\hat{\bar{y}}} \cdot 100 \quad (17)$$

[48] For a discussion of the estimation of the ratio variables via ratio of means or mean of ratios, see e.g. Rao (2002).

[49] The degrees of freedom for the stratified sample are corrected by the number of strata. Alternatively, the method of Satterthwaite (1946) could be used for the correct calculation of the degrees of freedom, but this has hardly any impact on the results.

In this section, the formulas and the approach for **calculating the variance** of the estimator introduced in Chapter 2.6.3 are presented.

The **variance of the post-stratification estimator of the mean** is calculated according to the following formula (Horvitz and Thompson, 1952):

$$\widehat{Var}(\hat{\bar{y}}^{PS}) = \frac{1}{N^2}\sum_{k\in S}\sum_{l\in S} r_k r_l \frac{(\pi_{kl}-\pi_k\pi_l)}{\pi_{kl}}\frac{y_k}{\pi_k}\frac{y_l}{\pi_l} \quad (18)$$

where π_k represents the overall inclusion probability: $\pi_k = \rho_k/d_k = Pr_k\rho_k$ with Pr_k, the selection probability according to selection plan ($Pr_k = 1/d_k$) and ρ_k is the probability of participation. In addition, π_{kl} is the pairwise inclusion probability that both farms *k* and *l* are contained in the net sample, and r_k is a dummy variable assuming the value 1 in the case of participation and the value 0 in the case of non-participation. With the post-stratification estimator, the overall inclusion probability is calculated as follows: $\pi_k = n_{hk}^R/N_{hk}$.
The variance of the calibration estimator of the mean is estimated according to the residual method of Deville und Särndal (1992) as follows:

$$\widehat{Var}(\hat{\bar{y}}^{Calib}) = \frac{1}{N^2}\sum_{k\in S}\sum_{l\in S} r_k r_l \frac{(\pi_{kl}-\pi_k\pi_l)}{\pi_{kl}}(w_k e_k)(w_l e_l) \quad (19)$$

where e_k and e_l represent the residuals of k and l resulting from linear regression on the variables used for the calibration (10)-(11) $e_k = y_k - \mathbf{x}'_k\hat{B}_{ws}$.

The **variance of the quotient** of two Horvitz-Thompson means is calculated according to Taylor-series linearisation (Woodruff, 1971) as follows:

$$\widehat{Var}(\hat{\theta}^{HT}) = \sum_{k\in S}\sum_{l\in S} r_k r_l \frac{(\pi_{kl}-\pi_k\pi_l)}{\pi_{kl}}\frac{\hat{z}_k}{\pi_k}\frac{\hat{z}_l}{\pi_l} \quad (20)$$

where $\hat{z}_k = \frac{y_{1k}-\hat{\theta}y_{2k}}{\sum_{k\in S}\tilde{w}_k y_{2k}}$ and $\hat{\theta}$ is defined as an estimator for the quotient according to the formula (13). If the means are calculated with the calibration weights, then the residuals e_k are used instead of $\hat{z}_k$ in the formula (20): $e_k = \hat{z}_k - \mathbf{x}'_k\hat{B}_{ws}$.
The variance of the absolute change is calculated as follows:

$$\widehat{Var}(\widehat{\Delta y}) = \widehat{Var}(\hat{\bar{y}}_j) + \widehat{Var}(\hat{\bar{y}}_{j-1}) - 2\widehat{Cov}(\hat{\bar{y}}_j, \hat{\bar{y}}_{j-1}) \quad (21)$$

with the variance $\widehat{Var}(\hat{\bar{y}}_j)$ and $\widehat{Var}(\hat{\bar{y}}_{j-1})$ being calculated according to formulas (18) and (19), respectively.

The covariance $\widehat{Cov}(\hat{\bar{y}}_j, \hat{\bar{y}}_{j-1})$ is calculated as follows (Qualité and Tillé, 2008; Berger and Priam, 2016):

$$\widehat{Cov}(\hat{\bar{y}}_j, \hat{\bar{y}}_{j-1}) = \frac{1}{n_C-1}\sum_{s_C}\left(w_{(j-1)k}y_{(j-1)k} - \bar{y}_{(j-1)C}\right)\left(w_j y_{jk} - \bar{y}_{jC}\right) \quad (22)$$

with

$$\bar{y}_{(j-1)C} = \frac{1}{n_C}\sum_{s_C} w_{(j-1)k}y_{(j-1)k}$$

$$\bar{y}_{jC} = \frac{1}{n_C}\sum_{s_C} w_{jk}y_{jk}.$$

S_C is a constant (overlapping) part of the sample, consisting of the farms which delivered their data in both years j and j-1, and n_C is the number of said farms.

Alternatively, the ratio-type estimator of covariance can be used:

$$\widehat{Cov}(\hat{\bar{y}}_j, \hat{\bar{y}}_{j-1}) = Cor_{j(j-1)C}\sqrt{\widehat{Var}(\hat{\bar{y}}_j)\cdot\widehat{Var}(\hat{\bar{y}}_{j-1})}. \quad (23)$$

An important property of this estimator for covariance is that it cannot assume any negative values (Qualité, 2009). Here, the correlation coefficient $Cor_{j(j-1)C}$ is calculated as follows, based on the constant sample data (Qualité, 2009, p. 85):

$$Cor_{j(j-1)C} = \frac{\Sigma_{s_C}\left(1-1/w_{cjk}\right)\left(w_{c(j-1)k}y_{(j-1)k}-\Sigma_{s_C}a_{ck}w_{c(j-1)k}y_{(j-1)k}\right)\left(w_{cjk}y_{jk}-\Sigma_{s_C}a_{ck}w_{cjk}y_{jk}\right)}{\left[\Sigma_{s_C}\left(1-1/w_{C(j-1)k}\right)\left(w_{C(j-1)k}y_{(j-1)k}-\bar{y}_{(j-1)C}\right)^2\Sigma_{s_C}\left(1-1/w_{cjk}\right)\left(w_{Cjk}y_{jk}-\bar{y}_{jC}\right)^2\right]^{1/2}} \quad (24)$$

Note that only the observations of the constant sample are used in the correlation coefficient formula (24), whilst the variances of the two full samples are used in the covariance formula (23).

The **variance of the relative change** is calculated as follows (Qualité and Tillé, 2008):

$$\widehat{Var}\left(\hat{\Delta}_R\right) = \frac{1}{\hat{\bar{y}}_{j-1}^{\,2}}\left[\widehat{Var}\left(\hat{\bar{y}}_j\right) + \left(\frac{\hat{\bar{y}}_j}{\hat{\bar{y}}_{j-1}}\right)^2 \widehat{Var}\left(\hat{\bar{y}}_{j-1}\right) - 2\frac{\hat{\bar{y}}_j}{\hat{\bar{y}}_{j-1}}\widehat{Cov}\left(\hat{\bar{y}}_j, \hat{\bar{y}}_{j-1}\right)\right] \quad (25)$$

The variance of the absolute or relative change is used to test for statistically reliable statements on the rise or fall in agricultural income on the level of Switzerland as a whole and on that of the three regions. If, according to the calculated t-statistic, the p-value lies above the 5% significance level, then the estimated change must be interpreted with caution.

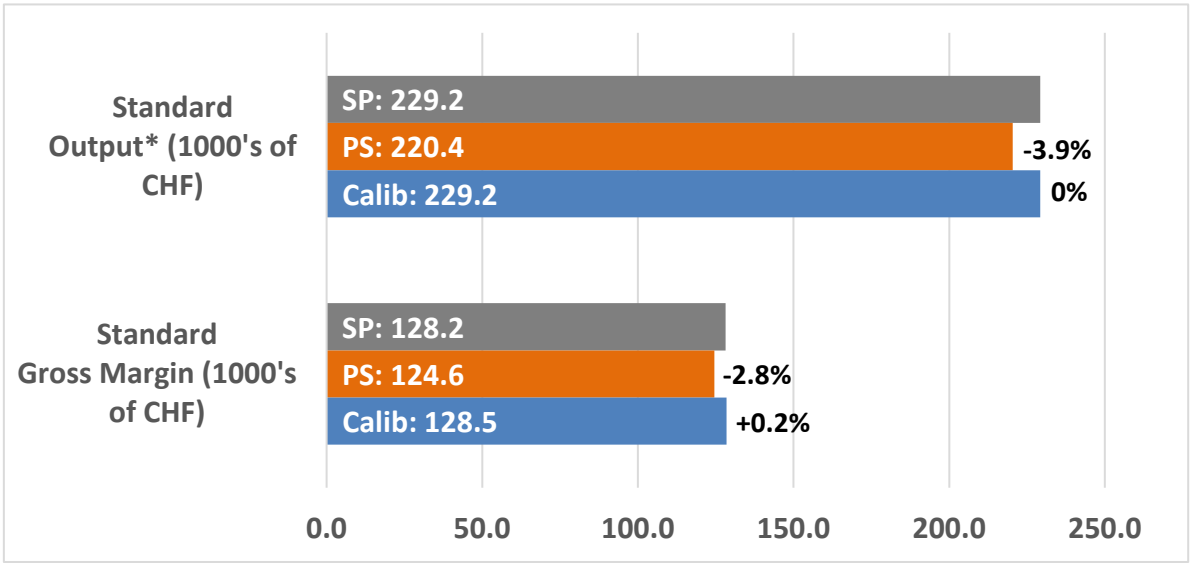
2.6.5 Accuracy and reliability of the estimate – comparison of methods

A **representative sample** enables to make reliable statistical inferences about the target population. Stratification and the corresponding weighting ensure that the sample drawn has a similar structure to the target population in terms of certain characteristics.

The post-stratification weighting ensures the representativeness of the sample in terms of the allocation of the farms to the plain, hill and mountain regions, as well as in terms of allocation according to farm type and economic size. Calibration additionally ensures representativeness with regard to the entire agricultural area and livestock numbers, both for Switzerland as a whole as well as for the plain, hill and mountain regions. Moreover, it ensures that the average economic size of the farms (measured according to standard output) on different aggregation levels (Switzerland as a whole, regions, farm types) corresponds to the circumstances of the target population. The stratification and calibration characteristics were chosen so as to estimate the most important variables of interest as precisely as possible. This chapter analyses the reliability of the estimation results from the sample and compares the properties of the weighting methods described above.

Mean squared error is a quality criterion commonly used to assess the accuracy of the sample estimate (cf. e.g. Cochran, 1977, p. 15 or Lohr, 1999, p. 28). A low mean squared error is achieved when both the systematic error (bias) and the variance of the estimator are small – in other words, the expected mean estimator is as close as possible to the true value, and has a low variance.

In the first step, we compare the means of some structural variables estimated on the basis of the AY 2016 sample with the known true values of the sampled population (selected variables from the 2016 FSS data) at the level of Switzerland as a whole. This allows a direct assessment of the accuracy of the alternative weighting procedures. Figures 11 to 13 show the true means of several selected structural variables in the sampled population (SP), as well as the estimated means and the percentage deviations from the true mean of the two weighting methods – post-stratification (PS) and calibration (Calib).

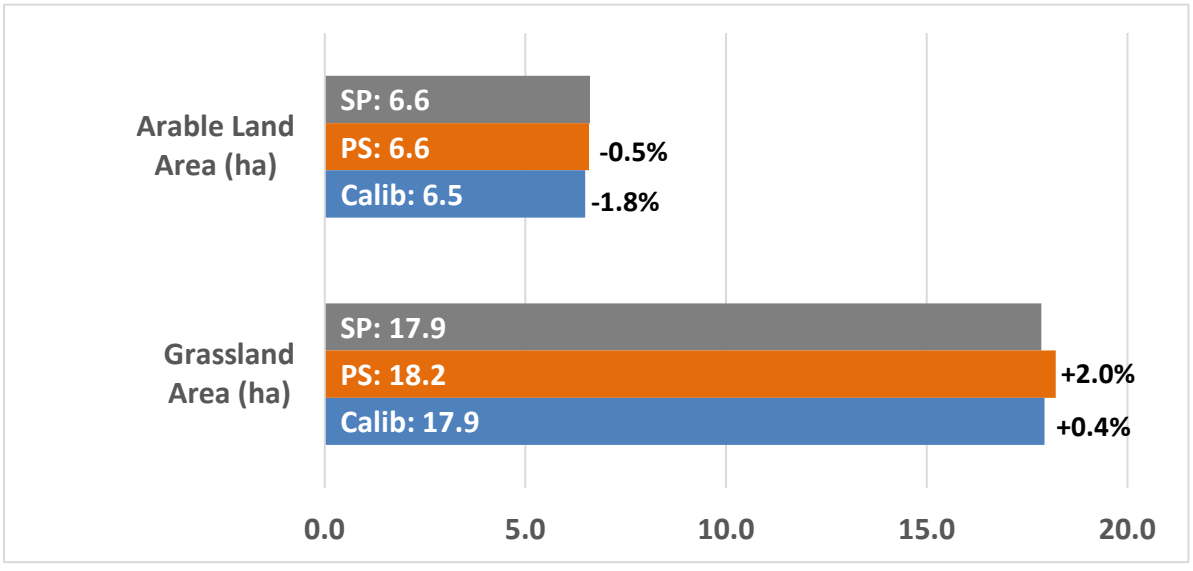


PS = estimation with post-stratification weights; Calib = estimation with calibrated weights; SP = sampled population.

Figure 11: Comparison of the estimated means (SO and SGM) with the true means of the sampled population at the level of Switzerland as a whole

Because the standard ouput was used as a calibration variable (owing to its high correlation with the main variable of interest 'agricultural income'), the calibration estimator has no bias (Figure 11).

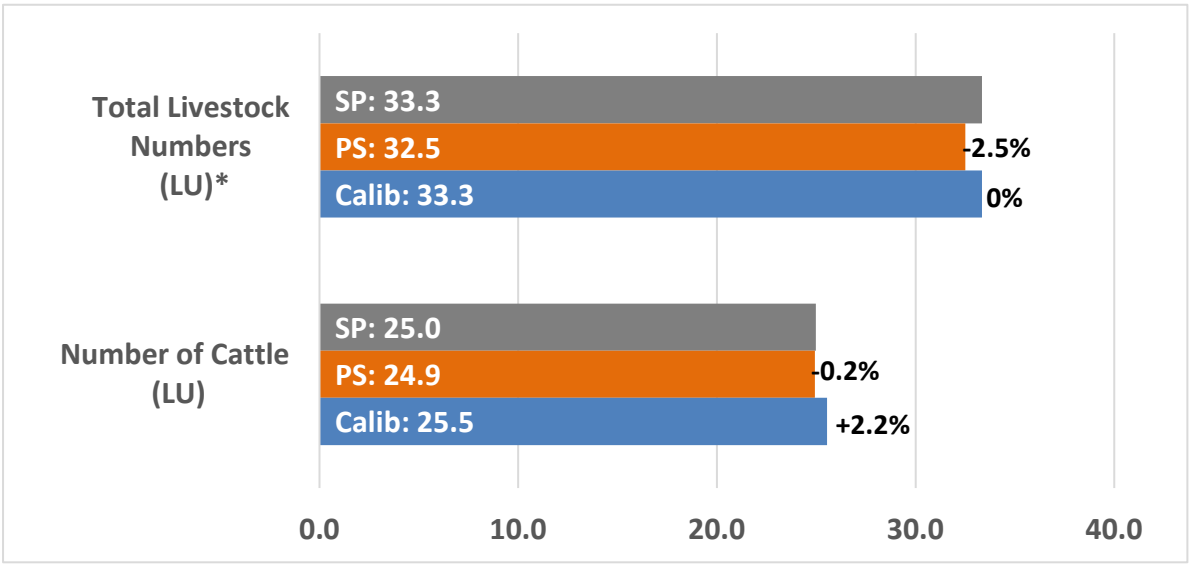
By contrast, the mean estimated with post-stratification has a downward bias of 3.9%. With the standard gross margin – another important economic variable – the deviation for the calibration estimator of +0.2% is also smaller than with post-stratification (-2.8%).



PS =estimation with post-stratification weights; Calib = estimation with calibrated weights; SP = sampled population.

Figure 12: Comparison of the estimated means (arable and grassland area) with the true means of the sampled population at the level of Switzerland as a whole

For the selected **areas** (arable area, grassland area), the bias for both estimates is relatively low, with calibration weights providing somewhat more accurate results for grassland area and the post-stratification estimator calculating arable area more accurately (Figure 12).



PS =estimation with post-stratification weights; Calib = estimation with calibrated weights; SP = sampled population.

Figure 13: Comparison of the estimated means (total livestock numbers and number of cattle) with the true means of the sampled population at the level of Switzerland as a whole

The estimated means for the **livestock numbers** (total and cattle) are also close to the true values known from the FSS database (Figure 13). Although calibration was applied only to total livestock numbers, the estimate is also fairly accurate for cattle (+2.2 %), with the post-stratification estimator being more accurate here (-0.2 %).

Looking at the bias for the lower aggregation levels (region, farm type), which for statistical reasons is generally greater, we can see that the errors with the calibration estimator are usually smaller than with post-stratification. For the sake of greater clarity, we have refrained from depicting the post-stratification estimate in the following representation of the accuracy for individual regions and farm types.

Table 12 shows the relative **deviations of the calibration estimator** of several selected structural variables on various aggregation levels (Switzerland as a whole, region, farm type). These results show that the estimate is fairly accurate at the level of Switzerland as a whole and at regional level in terms of the most important structural variables. Although the linguistic distribution of the sampling population was considered during the sampling, the farms from German-speaking Switzerland are over-represented in the net sample. This is true both for Switzerland as a whole (5% more than in the sampled population) and for lower aggregation levels (distribution according to language has a particularly strong bias in the case of Special Crops).

Age distribution – measured by the percentage of farm managers under 40 years of age – is also biased on the level of Switzerland as a whole, with younger farmers being under-represented in the sample, leading to greater errors at the lower aggregation levels. Generally, deviations for all variables are greater at lower aggregation levels, since the estimate is performed with a smaller number of farms. Greater biases in the estimates for farm types 1512 and 1531 can be attributed to their low response rates. In addition, the estimates for areas or livestock categories which are insignificantly small for certain farm types (e.g. open arable land or grassland in the case of farms specialising in pigs and poultry) are flawed.

Table 12: Accuracy of the estimate with calibration weights at different aggregation levels (percentage deviation from true mean of sampled population)

Aggregation level	SGM (1000s of CHF)	UAA* (ha)	Arable land (ha)	Grass-land (ha)	Livestock numbers (LU)*	Livestock numbers, cattle (LU)	Percentage of German-speaking farms	Percentage of farm managers < 40 years Old
Switzerland as a whole	0.2%	0.2%	-1.7%	0.2%	0.1%	2.2%	5.0%	-4.6%
Plain region	0.9%	0.1%	-3.1%	2.3%	-0.1%	5.4%	3.0%	-11.7%
Hill region	-0.2%	0.2%	2.3%	-0.6%	0.1%	0.3%	5.2%	9.3%
Mountain region	-1.4%	0.0%	25.0%	0.0%	0.0%	-0.5%	7.4%	-7.9%
Arable crops 1511	-0.8%	-5.3%	-5.5%	-3.4%	10.0%	16.8%	-11.6%	-10.1%
Special crops 1512	1.9%	6.1%	9.7%	6.9%	3.8%	4.1%	35.2%	-13.0%
Dairy cows 1521	-0.7%	1.1%	-1.8%	0.5%	-1.2%	-1.8%	1.5%	-6.3%
Suckler cows 1522	-3.0%	-11.1%	25.4%	-13.2%	-5.0%	-3.8%	11.7%	21.9%
Cattle, mixed 1523	-2.4%	1.6%	-7.8%	1.8%	-3.1%	-2.8%	8.5%	-18.1%
Horses/sheep/goats 1531	-18.3%	-1.9%	3.3%	-2.2%	-6.4%	8.6%	27.8%	29.0%
Pigs/poultry 1541	2.3%	30.3%	20.0%	30.9%	-0.3%	33.4%	5.2%	21.6%
Comb. dairy/arable crops 1551	-2.7%	-3.6%	-4.7%	-2.2%	-1.1%	-0.9%	1.7%	-21.0%
Combined suckler cows 1552	0.4%	-5.8%	-18.4%	2.7%	5.3%	7.0%	19.5%	-24.9%
Combined pigs/poultry	0.9%	0.1%	-7.3%	5.3%	0.0%	6.6%	0.4%	7.6%
Combined other 1554	4.4%	1.9%	6.7%	-0.4%	6.0%	10.8%	0.9%	-9.2%

* Variables used for calibration at the level of Switzerland as a whole, and at plain, hill and mountain region level

For the the economic indicators, which include the actual **variables of interest** such as agricultural income, there is no way to determine the deviation from the 'true' means, since this information is only known in the case of farms supplying data. For this reason, the variance of the estimator, i.e. the measure of spread of the estimated means around the expected value, is used in the second step to evaluate the **precision** of the estimator. Based on the square root of the variance (the standard error), and assuming a normal distribution, we can then determine the confidence intervals in which the true value is to be found with a given probability (see Chapter 2.6.4). Table 13 presents the estimated means for agricultural income, together with the absolute and relative confidence-interval ranges at different aggregation levels.

The results presented show that with post-stratification, agricultural income for the whole of Switzerland can be calculated with an expected accuracy of +/- 3.5 % around the mean. Calibration could improve the accuracy of the estimate for the most important variable to +/- 2.7 %. At regional level, the possible deviation from the expected mean lies in the range of +/- 4 % to 5 % for the calibration estimator and +/- 5 to 8 % for the post-stratification estimator, with the estimate for the plain region being somewhat more accurate. The 'farm type' level can produce larger errors, and results should be interpreted with caution for farm types 1512 (special crops), 1522 (suckler cows), 1531 (horses/sheep/goats) and 1552 (combined suckler cows) in particular.

Table 13: Comparison of the estimated means of agricultural income and the 95 % confidence interval, calculated with post-stratification and calibration weights

Aggregation level	Mean (CHF)		Conf. interval (+/- CHF)		Conf. interval (+/-%)	
	PS	Calib	PS	Calib	PS	Calib
Switzerland as a whole	63 813	64 275	2 230	1 739	3.5%	2.7%
Plain region	77 284	79 923	3 542	3 168	4.6%	4.0%
Hill region	55 811	54 684	3 680	2 578	6.6%	4.7%
Mountain region	52 291	51 155	4 100	2 689	7.8%	5.3%
Arable crops 1511	70 108	71 247	9 104	6 962	13.0%	9.8%
Special crops 1512	94 539	102 097	12 085	11 617	12.8%	11.4%
Dairy cows 1521	55 776	55 240	3 304	2 606	5.9%	4.7%
Suckler cows 1522	51 990	47 994	7 103	5 154	13.7%	10.7%
Cattle, mixed 1523	48 868	48 829	6 374	3 998	13.0%	8.2%
Horses/sheep/goats 1531	51 442	47 508	7 669	5 131	14.9%	10.8%
Pigs/poultry 1541	84 990	88 006	9 435	7 675	11.1%	8.7%
Comb. dairy/arable 1551	65 413	66 213	7 135	5 296	10.9%	8.0%
Comb. suckler cows 1552	59 308	57 098	12 681	10 215	21.4%	17.9%
Combined pigs/poultry 1553	78 402	78 337	5 521	4 327	7.0%	5.5%
Combined other 1554	61 067	62 114	6 006	4 643	9.8%	7.5%

2.6.6 Evaluation and selection of the weighting methods

Criteria for the evaluation

The weighting method was selected on the basis of the following criteria:

- *Practicability*: This criterion encompasses the complexity of the methodology and the scope of the information and variables necessary for the calculation of the weights.
- *Communicability:* 'Communicability' refers to the ease with which this methodology can be conveyed to the wider public and understood by the same.
- *Extreme weights*: 'Extreme weights' means that a very large number of farms from the sampled population are represented by the data of one farm from the sample. This can happen if the farm was drawn from a stratum with a low sampling fraction (large, relatively homogeneous strata) whilst at the same time representing a group of farms with a low probability of participation. Extremely large weights lead to the farms in question potentially having a very strong influence on the aggregated results. This can lead to undesired effects, especially when analysing subgroups (at the 'type' or 'type and region' level).
- *Accuracy*: The actual aim of the weighting is to correct the bias and obtain as accurate an estimate as possible. Using several known characteristics of the sampled population that are correlated both with the probability of participation and with the variables of interest, the weights are set so as to minimise the systematic errors. The true means of the sampled population are compared with the estimated means in order to evaluate accuracy.
- *Precision:* The estimator with the lowest variance (i.e. with the narrowest confidence interval) is preferable.

Evaluation and selection of the weighting methodology

The advantage of the **post-stratification** methodology is the ease with which it allows weights to be calculated. The method is easy to understand and convey to others, and leads to a balanced distribution of the weights. Extremely large or small values do not occur for the weights. For these reasons, post-stratification is often used in social research. Post-stratification weights were used in the earlier FADN reference farm sample by adjusting the distribution of the data-supplying farms to the distribution of the sampled population for all strata. This method is also often used in the international context. All of the EU's national farm accountancy data networks (the French INLB, Britain's FADN) use post-stratification for estimating the means. Stratification is usually carried out according to farm type and farm size, and sometimes additionally according to region.

When post-stratification weights are used, the structural features of the sampled population are fairly well represented by the sample. For most variables, the deviations from the true means lie in the range of +/- 2 % to +/- 4 % at the level of Switzerland as a whole and in the range of +/- 3 % to +/- 6 % for the regions. Larger biases occur on the 'farm type' level which cannot be eliminated completely by the weighting.

Using the **calibration weights** makes it easier to correct the bias of the structural variables. The model was formulated such that no deviations from the true mean occur for the most important auxiliarly variables, even at the 'farm type' level. With few exceptions, the deviations in the control variables, which were not taken into account during calibration, also remain small. Owing to its high precision, this estimator also has an advantage when evaluating the variables of interest. The drawbacks of this estimator lie in its higher extreme weights, its somewhat more complicated implementation, and in the difficulty of explaining the estimating procedure to others.

Because the advantages of increased precision and accuracy prevailed in our assessment, we decided to use the calibration weighting. The means and the ratio variables of AY 2016 were calculated with the calibration weights according to formulas (12a) and (13).

The **absolute** and **relative changes** to the previous year were calculated with the calibrated means of the years in question according to formulas (14) and (15). Since this estimate is influenced by the farm's annual entry into and exit from the sample, the annual changes are additionally compared with the means of the constant sample (or balanced panel). Here, only those farms participating in the sample both years were taken into consideration. To this end, special post-stratification weights were calculated. Owing to the smaller number of farms in the balanced panel (in AY 2016, the constant sample contained 1546 comparable farms at the level of Switzerland as a whole), a simplified allocation to the different strata according to Type x Region (with no additional allocation according to farm size) was applied.

2.7 Reporting survey results of the Income Situation sample

Because the financial accounting on which the survey is based is used for tax purposes and is accordingly optimised, certain key figures need to be interpreted differently. This is especially true for depreciations, repairs and own work capitalised, as well as deliveries in kind from the farm to the private household. Consideration of the cash flow offers the advantage that tax optimisations have less influence. Cash flow was only included in the standard publications from AY 2016 onwards. However, the new sampling concept, with random sampling and the linking of the FSS database, also provides new publishing opportunities on the subjects of accuracy of estimate and assessment of representativeness.
With the changeover to the new survey concept, **group farming businesses** (unregistered partnerships) are now included in the analysis. The profit and loss account jointly evaluates the sole proprietorships and group farming businesses, up to and including agricultural income and working income.

The representation of the key figures for cash flow as well as total income and private consumption includes just the sole proprietorships without the group farming businesses. The reason for this is that the private sphere of group farming businesses cannot be surveyed by all partners.
From AY 2015 onwards, the analysis of the economic situation of the Swiss agricultural sector was based on the random Income Situation sample. The changeover in reporting therefore also took place then.

2.7.1 Communication to the public

The **media release** is directed at the non-specialist press, and addresses the following issues for both Switzerland as a whole and for the three regions 'plain', 'hill' and 'mountain', bearing in mind the relevance[50] and significance of the changes:

- How have agricultural income, working income and overall income changed to the previous year?
- What are the main reasons for the change (e.g. expenditure, income, price developments)?
- How has the cash-flow changed?

More-detailed information, graphics and tables can be found in the Main Report, which is published simultaneously.

In 2017, the FADN's media releases were for the first time published simultaneously with the results of the Economic Accounts for Agriculture (EAA) of the FSO, in order to counteract any possible confusion. Since then, a brief description of the most important differences between the EAA and FADN methods has supplemented the media release. A detailed report by Murbach and Schmid (2017) on the differences in methods is also available.

The **Main Report** is geared to all stakeholders and members of the public interested in the agricultural sector (e.g. agricultural advisors, accountants, farmers, policy makers, media professionals) and is available as an additional source of information.

For Switzerland as a whole as well as for the three regions of plain, hill and mountain and for the farm types, key figures are published on the following topics:

- Farm structure
- Farm operating income and expenses
- Agricultural income
- Working income and comparison with reference salary of the second and third sectors
- Total income
- Cash-flow statement
- Dispersion and top/bottom quartile[51] of working income.

The **Basic Report** is an electronic tabular report in Excel and PDF formats which since AY 2015 has only been available in digital form. The target audience of the Basic Report consists of stakeholders in the agricultural sector.

[50] In this context, 'relevance' is taken to mean importance for the trend in agricultural income. If, for example, a key figure has changed significantly, but is only responsible for 0.1% of agricultural yield, it is not mentioned in the report owing to its low relevance.

[51] A quartile is taken to mean one of three points dividing a range of data (which has been sorted by size) into four parts. The top or third quartile demarcates the most successful 25% of farms, the bottom or first quartile the least successful 25%. A boxplot graph, the values of the first and third quartiles, the absolute dispersion range between the first and third quartiles and the relative dispersion as a quotient of the third and first quartiles are used to assess the dispersion.

Presentation of the survey results is adjusted to the usual presentation in financial accounting, so that comparison with individual-farm accounting data is as simple as possible for interested parties. Additionally, the harmonised calculation of agricultural and working income, which differs slightly from the financial accounting and is more suitable for purposes of comparison, is presented in the second part of the table results.
The content of the report consists of the data from the current accounting year and the two previous years, as well as the mean of these three years for the following subject areas:

- Work input
- Land resources
- Livestock numbers
- Closing balance sheet
- Profit-and-loss account
- Agricultural income and working income
- Production factors and performance
- Total income and private consumption (excluding group farming businesses)
- Cash-flow statement (excluding group farming businesses)

The published aggregation levels are Switzerland as a whole, the plain, hill and mountain regions, and all farm types. To the extent permitted by sample size in the strata, results for more-detailed stratifications are listed.

The Income Situation sample offers a relatively large number of key figures whose **accuracy of estimation** can vary greatly depending on aggregation level. Accuracy can be measured by the coefficient of variation, which is expressed as a percentage and defined as the ratio between the standard error (cf. Chapter 2.6.4) and the estimated mean. In the Basic Report, the estimated values from AY 2016 onwards are classified according to their quality by means of the coefficient of variation:

- a: Very good – coefficient of variation < 1 %
- b: Good – coefficient of variation ≥ 1 % and < 2 %
- c: Average – coefficient of variation ≥ 2 % and < 5 %
- d: Adequate – coefficient of variation ≥ 5 % and < 10 %
- e: Poor – coefficient of variation ≥ 10 % and < 30 %
- f: Unreliable – coefficient of variation ≥ 30 %.

These quality categories are displayed in the data table right next to the absolute value.

The Basic Report contains a glossary with definitions and explanations of important key figures, provided that the definition cannot be inferred from the representation in the table (e.g. in the profit-and-loss account). The glossary is constantly updated.

2.7.2 Individual-farm report for participating farms

Besides monetary compensation, individual-farm reports offer data-supplying farms of the Income Situation sample additional added value by enabling farm managers to compare their farms with structurally similar farms (mean, top/bottom 25 %) and with the average Swiss farm, and to analyse their own farm's trend over the last three years at most.
The individual-farm report contains important key figures from the annual financial statement, as well as the most important key structural figures. Some of the data is illustrated with diagrams.

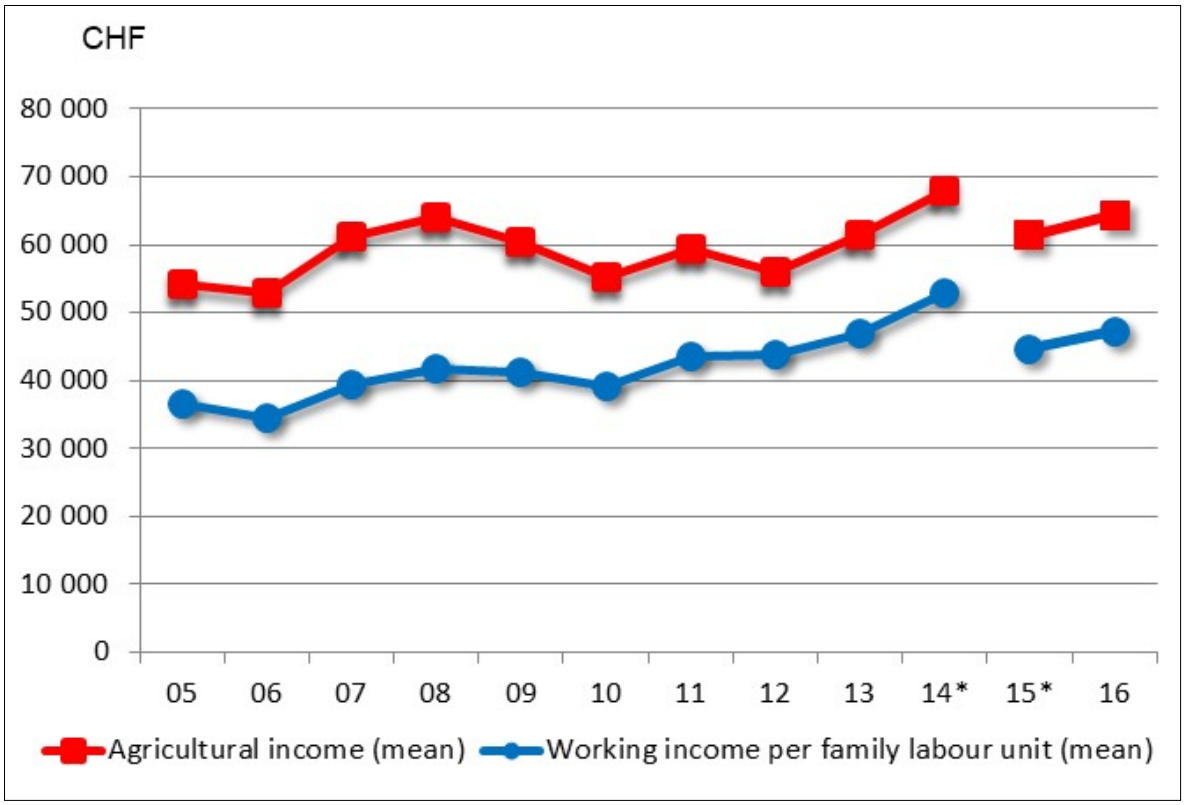
The following topics are covered:

- Structural data (workforce, livestock numbers, utilised agricultural area)
- Balance sheet
- Profit-and-loss account according to the 'Agricultural SME Chart of Accounts'
- Agricultural income and working income
- Balance-sheet and profit-and-loss account performance indicators
- Cash-flow statement
- Total income and private consumption (excluding group farming businesses)

Once the FADN has completed the data preparations, the individual-farm report is posted as a four-page colour document together with the payment cheque to all farm managers who supplied data. From AY 2016 the farm manager has been able to decide whether his data-supplying accounting offices should receive an electronic copy of his farm report.
For a better understanding of the manner of presentation and the key figures illustrated, a documentary report with explanations is published online (Agroscope 2017b).

2.8 Influence of reorganisation of survey system on income estimation

With the introduction of the new FADN system, the income estimate from the 2015 accounting year onwards has been based on the data of the Income Situation sample. Owing to the many methodological changes, there is a **break in the time series** between accounting years 2014 and 2015 (see Figure 14).



** Break in the time series between 2014 and 2015 owing to the changeover to a new survey design with the introduction of a random sample.*

Figure 14: Trend in agricultural income per farm and working income per family labour unit (2005-2016).

For the 2014 accounting year, it was possible to compare the results estimated with the old and new systems.[52]

The **break in the time series** was quantified by comparing the estimated means of the Income Situation sample with the previous Reference Farm sample. Since the data of the Income Situation sample were not yet available in full in the first survey year AY 2014 (*inter alia* owing to a number of missing or under-represented strata), two alternative estimates were made. In one, a weighted mean of the full sample (1934 farms) was calculated for AY 2014. For the alternative estimate, the relative change in the so-called balanced panel between 2015 and 2014 was first of all calculated, using only the data of of those farms which participated in the survey both years. Based on this relative change, the mean for AY 2014 was estimated backwards from the mean of the full sample for AY 2015 (almost 2200 farms). These alternative calculations yield a range for the estimated value of the respective indicator in AY 2014.[53]

After the break was quantified, it was noted that the two most important indicators – agricultural income and working income per annual family labour unit – were lower in the Income Situation sample than in the Reference Farm Sample in accounting year 2014. Below, we limit ourselves to presenting these most important key figures, and show the difference for Switzerland as a whole and for the three individual regions.[54]

There are numerous **causes for this level shift**, though a more precise quantification of the individual effects is only possible to a limited extent. The effects of the changeover to financial accounting were calculated using a partial sample of the reference farms which made the financial accounting results available in addition to the results according to the 'farm accounts'. With these farms, a direct comparison was possible. These results were generalised for the entire Reference Farm sample for AY 2014. The effect of the new income calculation methodology (in particular, the new accounting rules for employer contributions) was also quantified on the basis of the Reference Farms sample. Table 14 summarises the most important effects of these two methodological changes.

Table 14: Most important effects of the changeover to financial accounting and the new income calculation method

	Farm	Family
Rental of residence: *Old system:* cost rent *New system:* imputed rental value	Agricultural income **lower**	Private expenditure **lower**
Depreciation: *Old system:* straight-line *New system:* tax-optimising	Different depending on the year	-
Family social security contributions: *Old system:* Everything posted privately *New system:* Some items posted to the farm	Agricultural income **lower**	Private expenditure **lower**

[52] The first provisional calculations of the effect of the change in sampling procedure were conveyed to the public at an information event on 30 June 2016 (Hoop et al., 2016). The definitive results were presented at the Tänikon Agricultural Economics Conference on 15 September 2016.

[53] This range is not a confidence interval. It is based solely on the two alternative calculating methods, and does not take into account the additional error sources resulting from the sample (e.g. for sampling error or non-response error, see Chapter 2.6.1).

[54] For the difference for the individual farm types, see Hoop and Lips (2016).

In addition to the switch to a random sampling, the following methodological changes significantly influence income estimation:

- Firstly, the **income calculation method** plays a role. The main change is the posting in the farm's accounts of the employer's contributions for the first and second pillars of the pension scheme of the farm manager and his or her partner. The way in which the costs of the residence are calculated has also changed. Both changes lead to a 12% lower calculated income with a simultaneous reduction in private consumption of equal value.
- Secondly, the **switch from farm accounts to financial accounts** can also influence the results. This is particularly true in the case of depreciations, repairs, and own work capitalised, as well as services in kind from the farm to the private household.
- The previous system stipulated that all fixed assets be depreciated using a linear amortisation method (same amount of depreciation each year). With financial accounting, the amount of depreciation can be adjusted annually to the course of business, in order to even up income over the years and thereby reduce fluctuations in income tax over time.
- With financial accounting, this leads in good agricultural years to a lower agricultural income compared to farm accounting, whilst in poorer years the lower depreciation results in the higher agricultural income. The switch from straight-line depreciation to the depreciation reported in the financial accounting leads to an approximate 1% reduction in agricultural income in AY 2014.
- Furthermore, the income estimate is influenced by the **sampling criteria**, which determine which farms can actually participate in the FADN survey, and for which Swiss farms income is estimated. The new definition of the sampled population (see Chapter 2.2) results in a larger percentage of the small farms being excluded from the field of observation for the analyses. The average utilised agricultural area of the farms according to the new definition of the sampled population with the standard output rises by around 8 per cent. Total livestock numbers (in LU) are around 13 per cent higher compared to the old definition with the physical minimum thresholds. Hence, farm size – and correspondingly, agricultural income – tend to increase under the new criteria.

For Switzerland, the new estimate for **agricultural income** differs on average by minus 4% to minus 6% from the estimate in the old system. On average, agricultural income in the random sample ranges between CHF 63 700 and CHF 65 100 per farm, as compared to CHF 67 800 according to the old survey system. Whereas this difference amounts to minus 1 % to plus 2 % in the plain region (new estimate: CHF 77 400 – CHF 79 900), it comes to minus 11 to minus 12 % (CHF 57 600 – CHF 58 400) in the hill region, and to minus 5 to minus 6 % (CHF 49 900 – CHF 50 300) in the mountain region.

Working income is derived from agricultural income. After the equity invested in the farm has been remunerated at the ten-year-federal-bond interest rate, the remaining amount is divided by the number of family labour units. The working income per non-remunerated full-time labour unit amounts to CHF 42 400 – CHF 44 200 in the new sample as opposed to CHF 52 800 in the old sample, corresponding to a 16-20 % reduction. This results not only from the lower agricultural income, but also from the markedly higher use of family labour on the farms of the new sample.

As with agricultural income, the estimate for the three regions varies greatly. In the new sample, average working income per family labour unit in the plain region differs by 15 to 16 % (CHF 53 200 – CHF 54 300 instead of CHF 63 600), in the hill region by 23 to 26 % (CHF 38 300 – CHF 39 800 instead of CHF 51 500), and in the mountain region by 12 to 17 % (CHF 31 600 – CHF 33 800 instead of CHF 38 200) from the values of the Reference Farms sample.
In contrast to the agricultural income, **off-farm** income is 10 to 13 % higher in the new sample, standing at CHF 28 900 – CHF 29 600 as compared to CHF 26 300. This partially makes up for the lower agricultural income, reducing the difference in **total income** between the samples to just 3 to 4 % (now CHF 90 200 – CHF 90 900 as opposed to CHF 94 100, respectively).

3 Farm Management sample

The Farm Management sample includes monetary (accounting) and non-monetary (e.g. earnings in kind) data from Swiss farms, so that agricultural researchers can carry out **in-depth business analyses** on current topics.
The Farm Management sample is characterised by the following **more-detailed information** which is available over and above that provided in the Income Situation sample:

- Production-branch results (variable direct costing[55] with contribution margin)
- Internal deliveries
- Production-related information (such as crop or milk yield)
- More-detailed information on the household (such as private consumption or household assets)

Two forms of data usage can be distinguished: specific analyses for research purposes, as well as standard analyses that are compiled and published by Agroscope on a regular basis.
The **specific analyses** typically deal with the following topics:

- Determinants of the profitability of farms and production branches (e.g. structure, specialisation/diversification, organic/conventional, agricultural-policy measures, production technologies, sociodemographic factors)
- Economic and environmental performance and sustainability evaluation (e.g. through linking the FADN data and the data from the Swiss Agri-Environmental Data Network sample)
- Investments and investment promotion (e.g. through linking with the data from the agricultural credit information system E-MAPIS)
- Household income situation (importance of the various income sources, asset situation, pension scheme)
- Ex-ante evaluation of various agricultural-policy measures, using the data in the agent-based sector model SWISSland
- International comparisons (e.g analysis of milk-production costs in the International Farm Comparison Network).

The **standard analyses** meet the needs of various stakeholders:

1) Farmers, schools, extension and accounting offices:
   - Cross-comparison of own farm with reference group (benchmarking with cost/performance calculations based on full costs)
   - Reference data, for the comparison of different variants in farm management
2) Loan offices, assessment institutions (e.g. administration, insurance companies):
   - Reference data for assessing the profitability/sustainability of investments
   - Bases for estimation of capitalised earnings value
   - Basis for determining compensation (such as wayleave rights or feed-money standards)
   - Rents
3) Courts and social insurance institutions

[55] With variable direct costing, outputs (e.g. the revenue from crops or animals) and variable costs (e.g. fertilisers, feed) are allocated directly to the individual production branches (e.g. wheat, barley or livestock fattening, calf fattening). The result is the gross margin of the production branch, and is generally shown per area unit or output unit for purposes of the analysis.

4) Reference data for comparing a specific farm with a group of similar farms
5) Swiss Federal Statistical Office:
   - Data for Economic Accounts for Agriculture (EAA)
   - Data for calculating the standard gross margin
6) Federal Office for Agriculture and farmers' associations:
   - Development and evaluation of agricultural policy measures
   - International comparisons
   - Agricultural-engineering and workload analysis

Farms for the Farm Management sample can be sole proprietorships and group farming businesses that can supply monetary and non-monetary data at the required level of detail and quality. The farms must carry out **financial accounting with variable direct costing** which complies with the guidelines concerning the Farm Management sample and provides the minimum level of detail in accordance with the FADN accounting rules. As a technical requirement, data must be collected via compatible IT tools, with an interface to the Farm Management sample's DCollectZA survey tool (provided free of charge by the FADN). Currently, data is only supplied by accounting offices that have already provided data for the previous Reference Farm sample; however, a future expansion is planned in the event of interest on the part of new accounting offices.

For the following reasons, **data collection** for the Farm Management sample is not conducted by random sampling: (a) The percentage of farms with variable direct costing out of the total farm population is very small. Owing to a lack of information, the farms without variable direct costing cannot be excluded from the total farm population. Consequently, recruitment effort would be very high. (b) The level of detail involved in the data collection, and hence the associated time expenditure, is very high. Since participation in the survey is voluntary, we must therefore expect extremely low response rates, which would be much lower than those for the Income Situation sample.

We are also dispensing with a quota sample, which would involve the distribution of supply rights for certain strata among the accounting offices previously participating in the data survey for the Reference Farm sample. Past experience with the distribution of supply rights to the reference farms has shown that this control system would be very restrictive. In addition, the distribution of supply rights would be based on a great many assumptions (e.g. regarding the existing pool of farms) which might not always correspond to the current situation. Furthermore, participation should also be open to data suppliers who have not yet supplied data to the FADN.

In the new system, the data suppliers (potentially, any accounting office) are allowed to supply any farms to the FADN that meet the requirements in terms of **level of detail and quality**, as long as said farms lie above the **minimum specified threshold** in terms of size and belong to the **relevant strata** according to the selection plan (see Chapter 3.1). Financial incentives will be used to attempt to counteract the oversupply or undersupply of farms in certain strata (see Chapter 3.2).

3.1 Determining the minimum size and the relevant strata for the selection plan

The smallest farms which cultivate an insignificant share of the land or have an insignificant share of the livestock, and which are therefore less relevant for the analyses, are excluded from participation in the Farm Management sample survey. As in the previous Reference Farm system, physical farm-size indicators are used as **sampling criteria**.

However, these were limited to just two key figures: total **agricultural area** and total **livestock numbers** in LU. The farms were required to meet at least one of the following minimum thresholds in order to be considered for the FADN Farm Management sample and to qualify for compensation: utilised agricultural area of at least 10 ha, or livestock units (animals present on the farm) of at least 8 LU. Up until AY 2016, farm type 1531 (horses/ sheep / goats) constituted an exception, with a reduced minimum threshold of 4 sheep and/or goat LU.

The sampling plan for recruiting farms is allocated to strata according to **region** (plain, hill, mountain), **farm type** (according to S3 farm typology ZA2015, see Appendix 1), and **utilised agricultural area** (<= 20 ha, > 20 ha). Owing to limited financial resources, the Farm Management sample is limited to those strata that are typical or relevant for the Swiss agricultural sector. Here, the question arises as to what criteria are applied in selecting the relevant strata (i.e. combination of farm type and region).

The following criteria were used for the selection of relevant strata:

1) **Importance of the farm type** for the agricultural sector in the region in question according to the share of the total number of farms and the share of total potential production (measured on the basis of standard output).
2) Recruitability was assessed on the basis of both experience with the Reference Farm sample and the response rates for the Income Situation sample. Strata for which a low willingness to participate was expected were excluded in order to avoid inadequate representation of the stratum, as well as to make efficient use of available financial resources.
3) Relevance for the production-branch analyses was assessed on the basis of coverage of a production branch (measured on the basis of area under cultivation for plant crops and LU for livestock) by the individual farm types in the different regions.[56]
4) Relevance for research and extension was assessed on the basis of interest in specific production branches or farm types on the part of potential data users (expert groups, stakeholder surveys).

Taking account of the above-mentioned criteria, the **selection of the relevant combinations** (region and farm type) was made for the Farm Management Sample:

- Dairy farms (type 1521) are highly relevant according to all four criteria, and are the only farm type surveyed in all three regions.
- Arable farms (type 1511) in the plain region are included primarily owing to their high relevance for the plain region and for many production branches involving plant production.
- Although suckler-cow farms (type 1522) account for a relatively low share (especially in relation to standard output), these farms are included in the survey owing to a very high interest in them on the part of research and extension, their high recruitability, and their importance in terms of environmental services in the hill and mountain regions.
- 'Mixed cattle' farm types (type 1523) and 'horse/sheep/goats' (type 1531) should be surveyed in the mountain region only, owing to their importance for this region.[57]
- Coverage of the combined dairy and arable farms (type 1551) is limited to the plain region.

[56] A stratum (farm type and region) was defined as relevant for the production-branch analyses if it accounted for over 8% of the total area under cultivation or total livestock numbers.

[57] Due to a budget cut that came into force from 2017, farm type 1531 (horses / sheep /goats) will not be surveyed from AY 2017 onwards.

- The farms specialising in pigs and poultry (type 1541) as well as combined pig and poultry farms (type 1553) and other combined farms (type 1554) are only surveyed in the plain and hill regions.
- Two farm types (special crops and combined suckler-cow farms) are not surveyed. Despite special-crop farms (type 1512) being of great importance for the plain region and for several production branches (e.g. viticulture), this farm type was left out of the sample due to low recruitability. Combined suckler-cow farms (type 1552) are not surveyed, owing to their low relevance and poor recruitability.

Natural persons and unregistered partnerships (e.g. group-farming business) with the operating form of 'year-round farm' belong to the **surveyed population**. According to the FSS database available for 2016, after exclusion of the non-relevant strata and the farms lying under the stipulated minimum thresholds, the surveyed population consists of 34,198 farms.

Table 15: Coverage of the entire area under cultivation by the sampled population (for selected production branches involving plant production)

Selected strata as per sampling plan	Share of the **area under cultivation** covered by selected strata (in %)						
Farm type and region	Barley	Wheat	Silage maize	Oilseed rape	Sugar beet	Potatoes	Grapes
1511 Arable crops (plain region)	17.3	28.5	7.9	35.0	39.4	24.0	2.8
1521 Dairy cows (plain, hill, mountain)	6.5	2.7	12.5	0.6	0.4	2.2	0.3
1522 Suckler cows (hill, mountain)	0.9	0.5	1.1	0.1	0.0	0.2	0.2
1523 Mixed cattle (mountain)	0.6	0.1	0.4	0.0	0.0	0.1	0.1
1531 Horses/sheep/goats (mountain)	0.0	0.0	0.0	0.0	0.0	0.0	0.2
1541 Pigs/poultry (plain, hill)	0.5	0.2	1.1	0.1	0.1	0.2	0.0
1551 Comb. dairy/arable crops (plain)	13.3	13.7	16.4	12.1	16.9	18.3	0.7
1553 Combined pigs/poultry (plain, hill)	18.8	12.1	14.9	15.3	14.2	20.0	0.6
1554 Combined other (plain, hill)	25.1	21.8	29.5	19.0	14.4	12.6	2.5
Total share:	**83.0**	**79.7**	**83.9**	**82.1**	**85.3**	**77.6**	**7.4**

Source: FSS data, own calculations

Table 15 and Table 16 show the shares of the total area under cultivation and of total livestock numbers **covered by the sampled population**. As can be seen from these tables, after the exclusion of non-relevant strata and small farms from the Farm Management sample, over 80 per cent of the most important production branches can still be depicted. In fact, over 90 per cent of dairy-cow, fattening-pig, breeding-pig and poultry populations are covered by the selected (relevant) strata. The Farm Management sample therefore continues to be suitable for detailed analyses of these production branches. For certain production branches, however – viticulture, fruit and vegetable production – reliable analyses based on the Farm Management sample cannot be performed, since Special Crop farms are no longer surveyed.

Table 16: Coverage of total livestock numbers by the sampled population (for selected production branches involving livestock production)

Selected strata as per sampling plan	Share of **livestock numbers** covered by selected strata (in %)						
Farm type and region	Dairy cows	Suckler cows	Calf fattening	Beef-cattle fattening	Pig fattening	Pig breeding	Poultry
1511 Arable crops (plain)	0.5	1.1	2.7	7.6	0.7	0.8	2.0
1521 Dairy cows (plain, hill, mountain)	47.5	2.0	29.1	6.1	2.7	2.1	2.1
1522 Suckler cows (hill, mountain)	0.3	38.4	7.5	11.1	0.2	0.2	0.3
1523 Mixed cattle (mountain)	5.7	3.0	7.9	3.6	0.2	0.1	0.2
1531 Horses/sheep/goats (mountain)	0.1	0.5	0.2	0.2	0.1	0.0	0.1
1541 Pigs/poultry (plain, hill)	1.7	1.5	1.4	1.0	24.0	20.9	17.7
1551 Comb. dairy/arable crops (plain)	10.7	0.5	6.1	2.1	0.4	0.4	0.5
1553 Combined pigs/poultry (plain, hill)	10.7	7.9	9.9	10.4	51.3	60.0	63.9
1554 Combined other (plain, hill)	14.6	5.3	18.0	33.5	1.3	1.3	1.2
Total share:	**92.8**	**62.4**	**84.4**	**78.3**	**92.8**	**93.8**	**91.8**

Source: FSS data, own calculations

3.2 Sample size, sampling plan and compensation

In prepearing the selection plan, the desired **total sample size of 2100** farms is divided among the relevant strata.[58] Distribution is proportional to the number of farms in the sampled population per stratum. The advantage of a **proportional distribution** is that all farms are considered with the same weights in the calculation of the means. We speak of the so-called self-weighting sample, in which the mean of the sample corresponds to the mean of the sampled population. Of course, this only applies assuming that the sample can be taken exactly according to the selection plan, which is highly unlikely.

We dispensed with the optimal distribution as with the Income Situation sample (see Chapter 2.3.2). Unlike the Income Situation sample, the Farm Management sample is not used for analyses at the level of Switzerland as a whole. Consequently, the aim of an accurate estimate for the overall sample is irrelevant. Furthermore, the optimal distribution is not particularly useful if the composition of the farms cannot be ensured either by the random selection or by a quota sampling.

The proportional distribution of the sample according to the number of farms in the sampled population was calculated on the basis of the FSS data for 2016. The resultant sampling plan is shown in Table 17.

[58] Originally, the Farm Management sample had a target figure of 2100 farms; however, owing to a budget cut in 2017, the target figure was reduced to 2000 farms from AY 2017 onwards.

Table 17: Target number of farms for the Farm Management sample for AY 2016 (selection plan)

	Plain region		Hill region		Mountain region		Total
Farm type	UAA≤20 ha	UAA>20 ha	UAA ≤20 ha	UAA>20 ha	UAA ≤20 ha	UAA>20 ha	
Arable crops 1511	60	100					**160**
Special crops 1512							**0**
Dairy cows 1521	70	80	160	120	160	160	**750**
Suckler cows 1522			60		100		**160**
Mixed cattle 1523					80	100	**180**
Horses/sheep/goats 1531					70		**70**
Pigs/poultry 1541	30		40				**70**
Comb. dairy /arable crops 1551	30	100					**130**
Combined suckler cows 1552							**0**
Combined pigs/poultry 1553	160		90				**250**
Combined other 1554	100	140	90				**340**
Total	**870**		**560**		**670**		**2 100**

Compensation will be paid for all on-time, complete and plausible financial statements of farms lying above the stipulated minimum thresholds and delivered to the strata surveyed according to the selection plan. Only those farms fulfilling the requirements[59] of the Farm Management sample (*inter alia* financial accounting with variable direct costing, preparation and delivery with DCollectZA) can be delivered to the FADN.

The **compensation modality**[60] for the Farm Management sample was designed such that (a) the farms supplied are compensated according to the time and effort expended on data acquisition; (b) data delivery according to the selection plan is controlled by financial incentives. The first aim is achieved through the amount of the production-branch bonuses, which reflect the differing degrees of effort involved in collecting the data between one production branch and another. For the second aim, the stratum-specific basic payments are set such that a farm in an undersupplied stratum – i.e., one in which the number of supplied farms is less than the number stipulated by the selection plan – receives a higher basic payment than a farm in an oversupplied stratum. This creates the financial incentives for deliveries in under- or oversupplied strata.

The total amount of compensation per farm consists of the following components (as of the 2017 accounting year):

- A bonus of CHF 100 or CHF 50 for delivery before the stipulated **deadlines** in April and June, respectively;
- A **continuity** bonus of CHF 50 for farms which delivered complete and plausible data in both the previous and the current accounting year;
- A bonus ranging from CHF 10 to CHF 40 per fully delivered **production branch**;
- A bonus of CHF 50 per farm for **sole proprietorships**, as compensation for the complete recording of information on household assets/income;

[59] The requirements and prerequisites are described in detail and stipulated in the Farm Accountancy Data Network's Guidance Note. The Guidance Note can be accessed at www.agrarmonitoring.ch

[60] The current compensation modality can be accessed at www.agrarmonitoring.ch

- The stratum-specific **basic payment** determined each year by comparing the number of actually supplied farms with the distribution of the farms in the relevant strata in accordance with the selection plan. If fewer farms are delivered in a stratum than provided for in the selection plan, the basic payment per farm increases. In the case of oversupplied strata, the basic payment is reduced accordingly.

3.3 Recruitment process and tools used for data collection

The **selection** and **delivery** of the farm's accounting data by the accounting offices takes place following the selection plan stipulated by the FADN and the data-quality requirements.
The accounting offices are informed by the FADN about the **delivery process modalities** at the start of the data collection campaign. There are three possible delivery dates: April, June and August. After the official delivery dates, the quality of the datasets supplied up to that point is checked (see also Chapter 3.4). After the first two delivery deadlines, the accounting offices receive feedback on the fulfilment of the selection plan for the already-supplied farms at the overall sample level, and for their own accounting office (stratum information), as well as for farms with failed tests. This gives the accounting offices the opportunity to improve the data quality of individual farms by the final delivery deadline, and to optimise data delivery in terms of selection plan. At the end of the data collection campaign, the compensation is calculated according to the compensation concept and paid out to the accounting agencies by the FOAG.

The **collection tool** DCollectZA, which allows data from different accounting software programs to be imported and processed via an XML interface, was developed so that data could be supplied irrespective of the accounting software used in each case. DCollectZA is an access application that must be installed locally on the IT systems of participating accounting offices.

The accounting data (accounting journal) of a farm are imported into DCollectZA. For the WinBiz Agro software, there is an **interface** via which the data can be transferred directly from its database (datapool) into DCollectZA. In a so-called conversion, these accounting data are aggregated by DCollectZA to easy-to-handle characteristics for the Swiss FADN. For example, business processes recurring over the accounting year such as the monthly milk payments are consolidated into a 'milk yield' item. The conversion is based on rules that are in turn based on a fixed chart of accounts (FADN chart of accounts). However, since charts of accounts which differ in their details are used by the accounting offices, an allocation of the individual accounts to the accounts of the FADN chart of accounts – a so-called mapping – is needed. In DCollectZA, mapping templates for sole proprietorships and for group farming businesses are provided. The accounting offices can adapt these and provide their own customised templates. For the conversion, the allocation to the accounts (mapping) must be complete. The FADN chart of accounts for the Farm Management sample is based on the Agricultural SME chart of accounts (AGRO-TWIN AG, 2014).
In addition to the accounting data, non-monetary data and monetary data outside of the accounting are collected by the farms in DCollectZA.

Table 18: Mapping allocation example

Bookkeeping Account	FADN Account
1 000 – Cash	111 – Liquid assets
1 020 – Bank account	111 – Liquid assets
5 000 – Wage costs	51 – Wage costs
7 501 – Imputed rental value of private residence	75 115 - Imputed rental value of private residence

The non-monetary data consist of general information on the farm, such as ownership form, cooperation form, farming system, geographic location (zone, canton, municipality, altitude in metres), etc. In addition to the workforce information (year of birth, sex, education, working days, etc.), detailed information on surface areas and livestock numbers is collected. For the production-branch analyses, the produced and sold quantities (e.g. grains, milk, etc.) are important.

These non-monetary data are for the most part already available in other collection programs (e.g. Agro-Tech). The interfaces are implemented here, and data can be transferred directly, with the result that a double input is avoided. If no programs with interfaces are used for data collection, data can be entered directly in DCollectZA.

Asset, income and expenditure components of the farm managers' families not contained in the accounting should be added to DCollectZA in order to obtain a comprehensive picture of the farms. The tax return usually serves as a basis for this.

3.4 Plausibility check and quality assurance

The accounting offices carry out a **plausibility check** of the datasets centrally at the Swiss FADN and immediately after data transfer. This ensures that the tests used are always up-to-date, and that the results of the plausibility check are immediately available to the accountants. The data to be checked for plausibility are sent by the accounting offices via a secure file transport protocol (FTP) connection from DCollectZA. Alternatively, the data can be sent for plausibility checking via a web upload (www.za-dc.ch). When preparing the data, the plausibility-check step precedes the data-delivery step (Figure 15).

The **plausibility checks** encompass e.g. the following checks: If the farm in question is a livestock farm, the animals must be inventoried in the balance sheet; stated age must lie within a realistic range, or calculated price of the products of a specific production branch must lie within a plausible range, indicating that the physical and/or monetary yields were most likely entered correctly. In order to facilitate troubleshooting for the accountants, additional information is provided in the FADN's response to the accountants for each plausibility check. After checking once again, the accounting office can issue a message of correctness, thereby confirming that the data are correct, even if the plausibility check is effective.

After the official delivery deadlines, the **quality** and **completeness** of the datasets supplied up to that point are checked by the FADN. The check is performed using the same tests available to the accountants over the entire survey campaign for plausibility checking, as well as outlier tests. The outlier tests check key figures such as agricultural income for extremely small or large values.
Since correct account mapping (see Chapter 3.3) is very important for the contents of the conversion, the allocation of the farms' accounts is also checked. Feedback on these tests is given after the delivery deadlines.
The plausibility checks are constantly being improved with the aim of achieving high data quality whilst reducing the effort required for the accounting offices to complete the survey.

3.5 Data flow

The farms' annual financial statements are prepared by the accounting offices. In addition to the annual financial statement, a so-called analytical statement is prepared in which the services and variable direct costs are allocated to the individual production branches (variable direct costing). Furthermore, the key non-monetary figures such as livestock numbers, surface areas and earnings in kind are collected.

After completion of data collection for the farm in the accounting software, the data is imported into DCollectZA. The accounts of the bookkeeping program or of the chart of accounts specific to the accounting office are then allocated individually or via the mapping templates to the accounts of the FADN chart of accounts ('mapping'). Once all accounts are allocated, the entries are summarised into the characteristics used for the FADN via the so-called 'conversion'. Since not all of the characteristics required by the FADN are covered by the accounting programs, the next step involves adding these characteristics in the DCollectZA. Here we are dealing mainly with the income and assets section of the tax return, which is not covered by the bookkeeping. Once all the data have been entered, the plausibility check can take place.

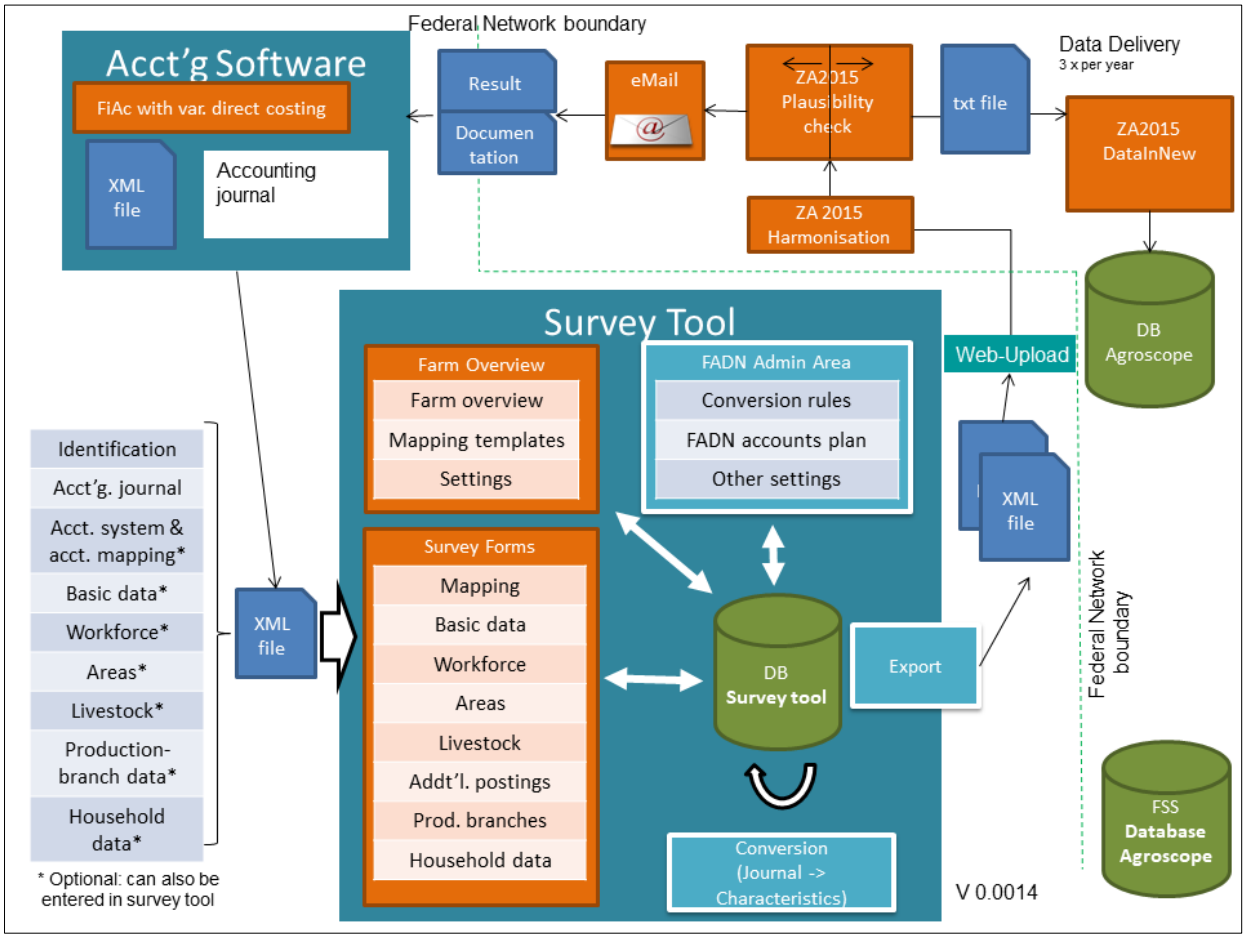


Figure 15: Farm Management Sample data-flow chart

For this, the data from the survey tool are sent to the FADN via FTP. At the FADN, the data are automatically prepared and plausibility-checked with a special IT tool; the accounting office is informed of the results by email. Once the data have been cleansed of errors, or have been reported as correct despite an error message, the accountant can deliver the farm data, again via FTP from the survey tool. Upon successful transfer, the accountant receives a receipt by email.

3.6 Sample for the 2016 Accounting Year

After a very low initial willingness to participate for AY 2016, the originally planned delivery deadline of 1 August 2017 was postponed to 31 August 2017, which significantly improved data delivery. Data from a total of 1535 farms in the various strata were thus available for the accounting year 2016 (Table 19), which were compensated with an average CHF 600 per farm.

Table 19: Number of farms in compensated strata of the Farm Management sample for AY 2016

	Plain region		Hill region		Mountain region		Total
Farm type	UAA≤20 ha	UAA>20 ha	UAA ≤20 ha	UAA >20 ha	UAA ≤20 ha	UAA>20 ha	
Arable crops 1511	12	26					**38**
Special crops 1512							
Dairy cows 1521	57	88	155	141	120	138	**699**
Suckler cows 1522			40		69		**109**
Mixed cattle 1523					47	44	**91**
Horses/sheep/goats 1531					13		**13**
Pigs/poultry 1541	23		20				**44**
Comb. dairy / arable crops 1551	15	70					**85**
Combined suckler cows 1552							
Combined pigs/poultry 1553	150		110				**260**
Combined other 1554	42	105	49				**196**
Total	**588**		**516**		**431**		**1 535**

Although full achievement of the target size was not expected owing to the changeover to the new delivery system, Table 20 nevertheless shows that in the 'dairy cow' and 'combined pigs and poultry' strata, the target was actually reached in the first year. Despite the highest basic payment being for the delivery of farms specialising in arable crops or horses/sheep/goats, these strata remain heavily underrepresented. Many other strata also failed to achieve the target number of farms.

Table 20: ACTUAL/TARGET ratio (in %) according to selection plan for AY 2016

	Plain region		Hill region		Mountain region	
Farm type	UAA≤20 ha	UAA>20 ha	UAA≤20 ha	UAA >20 ha	UAA ≤20 ha	UAA >20 ha
Arable crops 1511	20	26				
Special crops 1512						
Dairy cows 1521	81	110	97	118	75	86
Suckler cows 1522			67		69	
Mixed cattle 1523					59	44
Horses/sheep/goats 1531					19	
Pigs/poultry 1541	77		53			
Comb. dairy / arable crops 1551	50	70				
Combined suckler cows 1552						
Combined pigs/poultry 1553	94		122			
Combined other 1554	42	75	54			

The **representativeness** of the Farm Management sample can be assessed for individual strata by comparing the structural characteristics recorded by the participating farms (utilised agricultural area and aggregated livestock numbers) with the corresponding key figures from the FSS database (after definition of the sampled population).

This comparison was carried out for several selected strata with the aid of box plots (Figures 16-18). In general, it may be stated that the sample is not able to illustrate the total heterogeneity of the farms in the sampled population; this is apparent from the smaller spread of the distribution, and from the smaller interquartile range, which corresponds to the length of the box in the boxplot.

With heavily underrepresented strata, larger deviations of the median values are observed (see e.g. type 1511, 'arable crops', in Figure 16; type 1541, 'pigs/poultry', in Figure 17 or type 1523 in Figure 18). In most of the strata, the farms in the Farm Management sample are somewhat larger than in the sampled population. This can be explained by the fact that only farms with variable direct costing – usually a given for larger farms – can participate in the survey.

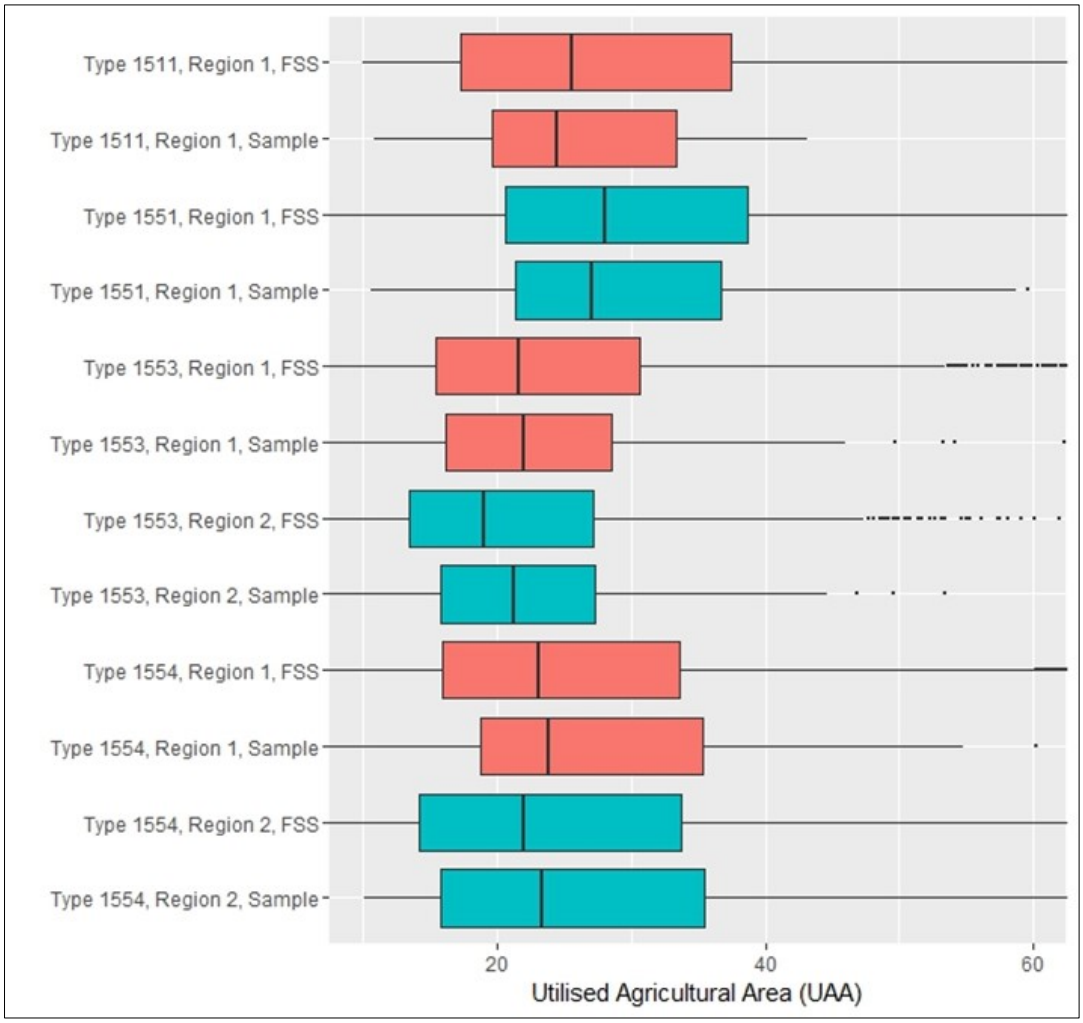


Figure 16: Distribution of the farms in the Farm Management sample compared to FSS data based on the utilised agricultural area (box plots for selected strata).

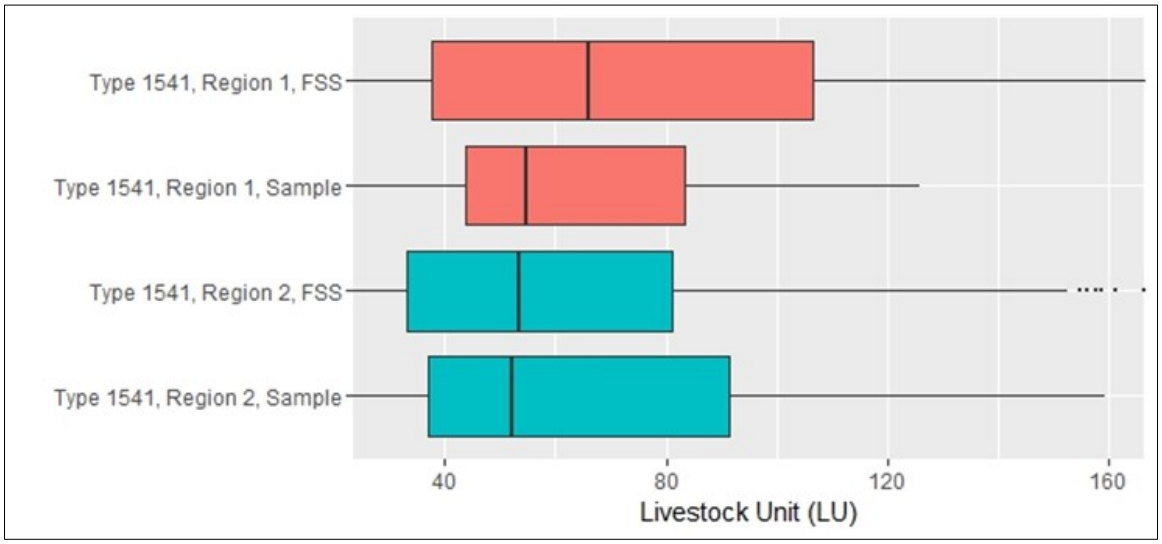


Figure 17: Distribution of the farms in the Farm Management sample compared to FSS data based on total cattle numbers in LU (box plots for selected strata).

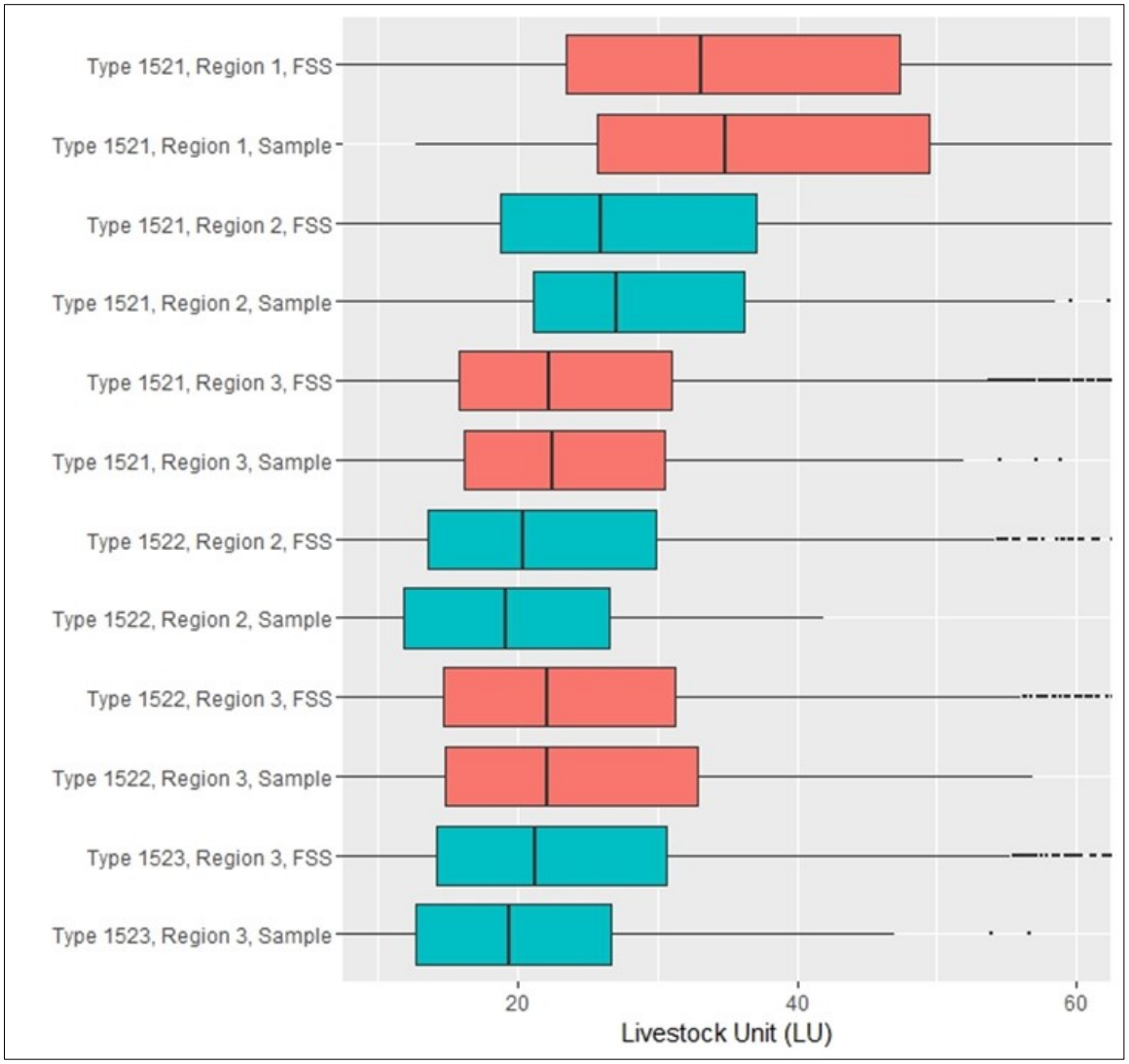


Figure 18: Distribution of the farms in the Farm Management sample compared to FSS data based on total cattle numbers in LU (box plots for selected strata).

3.7 Reporting

The standard publication based on the Farm Management sample is released in an electronic form (tables in Excel and PDF format), and contains the **results of the production branches**. For these, the directly apportionable services and variable costs are used to calculate the comparable **gross margin**. Results are depicted per production unit or per ha of crop area in plant production; in animal husbandry, the result is presented per animal or its physical output. The monetary figures are supplemented with further information on the production branch, such as crop yield and cultivation area.

Compared to the analysis of the reference farms, the method for the production-branch calculation has hardly changed. When switching from farm accounting to financial accounting with variable direct costing, the only change here concerns the valuation of deliveries to the private household, where the impact on results is gauged as minor. For both plant and animal production, numbers of farms supplied for the 2016 accounting year were substantially similar to those supplied for the reference farms of the 2015 accounting year. In the 'animal husbandry' production branches, the number of animals present is now used as a divisor instead of the number of animals owned. For the 'grazing stock' production branches, the key figures relating to the forage area are no longer shown.

The **farm-level indicators** are available in greater detail for the Farm Management sample than for the Income Situation sample. Thus, for example, for capital assets it is possible to differentiate between farm buildings, fixed installations, residential buildings, permanent crops, land, and melioration infrastructure. As with the Income Situation sample, it is not only sole proprietorships that are taken into account, but also group farming businesses. In the Household/Private section, in which no details are provided by the group farming businesses, only the results of the sole proprietorships are available.

With the Farm Management sample, we dispense with the standard publication of farm-level results. The reason for this is that there would be two results for the strata surveyed in the Farm Management sample: one from the Income Situation sample, and one from the Farm Management sample. There would thus be a risk of data being wrongly used or interpreted. Despite this, the farm-level data from the Farm Management sample are available in a standard format upon request. The farm-level results contain the groupings of the surveyed strata only. Owing to the sample design chosen, an aggregation of the results to the whole of Switzerland or to various regions is not possible, nor is it aimed for. If sufficient farms are available, groupings according to size categories can also be provided. Unlike the Income Situation sample, no precision information can be provided owing to the lack of random selection.

Appendix 1: Decision rules for the classification of farms according to S3-Typology ZA2015

S3	Type of farm	LU/ UAA	OAA/ UAA	SpC/ UAA	CaLU/ LU	DC/ CaLU	SC/ CaLU	HSG/ LU	PP/ LU	Additional conditions
1511	Arable crops	≤ 1	> 0.70	≤ 0.10						
1512	Special crops	≤ 1		> 0.10						
1521	Dairy cows		≤ 0.25	≤ 0.10	> 0.75	> 0.65	≤ 0.25			
1522	Suckler cows		≤ 0.25	≤ 0.10	> 0.75	≤ 0.25	> 0.25			
1523	Mixed cattle		≤ 0.25	≤ 0.10	> 0.75					Not 1521,1522
1531	Horses/sheep/goats		≤ 0.25	≤ 0.10				> 0.50		
1541	Pigs/poultry		≤ 0.25	≤ 0.10					> 0.50	
1551	Combined dairy/arable crops		> 0.40		> 0.75	> 0.65	≤ 0.25			Not 1511–1541
1552	Combined suckler cows				> 0.75	≤ 0.25	> 0.25			Not 1511–1541
1553	Combined pigs/poultry								> 0.25	Not 1511-1541
1554	Combined others									Not 1511-1553

All criteria in one line must be fulfilled simultaneously.

Abbreviations:

LU	Livestock units
UAA	Utilised agricultural area in ha
LU/UAA	Livestock units per ha UAA
OAA/UAA	Share of open arable area out of UAA
SpC/UAA	Share of special crops out of UAA
CaLU/LU	Share of cattle LU out of total livestock numbers
DC/CaLU	Share of dairy cows out of total cattle LU
SC/CaLU	Share of suckler cows out of total cattle LU
HSG/LU	Share of ‘horses, sheep and goats’ out of total LU
PP/LU	Share of ‘pigs and poultry’ out of total LU

References

Agroscope, 2017a. Pflichtenheft zum Projekt 17094 «Datenbeschaffung für die Zentrale Auswertung von landwirtschaftlichen Buchhaltungsdaten» mit den Losen Rekrutierungsstelle (Los 1) und Fachliche Hotline (Los 2). Agroscope, internal document, Ettenhausen.

Agroscope, 2017b. Erläuterungen zur betriebsindividuellen Rückmeldung der Stichprobe Einkommenssituation. www.einkommenssituation.ch → Datenlieferung → Betriebsindividuelle Rückmeldung.

AGRO-TWIN AG, 2014. Kontenrahmen KMU-Landwirtschaft, Revision 2013, Bern.

Bankier, M., 1988: Power Allocations: Determining Sample Sizes for Subnational Areas, *American Statistician*, 42 (3), 174–177.

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.

FSO, 2002. Charter of Swiss Official Statistics, Swiss Federal Statistical Office, Neuenburg.

FSO, 2012. Farm Census - additional survey. Published on 3 April 2012. Available at https://www.bfs.admin.ch/bfs/en/home/news/whats-new.assetdetail.4082354.html

BFS, 2016. Factsheet - Survey/Statistic. Farm Structure Survey. Published on 17 Oct. 2016. Available at https://www.bfs.admin.ch/bfs/en/home/statistics/catalogues-databases/surveys.assetdetail.2041588.html

Cochran WG, 1977: Sampling techniques, 3rd ed. New York: John Wiley & Sons

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

Dalenius, T. & Hodges, J.L., 1959. Minimum variance stratification, *Journal of the American Statistical Association*, 54, 88-101.

EU, 2011. Commission Implementing Regulation (EU) No 758/2011 of 1 August 2011 amending Regulation (EU) No. 1291/2009 concerning the selection of returning holdings for the purpose of determining incomes of agricultural holdings, available at http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32011R0758

EU, 2014. Commission Delegated Regulation (EU) No 1198/2014 of 1 August 2014 supplementing Council Regulation (EC) No 1217/2009 setting up a network for the collection of accountancy data on the incomes and business operation of agricultural holdings in the European Union, available at https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX:32014R1198

EU, 2015. Commission Implementing Regulation (EU) 2015/220 of 3 February 2015 laying down rules for the application of Council Regulation (EC) No 1217/2009 setting up a network for the collection of accountancy data on the incomes and business operation of agricultural holdings in the European Union, available at http://data.europa.eu/eli/reg_impl/ 2015/220/oj

Hausheer Schnider, J., Meier, B. & Schmid, D., 2004. Die Zentrale Auswertung von Buchhaltungsdaten im Umbruch, in Gysi, M. (Ed.): Landwirtschaftliche Forschung auf der Faktorseite, FAT-Schriftenreihe Nr. 60, Ettenhausen, pp. 25-33.

Hoop D., 2016. Harmonisierte Berechnung des landwirtschaftlichen Einkommens und des Arbeitsverdienstes in der Zentralen Auswertung von Buchhaltungsdaten. Agroscope, Ettenhausen.

Hoop D. & Lips M., 2016. Tieferes Einkommen. UFA-Revue 9-16.

Hoop D., Renner S., Dux D., Jan P., Schmid D., Weber A. & Lips M., 2016. Kurzinformation zum Stichprobenwechsel der Zentralen Auswertung von Buchhaltungsdaten. Agroscope, Tänikon.

Hoop D., Dux D., Jan P., Renner S. & Schmid D., 2017. Zentrale Auswertung von Buchhaltungsdaten – Grundlagenbericht 2016, Stichprobe Einkommenssituation. Agroscope, Ettenhausen.

Horvitz, D.G. and Thompson, D.J., 1952. A generalization of sampling without replacement from a finite universe. *Journal of the American statistical Association*, *47*(260), pp.663-685.

Lips, M., Mühlethaler, K., Roesch, A., Schmid, D. & Hausheer Schnider, J., 2009. Vorschlag der Arbeitsgruppe ZA 2015 für ein neues Konzept der Zentralen Auswertung von Buchhaltungsdaten, Abschlussbericht der Arbeitsgruppe ZA 2015, Agroscope Reckenholz-Tänikon, Ettenhausen.

Lips, M., Mühlethaler, K., Hausheer Schnider, J., Roesch, A. & Schmid, D., 2011. Stichprobenkonzept für das Schweizer Buchhaltungsnetz landwirtschaftlicher Betriebe, *Jahrbuch der Österreichischen Gesellschaft für Agrarwirtschaft und Agrarsoziologie*, 19(2): 131-138, Facultas Verlag, Vienna.

Lohr, S., 1999. Sampling: design and analysis. Duxbury Press, 1999

Lundström, S. & Särndal, C.E., 1999. Calibration as a standard method for treatment of nonresponse. *Journal of official statistics*, 15(2), p.305.

Meier, B., 2000. Neue Methodik für die Zentrale Auswertung von Buchhaltungsdaten an der FAT, Tänikon.

Moore, J.C. & Welniak, E.J., 2000. Income measurement error in surveys: A review. *Journal of official statistics*, 16(4), p.331.

Murbach F., Schmid D., 2017. Die zwei Barometer der Schweizer Landwirtschaft. *Swiss Federal Statistical Office*, Neuenburg.

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, University of Neuchâtel).

Rao, T. J., 2002. Mean of ratios or ratio of means or both? *Journal of Statistical Planning and Inference* 102, 129–138.

Renner S., Jan P., Hoop D., Schmid D., Dux D., Weber A. & Lips M., 2018. Das Erhebungssystem ZA2015 der Zentralen Auswertung von Buchhaltungsdaten: Stichprobe Einkommenssituation und Stichprobe Betriebsführung. Agroscope Science, No.68, Ettenhausen.

R Core Team, 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL *https://www.R-project.org/*.

Roesch, A., 2011. Stichprobeneffekt – Wie aussagekräftig ist der Vergleich mit dem Vorjahr? *Agrarforschung Schweiz 2 (5)*, 220-225.

Roesch, A., 2013. Assessment of the Swiss Farm Accountancy Data Network target populations, *Yearbook of Socioeconomics in Agriculture*, 6: 139-154.

Roesch, A., 2015. Impact of the SO Threshold on the Statistics of Economic Variables for the Swiss Agricultural Sector. *German Journal of Agricultural Economics*, 64(1), 33-41.

Roesch, A. & Lips, M., 2013. Sampling Design for two combined samples of the Farm Accountancy Data Network (FADN), *Journal of Agricultural, Biological, and Environmental Statistics*, 18(2): 178-203.

Satterthwaite, F., 1946. An Approximate Distribution of Estimates of Variance Components. *Biometrics Bulletin*, 2(6), 110-114.

Särndal, C.-E. & Lundström, S., 2005. Estimation in Surveys with Nonresponse, John Wiley & Sons.

Schmid, D. & Roesch, A., 2010. Die wirtschaftliche Entwicklung der schweizerischen Landwirtschaft 2009: Hauptbericht Nr. 33 der Zentralen Auswertung von Buchhaltungsdaten (Zeitreihe 2000-2009). ART-Bericht No. 734, Ettenhausen.

Tillé, Y. & Matei, A., 2016. sampling: Survey Sampling. R package version 2.8. https://CRAN.R-project.org/package=sampling

Woodruff, R., 1971. A Simple Method for Approximating the Variance of a Complicated Estimate. Journal of the American Statistical Association, 66(334), 411-414.

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössisches Departement für
Wirtschaft, Bildung und Forschung WBF
Agroscope