GJAE 64 (2015), Number 1

Impact of the SO Threshold on the Statistics of Economic Variables for the Swiss Agricultural Sector

Auswirkung des SO-Schwellenwertes auf ökonomische Kennzahlen des Schweizer Agrarsektors

Andreas Roesch
Institute for Sustainability Sciences, Zurich, Switzerland

Abstract

In future, standard output (SO) will be the economic variable used to define the target population from which the Swiss FADN sample is drawn. This study assesses the impact of the SO threshold on key economic variables at national level. The analysis demonstrates that raising the SO threshold will lead to higher average values of key economic variables such as work income per family labour unit. This result is confirmed by two entirely different approaches, the first of which takes the FADN data into account, and the second of which considers census data supplemented by imputed economic variables.

Key Words

target population; sampling; standard output; sensitivity; linear mixed regression model; model prediction

Zusammenfassung

In Zukunft wird die Grundgesamtheit, aus welcher die Stichprobe des schweizerischen Informationsnetzes Landwirtschaftlicher Buchführungen (INLB) gezogen wird, mithilfe des Standardoutputs (SO) definiert. Die vorliegende Studie untersucht den Einfluss des SO-Schwellenwertes auf den nationalen Mittelwert einiger wichtiger ökonomischer Variablen. Die Analyse zeigt, dass eine Zunahme der SO-Schwellenwerte zur Abgrenzung der Grundgesamtheit zu einer Erhöhung der Mittelwerte ökonomischer Parameter (z.B. Arbeitsverdienst pro Familienarbeitskraft) führt. Dieses Resultat wird durch zwei unterschiedliche Vorgehensweisen gestützt, wobei die erste auf den Daten des INLB beruht, während die zweite auf Daten der landwirtschaftlichen Strukturdatenerhebung und interpolierten ökonomischen Parametern basiert.

Schlüsselwörter

Grundgesamtheit; Stichprobe; Standardoutput; Sensitivität; gemischte lineare Modelle; Modellvorhersage

1 Introduction

The Farm Accountancy Data Network (FADN) is a European sample survey conducted annually to collect structural and accountancy data on farms, with the primary aim of evaluating the income of agricultural holdings (EU, 2010). The target population (TP) from which the sample is drawn typically only includes farms that exceed a specific threshold in terms of economic size. Put differently, agricultural holdings below a minimum threshold (EU, 2013a) are excluded from the TP. Given the scarcity of financial resources for data collection, farms that are run only as a hobby have also been excluded (VAN DER VEEN et al., 2012), as well as extremely small farms contributing only to a very limited extent to employment, produced output, arable acreage and other important key figures in agriculture (ROESCH, 2013).

To define the TP, the European Union (EU) used the European Size Unit (ESU) based on the Standard Gross Margin (SGM, EU, 2010) for measuring the economic size of a farm. Recently, the SGM was replaced by the Standard Output (SO, EU, 2008). The SO of a farm equals the monetary value of its output, and can be interpreted as a measure of the farm's gross agricultural production (EU, 2013a). The SO threshold for the economic size depends on the country, ranging between EUR 2000 in Bulgaria and EUR 25 000 in Germany (EU, 2013b). Assuming the same threshold for Switzerland as for Germany yields a figure of EUR 25 000 for Switzerland (approx. CHF 31,000 in 2013). Since 2010/2011, the German Federal Ministry of Agriculture (BMEL) has measured the economic size of a farm on the basis of the SO (BMELV, 2011).

In Switzerland, the TP is currently based on a non-financial criterion applying 11 physical thresholds relating to the agricultural area or livestock numbers, *at least one* of which must be exceeded (MEIER, 2005). These physical thresholds (e.g., UAA >10ha, open arable land (OAA) >6ha, number of dairy cows >6, number of goats >50) are to be replaced by regionally determined (plain, hill and mountain region) SO thresholds in the near future, ensuring that 95% of the accumulated SO per region is included in the TP (ROESCH, 2013). This means that the threshold is set at the inverse of the Lorenz curve of the SO at the value of 5%. As discussed in ROESCH (2013), the threshold value of 95% was selected for the following three reasons: (i) the threshold is intuitively clear from both a statistical and economic point of view; (ii) the measure is easily customisable on an annual basis; and (iii) most key structural variables are well covered in the TP.

Given that farm size is positively correlated with farm income (e.g. EL-OSTA and JOHNSON, 1998; JAN et al., 2011), the question arises as to the extent to which the estimated income depends on the definition of the threshold value. To the author's knowledge, there are no studies on the impact of the threshold on statistical measures of key economic variables. Recent literature on the FADN's sampling design (e.g. SKINNER et al., 1994; VROLIJK et al., 2006; KOKIC et al., 2010; ROESCH and LIPS, 2013) neither provides an accurate definition of the target population (TP) nor addresses the sensitivity of the SO threshold on the TP mean of key economic variables. In the present paper, we analyse how the threshold value affects key economic variables such as gross farm revenue (total revenues minus direct costs), agricultural income (remuneration of the farming family's own factors of labour, capital and land), and work income per annual family work unit at the national level for Switzerland.

For this analysis, we apply two different approaches for estimating the key economic variables – the first based on accountancy data from the Swiss FADN, and the second making use of the Farm Structure Survey (FSS), which provides structural and socio-demographic data for all Swiss farms. Since the key economic variables are not included in the FSS, they will be predicted by imputation using a linear mixed regression model.

The present paper is organised as follows: the data and method are described in Sections 2 and 3, respectively. The results are presented in Section 4, whilst conclusions are drawn in Section 5.

2 Data

The analysis of the SO threshold's impact on the TP's mean of key economic variables is based on the following two datasets:

(i) The Farm Structure Survey (FSS) data (FSO, 2013). This dataset provides a detailed insight into the structural, technical and socio-demographic situation of all Swiss holdings on an annual basis, but contains no economic data. The study is based on the FSS from 2012, comprising a total population of 56 575 farms. The Federal Statistical Office (FSO) has very recently started to include the SO as an additional (calculated) variable in the FSS data, compiling the SO values per farm (MURBACH, 2013). To calculate a given farm's SO, the standardised SO coefficients are multiplied by the number of hectares (for crops) or head of livestock (for animals) of the farm. Simply put, SO coefficients – which are calculated for over 90 separate crop and livestock categories as five-year averages – represent the monetary value of the output from one hectare of land or one head of livestock at farm-gate prices.

(ii) The Swiss FADN data from 2010-2012. This comprehensive database contains detailed information on annual cost accounting from approx. 3000 farms. The holdings are *not* randomly selected, which may result in serious biases. The FADN is administered by Agroscope (Institute for Sustainability Sciences ISS, Agricultural Economics).

The Swiss FADN data do not include the SO variable. SO per farm was computed from the number of livestock units (LSU) and crop area, based on the approach used by SCHÜRCH and SCHMID (2010). An analysis showed substantial differences in the descriptive statistics between the approach developed by SCHÜRCH and SCHMID (2010) and MURBACH (2013). In order to harmonise the two approaches, the SO coefficients suggested by SCHÜRCH and SCHMID (2010) and MURBACH (2013) were critically reviewed and slightly harmonised, e.g. striking differences were found for the SO coefficient for horses, since MURBACH (2013) ignores the fact that the coefficient for boarding horses should be much higher than that for horses intended for slaughter. Furthermore, MURBACH (2013) determines SO coefficients separately for each individual Swiss canton, whilst SCHÜRCH and SCHMID (2010) do not consider SO coefficients on a cantonal basis. Since the current Swiss data law does not allow the FADN farm to be linked to the corresponding FSS farm, differences at farm level cannot be analysed. This is reflected in substantial biases in the SO statistical measures (see Tables 1 and 3).

Descriptive statistics on selected structural variables for FSS and FADN data are provided in Table 1 for the year 2012. For purposes of comparison, the FSS population was delimited by the same physical thresholds as those applied to the FADN sample. The statistical t-test (p-values are given in the third column in Table 1) reveals that significant differences occur between the FSS and FADN means of a number of structural and socio-demographic variables. On average, FADN farms are significantly larger than FSS farms in terms of both area and livestock numbers. Highly significant differences are also found between the two categories for the number of employees (EMPL) and family workers (FAM). This is probably due to the fact that for FSS farms, the two variables FAM and EMPL had to be estimated from the number of persons in the three categories of employment rate <50%, employment rate 50%-75%, and employment rate >75%.

Table 1. Descriptive statistics (means) of FSS and FADN data (year 2012)

Variable (unit)	FSS	FADN (2012)	p
UAA (ha)	*21.10*	*21.80*	*<0.001*
OAA (ha)	5.47	5.55	0.37
GRASS (ha)	*14.95*	*15.61*	*<0.001*
WHEAT (ha)	1.61	1.64	0.29
MAIZE (ha)	1.26	1.25	0.11
STOCK (LSU)	26.0	26.1	0.24
CAT (LSU)	*19.7*	*20.7*	<0.001
FAM [-]	*1.55*	*1.21*	<0.001
EMPL [-]	*0.32*	*0.48*	<0.001
SO [kFr]	*197.3*	*224.5*	<0.001

Note: For purposes of comparison, the FSS population was delimited by the same physical thresholds as those applied in the FADN sample. Weighting has been applied to the FADN data where the sampling weight represents the number of farms in the TP represented by the sample member. Variable names are listed in Table 2.
Source: own calculation

3 Statistical Matching of Economic Data

The primary aim of this paper is to demonstrate the relationship between the SO threshold and the national mean for economic variables. In order to achieve this objective using FSS data, economic parameters were imputed to all FSS holdings. The imputation algorithm ('statistical matching') for predicting economic variables of FSS holdings was deduced from FADN data, since the latter is the only data to include both economic and structural information. Several methods for statistical matching have been suggested in the literature (SINGH et al., 1993; RÄSSLER, 2002; VROLIJK, 2004; VROLIJK et al., 2005; D'ORAZIO et al., 2006). Among others, these include single and multiple imputation, regression-based estimates, and Bayesian approaches. After extensive validation of imputation algorithms such as multiple imputation, it was decided to estimate economic variables for all FSS holdings from the outcome of linear mixed regression models applied to FADN data. For the notation, let

$$(y_{ij}, x_{pij}),\ i=1,\dots,m;\ j=1,\dots,n_i;\ p=1,\dots,k$$

denote the values of the response variable y and covariates (explanatory variables) x_p observed at times $t_{i1} < \cdots < t_{ij} < \cdots < t_{in_i}$ for farm i=1,…, m. The classical linear model

$$y_{ij} = \beta_0 + \beta_1 \cdot x_{1ij} + \beta_2 \cdot x_{2ij} + \cdots + \beta_k \cdot x_{kij} + \varepsilon_{ij}\,, \quad (1)$$

with i.i.d errors $\varepsilon_{ij} \sim N(0, \sigma^2)$ does not take into account the fact that we have repeated measures j=1,…, n_i on the same individual farm i (FAHRMEIR et al., 2013). Given that the error terms are correlated for individual farms over several years, the following linear mixed model with random intercept was applied for the prediction of economic variables:

$$y_{ij} = \beta_0 + \beta_1 \cdot x_{1ij} + \beta_2 \cdot x_{2ij} + \cdots + \beta_k \cdot x_{kij} + \gamma_{0i} + \varepsilon_{ij}\,, \quad (2)$$

where β_0 is the fixed population intercept, β_p, p=1,…, k are the fixed population slope parameters, and γ_{0i} is the individual deviation from the population intercept β_0. The term $\beta_0 + \gamma_{0i}$ is called "random intercept" for farm i (FAHRMEIR et al., 2013). Since the FADN data is a subset of a larger population, the farm-specific parameters are assumed to be random with i.i.d. deviations $\gamma_{0i} \sim N(0, \tau_0{}^2)$. The random intercept $\beta_0 + \gamma_{0i}$ is thus distributed according to $N(\beta_0, \tau_0{}^2)$.

The imputation of economic data to all FSS holdings was accomplished via the following two steps:

In a *first* step, the linear mixed regression model (Eq. 2) with a random intercept was applied to the FADN sample in order to predict the dependent (economic) variable y_i of farm i, given a set of k explanatory variables ($x_1, x_2, \dots, x_k$). A list of the complete set of structural variables x_p used in the regression model is provided in Table 2. Note that this variable set must be available in both the FADN and FSS data. All

Table 2. List of structural variables x_p used in the regression model

Explanatory Variables	Description
REG	Region: Plain region (REG=1), Hill region (REG=2), Mountain region (REG=3)
FAM	Number of workers belonging to the family (No.)
EMPL	Number of employees (No.; Part-time employees weighted according to hours worked)
ORG	Organic farm (ORG=1) or Non-organic farm (ORG=2)
UAA	Utilised agricultural area (ha)
OAA	Open arable land (ha)
POT	Potatoes (ha)
WHEAT	Wheat (ha)
MAIZE	Maize (ha)
SUG	Sugar beet (ha)
RAP	Oilseed rape (ha)
GRASS	Grassland (ha)
SPEC	Special crops: Vegetables, fruit, vineyards (ha)
FRUIT	Fruit (ha)
VEG	Vegetables (ha)
VINE	Vineyards (ha)
BER	Berries (ha)
FOR	Forest (ha)
STOCK	Livestock (LSU)
CAT	Cattle (LSU)
HOR	Horses (LSU)
PIG	Pigs (LSU)
SHEP	Sheep (LSU)
GOAT	Goats (LSU)

LSU = livestock unit

explanatory variables are treated as fixed effects. The random effect characterises the idiosyncratic variations at individual-farm level. Applying the linear mixed model separately for each farm type has been shown to be advantageous. This makes sense, given that some farm types (e.g. no. 21, 'Dairy Farms', and no. 22, 'Suckler Cows') focus on livestock, whilst others (e.g. no. 11, 'Arable Crops', and no.12, 'Fruit/Vegetables/Vines') mainly cultivate crops. Individual differences at farm level are modelled by assuming random intercepts for all farms. The economic variable y (e.g. 'agricultural income' (AI) or 'cashflow' (CF) is regressed against key structural variables x_p (predictors) which (i) were previously tested via simple robust linear regression for a statistically significant impact on the economic variable, and (ii) have – according to the criteria of HOOP and SCHMID (2013) – a certain relevance for the selected farm type (e.g. it is pointless to include the number of goats in the regression model for the farm type 'Special Crops'). For each farm type, the best model was selected by automatic backward elimination of all effects using the AIC criterion. First backward elimination of the random part is performed following by backward elimination of the fixed part.

The impact of multicollinearity (correlation among three or more independent variables) was analysed using the variance inflation factor (*VIF*) and tolerance values (HAIR et al., 1998). Severe multicollinearity was not detected. Heteroscedasticity (non-constant variance of the error term ε_{ij}) was thoroughly tested for using the Breusch-Pagan-Test and demonstrated via the robust covariance matrix estimator (GREENE, 2011).

In a *second* step, we use the regression coefficients β_p and the mean "fixed" population intercept β_0 derived from the FADN data for predicting economic variables such as gross farm revenue (GFR), AI, and work income (WI) per annual family work unit (AWU) for all FSS enterprises. This second step thus allows us to enhance the FSS dataset with key economic variables.

Table 3 gives a summary of a number of statistical measures for the four variables SO, GFR, AI, and WI per AWU derived from both the FADN sample and the FSS survey. The values listed reveal that in some cases the statistical measures of these four variables differ markedly between the FADN and the FSS. Table 3 reveals that figures for the GFR variable were markedly higher for the FADN sample than for the FSS sample. The tendency towards underestimation may be influenced by the following three factors: (i) the distribution of GFR differs between the FADN and FSS samples; (ii) only a limited number of explanatory variables were retained in the final mixed linear model; and (iii) the linear model ignores any non-linear effects such as economics to scale and/or interactions between production factors. Regarding the variable WI per AWU, a further uncertainty factor is produced by the inaccurate specification of the employment rate for the variables FAM and EMPL in the FSS dataset (cf. Chapter 2). Rather than being based on linear mixed model predictions, SO estimates were derived according to SCHÜRCH and SCHMID (2010) and MURBACH (2013) for FADN and FSS farms, re-

Table 3. Statistical measures of the four variables SO, GFR, AI, and WI/AWU for both the FADN sample and the FSS data. Unit: CHF

		SO	GFR	AI	WI/AWU
FADN	1st quartile	96 000	143 000	24 700	20 100
	median	166 200	219 000	50 300	39 900
	3rd quartile	288 300	324 300	79 900	63 900
	mean	224 500	259 800	56 000	43 700
FSS	1st quartile	79 100	92 000	38 600	23 500
	median	139 400	146 800	57 400	35 300
	3rd quartile	234 000	248 800	81 800	53 500
	mean	197 300	197 600	63 200	38 500

Note: For purposes of comparison, the FSS population was delimited by the same physical thresholds as those applied in the FADN sample.
Source: own calculation

spectively. Differences between FADN and FSS may thus be attributed to differences in the applied SO coefficients and the sampling error induced by the non-randomly selected FADN farms.

4 Results

4.1 Evaluation of the Statistical Matching of Economic Data

This section describes both the properties and performance of the linear mixed model. First of all, we illustrate the regression coefficients of the linear mixed regression model for the most common farm type, 'Dairy'. The regression coefficients β_p are listed in Table 4 for the three key economic variables GFR, AI, and WI per AWU. The regression coefficients for REG show that the region negatively affects all three economic variables specified in Table 4. The decrease in economic performance accompanying an increase in altitude is a well-known phenomenon associated with the more favorable climate for agricultural production and the flatter topography of the plain region as compared to the mountain region (e.g. HOOP and SCHMID, 2013). The number of family members (FAM) positively influences GFR and AI, whilst negatively influencing WI per AWU: whereas the linear model predicts an increase of CHF 6 985 per unit increase in EMPL, AI decreases by approximately CHF 7 900. This makes sense, as personnel costs – along with machinery and buildings – contribute significantly to external costs. The impact of UAA on GFR (the latter increasing by just CHF 607 per ha increase in UAA) is surprisingly low when compared to the sizeable impact of UAA on both AI and WI per AWU. This outcome is obvious when one considers that the variable STOCK is kept in the final model that predicts GFR for dairy

Table 4. Regression coefficients β_j and mean fixed intercept from the final linear mixed model for predicting the three economic variables GFR, AI, and WI per AWU applied to dairy farms (farm-type 21)

Variable	Fixed Regression coefficients β					
	GFR		AI		WI/AWU	
Mean fixed intercept	39 428	***	15 419	***	23 603	***
REG	-17 148	***	-7 355	***	-5 208	***
FAM	8 060	***	7 981	***	-4 809	***
EMPL	6 985	***	-7 924	***		
UAA	607	***	4 362	***	3 460	***
OAA	17.4	***				
GRASS			-42.2	***	-26.2	**
MAIZE	18.7	***	-41.5	**	-29.8	**
FOR	31.5	**				
STOCK	3 948	***	1 157	***	523	***

Note: See Table 2 for the abbreviations of the structural variables x_p used in the regression model. Data basis: Swiss FADN sample, 2010-2012
*** indicates statistical significance at the 1% level, ** indicates statistical significance at the 5% level
Source: own calculation

farms. This is reasonable, given that the number of dairy cows directly influences gross farm revenue in terms of the amount of milk and meat produced. Table 4 shows that the predicted GFR increases by CHF 3 948 per unit increase in STOCK. It is hardly surprising to see that the variable FOR is kept in the final model (at the 5% significance level) for predicting the GFR of dairy-farms, as the average forest area in 2012 came to 3.47 ha (HOOP AND SCHMID, 2013).

An overview of the model's predictive power for the above-mentioned key economic variables is given in Table 5 for all 11 farm types. Note that the computation of the coefficient of determination R^2 in mixed models differs from that of multiple linear regressions with fixed effects only. Here, we use the conditional pseudo-R^2 that can be interpreted as variance explained by both fixed and random factors, i.e. the full model (NAKAGAWA and SCHIELZETH, 2012). Furthermore, the number m of selected (independent/explanatory) variables x varies both among the farm types for a given economic variable and between different economic variables. This characteristic is related to the AIC criterion, which guarantees an optimised trade-off between the goodness of fit and the complexity of the model. We conclude from Table 5 that the GFR score is significantly higher than the AI and WI per AWU scores. For most farm types, the correlation coefficient R for GFR is approximately 0.98, implying that close to 96% ($= 0.98 \cdot 0.98 \cdot 100\%$) of the total variance can be explained when predicting GFR. This high score for GFR is not surprising, given that this quantity is closely linked to SO, and hence to the livestock numbers and arable crop areas. Table 5 shows that the number of variables incorporated in the final regression model varies distinctly between 3 (farm type no. 23) and 10 (farm type nos. 11 and 53) explanatory variables. A closer analysis of the selected independent variables in the final (i.e. after variable elimination) linear mixed model for GFR reveals that only the variable FAM (see Table 2 for an explanation of this abbreviation) remains in the final linear mixed model of all 11 farm types. The four variables UAA, STOCK, and OAA are included in the final model of seven farm types, whilst the three variables REG, EMP, and CAT are incorporated into the final specification of the model for six farm types. The other covariates are of minor importance for predicting GFR, as they are taken into account in the final model of three or fewer farm types.

The linear mixed model is also suitable for explaining a substantial percentage of the total variance in AI, e.g. 90% ($0.95 \cdot 0.95 \cdot 100\%$) for farm-type no. 31 and 79% ($0.89 \cdot 0.89 \cdot 100\%$) for arable-crop farms. For the variable WI per AWU, the averaged R (over all farm types) equals 0.83. As for AI, the prediction score for work income is the highest for farm-type no. 31, whilst it is lower for farm-type nos. 12 and 52, probably due to their higher intra-type variability. These results lead us to believe that the linear mixed regression model can serve as a powerful yet simple tool for predicting economic variables from the structural characteristics of the farm.

After predicting the key economic variables for all FSS farms (using the regression coefficients β_κ derived from the FADN sample), the sensitivities of mean key economic variables on the SO-threshold will be assessed on the basis of the FSS dataset as described in the following section.

Table 5. Conditional pseudo-R for the linear mixed regression model using FADN data from 2010-12

Farm type	GFR	AI	WI/AWU
11 (173)	0.99 (10)	0.89 (7)	0.78 (9)
12 (153)	0.98 (8)	0.81 (6)	0.72 (6)
21 (1641)	0.98 (9)	0.89 (7)	0.86 (6)
22 (270)	0.96 (8)	0.87 (7)	0.82 (5)
23 (252)	0.97 (3)	0.9 (4)	0.87 (7)
31 (68)	0.97 (7)	0.95 (6)	0.94 (5)
41 (127)	0.98 (6)	0.83 (5)	0.78 (4)
51 (327)	0.99 (9)	0.91 (8)	0.9 (7)
52 (92)	0.98 (8)	0.89 (6)	0.74 (5)
53 (649)	0.99 (10)	0.89 (7)	0.9 (7)
54 (603)	0.99 (8)	0.89 (8)	0.85 (7)

Note: The model is applied to the three dependent variables GFR, AI, and WI per AWU with separate treatment of the 11 farm types. The number of independent variables x (after variable elimination based on the AIC criterion) is given in brackets. Farm types and number of farms (in brackets, after adjustment for multiple counting) are given in the first column. Farm types: 11: Arable Crops; 12: Special Crops; 21: Dairy; 22: Suckler Cows; 23: Other Cattle; 31: Horses/Sheep/Goats; 41: Pigs/Poultry; 51: Comb. Dairy/Arable; 52: Combined Suckler Cows; 53: Combined Pigs/Poultry; 54: Combined Others

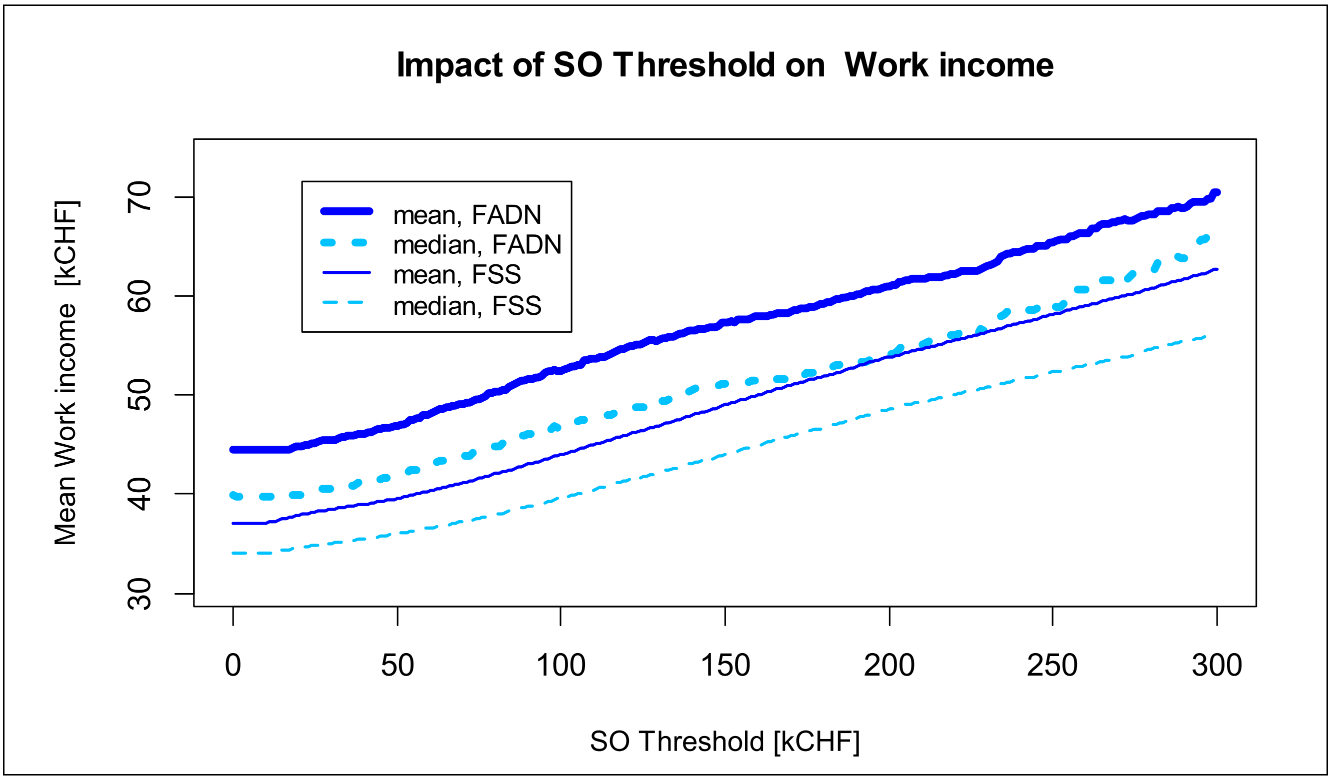
Source: own calculation

4.2 Impact of SO Threshold on Work Income per Family Member

This section provides a detailed analysis of the effect of the SO threshold on the national average of WI per AWU, a variable which is of the utmost importance for both the agricultural sector and policy decision-makers. The impact of the SO threshold on the national average for WI per AWU can be derived by using either the FADN data (enhanced by the SO) or the FSS data (enhanced by economic variables as outlined in Chapters 3 and 4.1). In order to avoid different characteristics owing to different TP definitions, only FSS farms within the FADN TP were considered for the following investigations.

Both approaches have their advantages and disadvantages. The use of FADN data to derive sensitivities has the drawback of limited sample size and the (expected) bias of non-random sampling. As regards the estimation of sensitivities from FSS data, although the TP comprises all agricultural holdings, the estimated economic data contain the prediction error from the linear mixed regression model. Hence, in order to compute the impact of the SO threshold on, say, the national mean for WI per AWU (for brevity's sake hereinafter referred to simply as WI), it is advantageous to use both the FADN data (augmented by the calculated SO) and the FSS data enhanced by the calculated WI. The computation of the mean and median WI were performed for the TP delimited by SO threshold values ranging from CHF 0 to CHF 300 000 in equal increments of CHF 1000. The result for 2012 is given in Figure 1 for both the WI mean and median. For both the FADN sample and the FSS data, the estimated national WI mean lies significantly above the median. This makes sense, given that the distribution of WI is significantly left-skewed. The difference between national mean and median WI is much more pronounced for the FADN data, however, indicating that the distribution of WI in the FADN differs considerably from that in the FSS data. The diverse characteristic of the distribution is also confirmed by the noticeable differences between central tendencies such as mean and median. Furthermore, Figure 1 clearly shows that for both FSS and FADN, the national WI mean increases along with the rising value of the SO threshold. Assuming an SO threshold of CHF 50 000, the estimated WI mean stands at CHF 46 900 and CHF 39 600 for the FADN and the FSS, respectively. Setting the SO threshold to CHF 50 000 leads to the exclusion of approximately 15 200 farms (27% of all FSS farms). A doubling of the SO threshold to CHF 100 000 raises the national WI mean to CHF

Figure 1. Impact of the SO threshold on the national median and mean of the work income per annual (family) work unit (WI per AWU)



Note: Ratios derived from FADN and FSS data are given as thick/thin lines. Basis: FSS, 2012; FADN, 2012. According to MEIER (2005), FADN and FSS target populations are identical.
Source: own calculation

52 400 for the FADN and to CHF 44 000 for the FSS. Roughly half of the FSS farms (27 500 farms or 48.5%) generate a total SO of less than CHF 100 000. Assuming a linear relationship between the SO threshold (within the interval CHF 50 000 to CHF 100 000) and the computed national WI, this indicates an increase in income for the FADN of approx. CHF 110 for each CHF 1000 increase in the SO threshold. The respective value for the FSS data is slightly lower, representing an increase of approx. CHF 90 for each CHF 1000 increase in the SO threshold. The fact that the positive relationship between the SO threshold and WI is found for both the FADN sample and FSS data strongly suggests that the specification of he SO threshold impacts on the national (and regional) statistics of the WI.

4.3 Impact of SO Threshold on Economic Variables

The procedure described in Section 4.2 above can easily be applied to other key economic variables. As outlined in Section 4.2, the computation of the sensitivities is based on the estimated means at the two SO thresholds of CHF 50 000 and CHF 100 000, assuming a constant sensitivity synonymous with the curves in Figure 1 being straight lines within the interval of interest. This assumption of linearity is reasonable, since the graphic inspection (not shown) shows that the sensitivities are fairly constant within the SO interval between approx. CHF 40 000 and CHF 200 000.

Table 6 summarises the impact of the SO threshold on the mean of selected economic variables. The analysis shows that the national mean of all investigated economic variables increases with rising SO thresholds. An increase of CHF 1000 in the SO threshold increases AI and TI by CHF 150 for the FADN sample (FSS: CHF 190) and CHF 120 (FSS: CHF 130), respectively. Since GFR is closely related to SO, it is obvious that the value of the SO threshold significantly influences the expected national average for GFR. The simulation shows that each CHF 1000 increase in the SO threshold will lead to an increase of approx. CHF 580 and CHF 720 for FADN and FSS in the national GFR mean, respectively. In summary, the national statistics for the economic situation of the Swiss agricultural sector will depend heavily on the specified SO threshold. Indeed, the results show that the national mean of the analysed key economic variables is likely to rise when the SO threshold is increased.

Table 6. National (arithmetic) mean of key economic variables based on FADN and FSS data (in 1 000`s of CHF) for an SO threshold of CHF 50 000 (column 2) and sensitivities per CHF increase in SO threshold for FADN (column 3)[1)]

Parameter	Arith. mean FADN[1)] *kCHF*	Sensitivity for FADN[1)] *CHF*
Agricultural Income (AI)	60.6 (66.2)	150 (190)
Total Income (TI)	86.3 (90.4)	120 (130)
Gross Farm Revenue (GFR)	159.0 (162.2)	580 (720)
Cash Flow (CF)	50.1 (54.2)	110 (130)
Farm Assets (FA)	825.5 (901.3)	1 580 (1940)

Note: Estimates are given per CHF 1000 increase in the SO threshold. Computation is based on the estimated means for the two SO thresholds of CHF 50 000 and CHF 100 000, assuming a constant sensitivity between these two threshold values.

[1)] FSS in brackets

Source: own calculation

5 Conclusions

In the near future, the sampling design of the Swiss FADN will be based on random sampling (ROESCH and LIPS, 2013). An important feature of the design will be the delimitation of the TP by the SO threshold, i.e. farms with an SO below this threshold will be excluded from the sample. This study assesses the impact of the SO threshold on key economic variables, mainly at the national level. The analysis demonstrates that raising SO thresholds will result in higher averages for key economic variables such as work income, gross farm revenue, cash flow, and farm assets. This finding is supported by both FADN and FSS data, the latter being enhanced by economic variables using mixed linear regression models for data imputation. This distinct sensitivity of the SO threshold on the statistics of economic parameters should therefore be borne in mind when interpreting the national and regional statistics describing the economic status of the Swiss agricultural sector. Furthermore, the results should encourage the responsible parties in Switzerland and the EU Member States to critically scrutinize the specified value of the SO threshold in their country, as well as to refrain from adjusting the value of the SO threshold unless the SO output coefficients have undergone significant adjustment.

References

BMELV (Federal Ministry of Food, Agriculture and Consumer Protection) (2011): Methodische Erläuterungen und Definitionen der Kennzahlen für die Buchführungsergebnisse der land- und forstwirtschaftlichen Testbetriebe. URL: http://berichte.bmelv-statistik.de/BFB-4114002-2011.pdf, accessed 31 August 2014.

D'ORAZIO, M., M. DI ZIO and M. SCANU (2006): Statistical Matching: Theory and Practice. Wiley, New York.

EL-OSTA, H.S. and J.D. JOHNSON (1998): Determinants of Financial Performance of Commercial Dairy Farms. Technical Bulletin No. 1859. Economic Research Service, U.S. Department of Agriculture, Washington, D.C.

EU (2008): Commission Regulation (EC) No 1242/2008, Annex IV Available at http://faolex.fao.org/docs/pdf/eur84629.pdf, accessed 31 August 2014.

– (2010): European Commission, Farm Accounting Data Network: An A to Z of methodology. Version 04/11/2010, DG Agriculture. URL: http://ec.europa.eu/agriculture/rica//pdf/site_en.pdf, accessed 11 April 2014.

– (2013a): Glossary: Standard output (SO). Available at http://epp.eurostat.ec.europa.eu/statistics_explained/index.php/Glossary:Standard_output_(SO), accessed 11 April 2014.

– (2013b): European Commission, Agriculture and Rural Development. Farm Accounting Data Network, Methodology. Available at http://ec.europa.eu/agriculture/rica/methodology1_en.cfm, accessed 11 April 2014.

FAHRMEIR, L., T. KNEIB, S. LANG and B. MARX (2013): Regression: models, methods and applications. Springer Verlag, New York.

FSO (Swiss Federal Statistical Office) (2013): Agricultural Census from the Swiss Federal Office for Agriculture 2012. Neuchatel.

GREENE, W. (2011): Econometric Analysis. 7th edition. Pearson Education, Upper Saddle River, New Jersey.

HAIR JR., J.F., R.E. ANDERSON, R.L. TATHAM and W.C. BLACK (1998): Multivariate data analysis. Prentice Hall, Upper Saddle River, New Jersey.

HOOP, D. and D. SCHMID (2013): Zentrale Auswertung von Buchhaltungsdaten. Grundlagenbericht 2012. Tabellenbericht. Agroscope Reckenholz-Tänikon Research Station ART, Ettenhausen.

JAN, P., M. LIPS and M. DUMONDEL (2011): Synergies and trade-offs in the promotion of the economic and environmental performance of Swiss dairy farms located in the mountain area. In: Yearbook of Socioeconomics in Agriculture 4: 135-161.

KOKIC, P.N., N. CHE and R.L. CHAMBERS (2010): Analysis of Economic Data Collected in Farm Surveys. In: Benedetti, R., M. Bee, G. Espa and F. Piersimoni (eds.): Agricultural Survey Methods. John Wiley & Sons Ltd., Chichester, UK. DOI: 10.1002/9780470665480.ch20.

MEIER, B. (2005): Analyse der Repräsentativität im schweizerischen landwirtschaftlichen Buchhaltungsnetz. Messung und Verbesserung der Schätzqualität ökonomischer Kennzahlen in der Zentralen Auswertung von Buchhaltungsdaten. ETH dissertation no. 15868. Zurich, Switzerland.

MURBACH, F. (2013): Table of SO coefficients used at the FSO for computing the SO per holding (not freely available). Personal communication.

NAKAGAWA, S. and H. SCHIELZETH (2012). A general and simple method for obtaining R2 from generalized linear mixed-effects models. Methods in Ecology and Evolution. DOI:10.1111/j.2041-210x.2012.00261.x.

RÄSSLER, S. (2002): Statistical Matching: A Frequentist Theory, Practical Applications and Alternative Bayesian Approaches. Springer Lecture Notes in Statistics No.168. Springer, New York.

ROESCH, A. and M. LIPS (2013): Sampling Design for Two Combined Samples of the Farm Accountancy Data Network (FADN). In: Journal of Agricultural, Biological, and Environmental Statistics. DOI: 10.1007/s13253-013-0130-5.

ROESCH, A. (2013): Assessment of the Swiss Farm Accountancy Data Network target populations. In: Yearbook of Socioeconomics in Agriculture 6 (1): 139-154.

SCHÜRCH, D. and D. SCHMID (2010): Standardoutput-Koeffizienten für Schweizer Landwirtschaft. In: Agrarforschung Schweiz 1 (11+12): 424-429.

SINGH, A.C., H. MANTEL, M. KINACK and G. ROWE (1993): Statistical Matching: Use of Auxiliary Information as an Alternative to the Conditional Independence Assumption. In: Survey Methodology 19 (**1**): 59-79.

SKINNER, C.J., D.J. HOLMES and D. HOLT (1994): Multiple Frame Sampling for Multivariate Stratification. In: International Statistical Review/Revue Internationale de Statistique 62 (3): 333-347.

VAN DER VEEN, H.B., I. BEZLEPKINA, R. DE HEK, R. VAN DER MEER and H.C.J. VROLIJK (2012): Sample of Dutch FADN 2009-2010. Design principles and quality of the sample of agricultural and horticultural holdings. LEI report 2012-061, LEI Wageningen UR, The Hague.

VROLIJK, H.C.J. (2004): STARS: Statistics for regional studies. Proceedings of Pacioli: New roads for farm accounting and FADN. Report 8.04.01. LEI, The Hague.

VROLIJK, H.C.J., W. DOL and T. KUHLMAN (2005): Integration of small area estimation and mapping techniques. Tool for Regional Studies. LEI, The Hague.

VROLIJK, H.C.J., H.B. VAN DER VEEN and J.P.M. VAN DIJK (2006): Sample of Dutch FADN. Design principles and quality of the sample of agricultural and horticultural holdings. Report 1.06.03.2006. LEI, The Hague.

DR. ANDREAS ROESCH
Institute for Sustainability Sciences ISS
Life Cycle Assessment
Reckenholzstr. 191, 8046 Zurich, Switzerland
e-mail: andreas.roesch@agroscope.admin.ch