Sample of Dutch FADN 2005

Design principles and quality of the sample of agricultural and horticultural holdings

H.C.J. Vrolijk
H.B. van der Veen
J.P.M. van Dijk

Project number 31002

April 2008

Report 1.08.01

LEI, The Hague

This publication has been produced on the basis of activities included within the statutory research tasks. The execution of these tasks is coordinated by the 'Centrum voor Economische Informatievoorziening' (Centre for Economic Information, CEI). The CEI monitors the independent character of the execution and the quality assurance.

The Agricultural Economics Research Institute (LEI) is active in a wide array of research which can be classified into various domains. This report reflects research within the following domain:

- ☑ Statutory and service tasks
- ☐ Business development and competitive position
- ☐ Natural resources and the environment
- ☐ Land and economics
- ☐ Chains
- ☐ Policy
- ☐ Institutions, people and perceptions
- ☐ Models and data

Sample of Dutch FADN 2005; Design principles and quality of the sample of agricultural and horticultural holdings
Vrolijk, H.C.J., H.B. van der Veen and J.P.M. van Dijk
The Hague, LEI, 2008
Report 1.08.01; ISBN/EAN 978-90-8615-214-8; Price €14.50 (including 6% VAT)
65pp., tab., app.

Mede voor de Europese Unie organiseren het CEI en het LEI jaarlijks de verzameling van technische en financieel economische gegevens van circa 1.500 bedrijven in de akkerbouw, tuinbouw en veehouderij. Voor nationaal beleidsgericht onderzoek wordt die informatie aangevuld met gegevens over bijvoorbeeld milieubelasting, natuurbeheer en plattelandsontwikkeling. Alle gegevens worden vastgelegd in het Bedrijven-Informatienet. In dit rapport wordt verantwoording afgelegd over de steekproef 2005, toegespitst op de Nederlandse bijdrage aan het Farm Accountancy Data Network van de Europese Unie. De diverse fasen, van het opstellen van het selectieplan, het werven van de bedrijven tot het beoordelen van de kwaliteit van de resulterende steekproef worden beschreven.

The EU Farm Accountancy Data Network (FADN) requires the Netherlands to yearly send bookkeeping data of 1,500 farms to Brussels. This task is carried out by LEI and CEI. The data sent to Brussels mainly involves technical and financial economic information. For national policy purposes additional data is collected, such as pesticide use, manure production, nature management, non-farm income and rural development. This report explains the background of the farm sample for the year 2005. The report mainly focuses on the Dutch contribution to the European Farm Accountancy Data Network. All phases from the determination of the selection plan, the recruitment of farms to the quality control of the final sample are described in this report.

Orders:
Phone: 31.70.3358330
Fax: 31.70.3615624
E-mail: publicatie.lei@wur.nl

Information:
Phone: 31.70.3358330
Fax: 31.70.3615624
E-mail: informatie.lei@wur.nl

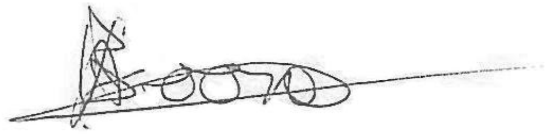
Contents

Page

Appendices

Preface

The EU Farm Accountancy Data Network (FADN) requires the Netherlands to yearly send bookkeeping data of 1,500 farms to Brussels. This task is carried out by LEI and CEI. This report explains the background of the sample for the year 2005. All phases from the determination of the selection plan, the recruitment of farms to the quality control of the final sample are described in this report. This report provides essential background information for the European Commission the Dutch Ministry and researchers of LEI and other organisations to fully understand the statistical aspects of the Dutch FADN sample.



Prof. Dr. R.B.M. Huirne
Director general LEI

Koen Boone
Head CEI

Samenvatting

1. Inleiding

Mede voor de Europese Unie organiseren het CEI en het LEI jaarlijks de verzameling van technische en financieel economische gegevens van circa 1.500 bedrijven in de akkerbouw, tuinbouw en veehouderij. Voor nationaal beleidsgericht onderzoek wordt die informatie aangevuld met gegevens over bijvoorbeeld milieubelasting, natuurbeheer en plattelandsontwikkeling. Alle gegevens worden vastgelegd in het Bedrijven-Informatienet. In dit rapport wordt verantwoording afgelegd over de steekproef 2005, toegespitst op de Nederlandse bijdrage aan het Farm Accountancy Data Network van de Europese Unie. De diverse fasen, van het opstellen van het selectieplan, het werven van de bedrijven tot het beoordelen van de kwaliteit van de resulterende steekproef worden beschreven.

2. Populatie en selectieplan 2005

De onderzoekspopulatie van het Bedrijven-Informatienet is gedefinieerd als alle bedrijven groter dan 16 Europese grootte-eenheden (ege) en kleiner dan 1.200 ege (tabel 3.1). Uit het steekproefkader (alle bedrijven in de landbouwtelling tussen 16 en 1.200 ege) wordt een gestratificeerde random steekproef getrokken. Economische omvang en het type bedrijf worden gebruikt als stratificatievariabelen. Voor het jaar 2005 omvat de totale agrarische populatie 81.830 bedrijven. (opgenomen in de Landbouwtelling). Het steekproefkader omvat 62.475 bedrijven. Deze bedrijven zijn verantwoordelijk voor 88% van de totale productiecapaciteit (tabel 3.1). Het selectieplan 2005 is in grote lijnen gelijk aan de selectieplannen van de jaren daarvoor. Het selectieplan geeft aan dat er 1.500 bedrijven in administratie dienen te worden genomen. Het daadwerkelijke aantal bedrijven is de afgelopen jaren lager geweest door capaciteitsproblemen.

3. Resultaat van de werving en kwaliteit van de resulterende steekproef 2005

Voor het jaar 2005 zijn 1.458 bedrijven uitgewerkt en aangeleverd aan Brussel (tabel 5.7). Hoofdstuk 6 geeft een kwantitatieve evaluatie van de resulterende steekproef. Een vergelijking tussen de onderzoekspopulatie en de totale agrarische populatie zoals beschreven in de landbouwtelling laat zien dat 23% van de bedrijven zich beneden de benedengrens bevinden. Deze bedrijven zijn echter verantwoordelijk voor slechts een klein deel van de totale productie. De onderzoekspopulatie dekt circa 90% van de productie van de meeste agrarische activiteiten. In de tuinbouw ligt het probleem bij de grotere bedrijven. Om dit probleem voor de toekomst te verminderen is de bovengrens van de steekproef opgetrokken. Tabel 6.2 geeft een nadere uitwerking van de dekking voor een groot aantal activiteiten. Tabel 6.4 geeft de samenhang weer tussen typen en agrarische activiteiten. Uit de tabel blijken grote verschillen in de mate van specialisatie van activiteiten. Slechts een beperkt percentage van alle vleesvarkens worden geproduceerd op gespecialiseerde vleesvarkensbedrijven. Aan de andere kant geldt dat bijna alle paddenstoelen worden geproduceerd door gespecialiseerde paddenstoelbedrijven. Twee belangrijke aspecten van steekproeven, de representativiteit en de betrouwbaarheid van schattingen worden geëvalueerd in paragraaf 6.3.3 en 6.3.4. Tabel 6.3 geeft voor een groot aantal variabelen een vergelijking tussen de waarde volgens de landbouwtelling en de schatting op basis van het Bedrijven-Informatienet. Deze informatie stelt de onderzoeker in staat om te beoordelen in hoeverre de steekproef representatief is voor zijn of haar specifieke onderzoeksproject.

Summary

1. Introduction

The EU Farm Accountancy Data Network (FADN) requires the Netherlands to yearly send bookkeeping data for 1,500 farms to Brussels. This task is carried out by the Agricultural Economics Research Institute (LEI) and the Center for Economic Information (CEI). The legislation of the FADN demands that the member states prepare a selection plan and a report on the results of the selection. This report fulfils this obligation. Furthermore the report gives an analysis of the quality of the sample.

2. Population and Selection plan 2005

The population (field of survey) of the FADN is defined as all farms above the threshold of 16 European Size Units (ESU). In the Netherlands farms between 16 and 1,200 ESU are included in the population (table 3.1). A stratified random sample is drawn, in which economic farm size and type of farming are used as stratification variables. The scheme for the types of farming is based on a Dutch version of the Common Agricultural Typology that is also used by EUROSTAT. The total agricultural population contains 81,830 farms according to the agricultural census. The field of survey contains 62,475 farms. These farms cover an important part (88%) of the production capacity (table 3.1). In the selection plan, LEI plans to select 1,500 farms for the 2005 accounting year. The real number has been lower in the last few years due to capacity problems.

3. Result of recruitment and quality of the sample 2005

For 2005, 1,458 farms were included in the sample and were delivered to Brussels (table 5.7). Chapter 6 gives a quantitative evaluation of the resulting sample. A comparison of the field of survey with the total agricultural population shows that 23% of the farms are below the lower threshold. These farms are responsible for a small percentage of production only. The sample results in a coverage of 90% of the production for most of the agricultural activities. In horticulture, part of the production is not covered because it takes place on farms above the upper threshold. Table 6.2 gives a description of the coverage of a large number of activities. Table 6.3 shows the relationship between types of farming and agricultural activities. The numbers show that only a limited percentage of pigs are produced on specialised pig farms, while at the other extreme almost all mushrooms are produced on specialised mushroom farms. Two important aspects of a sample, the representativeness of the sample and the reliability of estimates are evaluated in section 6.3.3 and 6.3.4. Table 6.4 evaluates for many variables whether there is a difference between the agricultural cen-

sus and the estimate based on the FADN sample. These tables provide useful information for specific research projects enabling the researcher to determine whether the sample is representative for his or her topic.

1. Introduction

1.1 Objective of the report

In 1965 the European Commission adopted a regulation (nr. 79/65/EEG) in which member states were obliged to set up a network for the collection of accountancy data on the incomes and business operation of agricultural holdings in the European Economic Community. The purpose of the data network is defined as the annual determination of incomes on agricultural holdings, and a business analysis of agricultural holdings. The Netherlands were required to provide financial economic information on 1,500 farms to Brussels.

For the management of the system, the EU requires information on the selection of farms that included in the national FADN systems. In particular the regulation prescribes the provision of data on the establishment of a selection plan and the recruitment of farms.

With respect to the selection plan the regulation EEG 1859/82 prescribes (article 6): Each Member State shall appoint a liaison agency whose duties shall be: ……to draw up and submit to the National Committee for its approval, and thereafter to forward to the Commission:

- the plan for the selection of returning holdings, which plan shall be drawn up on the basis of the most recent statistical data, presented in accordance with the Community typology of agricultural holdings;
- the report on the implementation of the plan for the selection of returning holdings.

This report provides all the relevant background information on the population, the selection plan, implementation of the selection plan and quality of the sample of data that is to be provided to Brussels and which forms the basis for a wide range of national research projects.

1.2 Structure of the report

Chapter 2 gives a description of the background of the Dutch FADN system. Chapter 3 describes the agricultural population in the year 2005. This chapter will also consider the demarcation of the population as used in the Dutch FADN. Also the design of the sample of the Dutch FADN system is described. Chapter 4 reports on the selection plan 2005. Chapter 5 provides information on the implementation of the selection plan and the recruitment of new farms. Chapter 6 provides a qualitative and quantitative evaluation of the sample 2005.

2. Statistical background of the Dutch FADN sample

2.1 Introduction

In the Dutch FADN detailed records on 1,500 agricultural and horticultural farms are kept. Besides financial economic information, a broad set of technical-economic, socio-economic and environmental-economic data is collected. One of the reasons for the Dutch FADN system is the legal obligation to provide information on the financial economic situation of farms to Brussels. However, an even more important use of the data can be found at the national level. Data from the FADN system are used for many national policy evaluations and research projects.

Based on a sample of farms estimations are made for the whole population. This might raise the question 'how can conclusions be drawn for the whole population if only a limited number of farms are observed'. The answer to this question can be found in the selection of farms that are included in the sample. A cook also doesn't eat all the soup to judge the quality of the soup. It is important to stir well before tasting; the spoon of soup should reflect all flavors in the pan of soup. The spoon of soup should be representative for the whole pan of soup. The same is true for the FADN sample. The farms that are included in the FADN should be representative for the whole population. In this way a sample can provide better information than a census (in which all units are observed). With a fixed budget it is much easier to collect good data on a limited number of farms instead of collecting information on all farms. With a limited number of farms and thus a limited number of data collectors, it is easier to ensure good procedures and good training to collect reliable data.

An important issue is how to ensure that the farms that are included in the FADN sample are representative for the whole population. Use is made of a disproportional stratified random sample. A stratified sample implies that the population is divided into a number of groups. Subsequently farms are selected from each of the groups. The variables on which the groups are defined should be relevant variables to make sure that the farms that are included in one group are similar (at least in the important aspects). Using this stratification, and selecting farms from each group, ensures that farms from all groups and thus with different characteristics are included in the sample.

Disproportional means that not all farms have the same chance of being included in the sample. Groups which are relatively homogeneous, i.e. farms which show large similarities, have a lower chance of being included in the sample. After all if all the farms are very similar, a limited number of observations is enough to draw reliable conclusions (in the extreme case that all farms are exactly identical, it would be enough to have only one observation). In case of less homogeneous groups it is important to have a larger number of observations to make reliable estimates.

The choice of the stratification variables has therefore an important impact on the representativeness of the sample.

This way of selecting farms make it possible to make unbiased estimates for the whole population of farms. Based on the sample farms in a certain group, estimations can be made for all the farms in that group. Stratification assures that farms are selected from all groups and therefore allowing estimations for all groups. All groups together make up the whole population. In the Dutch FADN this is achieved by assigning a weight to each sample farm. The weight is calculated by dividing the number of population farms in a group by the number of sample farms in this same group.

Stratification also improves the representativeness in case of non-response. If a farm, which is asked to join the FADN system, refuses, another farm in the same size class and of the same type of farming can be selected. If there is a difference between the selection plan and the actual implementation, stratification helps to improve the representativeness by taking into account the real sampling fraction.

Finally, stratification makes the maintenance of the sample easier. Due to attrition and changes in the population it is sometimes necessary to supplement certain groups. Stratification makes a more focused replacement possible.

3. Population 2005

3.1 Introduction

This chapter will describe the population or more precisely the field of observation as covered by the FADN sample. A lower threshold is used to define the field of observation. This threshold and the consequences of this threshold will be described in section 3.2. Section 3.3 describes the strata which are used to subdivide the population. Section 3.4 reports the number of farms in each of the strata.

3.2 Defining the field of observation

Collecting detailed information at farm level requires considerable time and money. To assure an efficient and effective allocation of the available budget, the sample design focuses on certain groups in the population (demarcation of the population). Given limited capacity it is important to apply a sampling procedure that optimises the reliability of the sample estimates (through stratification).

Regulation 1859/82 of the EU Commission (adapted by regulation EEG nr. 3548/85) defines the population (field of observation) for the Dutch FADN as those farms with a size of more than 16 European size units (esu). Until 2001 this threshold was translated into 16 Dutch size units (dsu), which is roughly similar to 18.7 esu. For the statistical use of the data and the comparability of results it was considered advisable to apply the esu threshold. Therefore the lower limit of the Dutch FADN system has been 16 esu since the year 2001.

In addition to a lower threshold there is also an upper threshold. This upper threshold has been adjusted every few years to take into account the growth of the average size of farms. Until 2001 the upper threshold was 800 dsu. In 2001 the upper threshold was raised to 1,200 esu. The percentage of farms and the agricultural output excluded due to this upper threshold has been growing since 2001. This is the reason why the upper threshold will be increased again from 2006.

Table 3.1 Number of farms and their relative economic importance (measured in European size units - esu) in the agricultural census 2005

	Number of farms	Percentage esu
All farms in the agricultural census (a)	81,830	100.00
Minus farms less than 16 esu	18,942	1.96
Minus farms larger than 1200 esu	413	10.09
Total of non covered farms (b)	19,355	12.06
Total of covered farms (a) - (b)	62,475	87.94

In 2005, 413 farms were excluded from the field of observation because of the upper threshold. These farms were responsible for 10.09% of the total production. There has been a strong increase in the production above the upper threshold from 6.96% in 2004 till 10.09% in 2005. This clearly shows that there is a strong growth in the size of the largest farms. Due to the lower threshold 18,942 farms were not covered by the FADN sample. Although this is a large number of farms, they are only responsible for 1.96% of the total production capacity. The number of farms and the share of economic production of these farms have slightly decreased compared to 2004. The population (field of observation) of the Dutch contribution to the EU FADN system is displayed in table 3.1.

3.3 Design of the stratification scheme

Farms are allocated to strata according to the following stratification variables: type of farming and size class. In the past a more detailed stratification scheme was used, but this resulted in numerous practical problems due to empty or nearly empty cells. Combining cells can easily lead to a distortion in the calculated results (a bias). Farms of a certain type of farming are divided into 3 size classes. In the past 4 size classes were used. The reduction of size classes can be explained by the problem of empty or nearly empty cells and the conclusion that a fourth size class only provided a very limited value in increasing the efficiency of the estimators (Vrolijk and Lodder, 2002).

In total 29 types of farming are distinguished (see table 3.2). For a number of types of farming a distinction is made between organic farm and non-organic farming. A compromise was found to fulfill the increasing demand for research on organic farms. Random selection of organic farms from the total population would result in a very low number of observations because of the low proportion of organic farms. The definition of separate strata would result in many practical problems. The number of strata would double. The problem of empty or nearly empty strata would increase seriously. In line with the existing stratification, a number of types of farming were selected where organic farming is especially relevant. The types that were originally selected were: field crop farms, dairy farms, field vegetables and combined crop farms (Vrolijk and Lodder, 2002). The growth in the organic sector was however lower than expected and aimed for by policy makers. This resulted in practical problems in the recruitment of organic farms, for example due to the fact that the number of farms according to the selection plan was close to or even higher than the actual number of farms in the population. To deal with this problem a number of organic strata have been combined. Organic field crops farms, field vegetables and combined crop farms have been integrated in one stratum organic crop farms (Vrolijk, 2006).

The break down in subtypes is as follows: field crop farms have been itemised in starch potato farms, organic crops and all other field crop farms. The vegetables under glass farms have been broken down in paprika, cucumber, tomato and other. Cut flowers under glass are divided in roses, chrysanthemums and other cut flowers. The dairy farms are split into organic and non-organic dairy farms. Within field vegetables and the combined crop farms the organic farms have been separated. These are subsequently combined with the organic field crop farms.

The final stratification and the size thresholds for each of the strata are displayed in table 3.2. The thresholds were determined by optimal stratification in 2000 (Vrolijk and Lodder, 2002) and have remain unchanged since then.

Table 3.2 Stratification of the Dutch FADN sample

Type of farming	Size class		
	1	2	3
Field crop farms			
- starch potatoes	16.0-73.2	73.2-177.9	177.9-1200.0
- organic crops	16.0-45.0	45.0-90.0	90.0- 1200.0
- other field crop farms	16.0-66.3	66.3-139.7	139.7-1200.0
Horticulture			
Vegetables under glass			
- paprika	16.0-245.1	245.1-479.5	479.5-1200.0
- cucumber	16.0-201.3	201.3-392.7	392.7-1200.0
- tomato	16.0-268.5	268.5-518.0	518.0-1200.0
- other	16.0-106.1	106.1-335.8	335.8-1200.0
Cut flowers under glass			
- rose	16.0-260.2	260.2-494.7	494.7-1200.0
- chrysanthemum	16.0-193.7	193.7-373.4	373.4-1200.0
- other	16.0-141.9	141.9-342.2	342.2-1200.0
Plants	16.0-185.4	185.4-463.5	463.5-1200.0
Other glass	16.0-107.5	107.5-292.3	292.3-1200.0
Field vegetables	16.0-85.8	85.8-256.5	256.5-1200.0
Fruit	16.0-63.9	63.9-139.2	139.2-1200.0
Nurseries	16.0-84.9	84.9-250.7	250.7-1200.0
Mushroom	16.0-187.5	187.5-444.6	444.6-1200.0
Bulbs	16.0-185.4	185.4-476.9	476.9-1200.0
Other open air	16.0-116.3	116.3-356.1	356.1-1200.0
Grazing livestock			
Dairy			
- organic	16.0-86.0	86.0- 127.5	127.5-1200.0
- non-organic	16.0-88.7	88.7-159.0	159.0-1200.0
Calf fattening	16.0-63.7	63.7-150.1	150.1-1200.0
Other grazing livestock	16.0-46.6	46.6-145.5	145.5-1200.0
Intensive livestock			
Breeding pigs	16.0-115.5	115.5-263.0	263.0-1200.0
Fattening pigs	16.0-60.4	60.4-160.5	160.5-1200.0
Integrated pig farms	16.0-128.8	128.8-252.9	252.9-1200.0
Laying hens	16.0-137.6	137.6-344.8	344.8-1200.0
Poultry	16.0-100.2	100.2-203.2	203.2-1200.0
Other intensive livestock	16.0-113.0	113.0-261.1	261.1-1200.0
Combined	16.0-81.1	81.1-205.5	205.5-1200.0

3.4 Number of farms in the population 2005

Table 3.3 presents the number of farms in the population (agricultural census 2005). In this table the stratification according to size class and type of farming is applied.

Table 3.3 The number of farms per stratum according to the agricultural census 2005

Type of farming	Size class			
	1	2	3	total
Field crop farms				
- starch potatoes	469	418	181	1,068
- organic crops	73	79	85	237
- other field crop farms	4,299	2,273	660	7,232
Horticulture				
Vegetables under glass				
- paprika	110	169	139	418
- cucumber	93	109	66	268
- tomato	68	128	105	301
- other	457	256	90	803
Cut flowers under glass				
- rose	98	133	166	397
- chrysanthemum	80	75	84	239
- other	850	693	295	1,838
Plants	566	408	256	1,230
Other glass	376	247	164	787
Field vegetables	490	272	77	839
Fruit	636	626	218	1,480
Nurseries	1,035	672	234	1,941
Mushroom	186	71	38	295
Bulbs	447	304	175	926
Other open air	813	392	130	1,335
Grazing livestock				
Dairy				
- organic	157	106	69	332
- non-organic	8,221	9,390	2,860	20,471
Calf fattening	344	515	171	1030
Other grazing livestock	5,215	2,047	302	7564
Intensive livestock				
Breeding pigs	911	487	101	1,499
Fattening pigs	891	457	95	1,443
Integrated pig farms	571	423	101	1,095
Laying hens	575	252	39	866
Poultry	189	200	62	451
Other intensive livestock	188	112	66	366
Combined	3,222	1,893	609	5,724
Total				62,475

This table shows that 62,475 farms fall within the field of observation. Dairy farms are clearly the largest group of farms. Almost one in every three farms is classified as a dairy farm. Compared to the 64,483 farms in the agricultural census in 2004, it is clear that the decrease of farms continues. In one year time 3% of the farms in the field of observation disappeared.

4. Selection plan 2005

4.1 Introduction

The allocation of the total capacity of sample farms is based on the relative importance and the heterogeneity of the different types of farming (Dijk et al., 1995a; Vrolijk and Lodder, 2002). Within each type of farming an optimal stratification (determination of thresholds of size classes) and optimal allocation is applied (distribution of sample capacity over the different size classes).

4.2 Selection plan 2005

The EU regulation prescribes the use of size class and type of farming as important variables in the stratification and the choice of farms. Due to differences in the exact stratification scheme it is necessary to take into consideration the different weights of farms in different strata (Dijk et al., 1995b).

The design principles of the sample of the FADN system facilitate an efficient alignment with the goals of the system (see chapter 2). A summary of the selection plan 2005 is provided in table 4.1. Given the goals of the FADN system the numbers provided in the table are the required number of observations per type of farming.

Table 4.1 Desired sampling size per type of farming (selection plan) 2005

		Number of farms		
		Main		Sub
type of farming	code	type	type	type
Field crop farms	*1*	*210*		
- starch potatoes			30	
- organic crops			30	
- other field crop farms			150	
Horticulture	*2 + 3*	*520*		
Vegetables under glass	2012		120	
- paprika				30
- cucumber				30
- tomato				30
- other				30
Cut flowers under glass	2022		100	
- rose				30
- chrysanthemum				30
- other				40
Plants	2022		30	
Other glass	other 2022 and 2013, 2023, 2039, 349 (> 50% glass)		30	
Field vegetables	2011		60	
Fruit	3210		40	
Nurseries	3480		40	
Mushroom	2033		30	
Bulbs	2021		40	
Other open air	other 2022 en 2013, 2023, 2039, 349 (< 50% glass)		30	
Grazing livestock		*420*		
Dairy	4110, 4120, 4370		340	
- non-organic				310
- organic				30
Calf fattening	4380		30	
Other grazing livestock	4410, 4420, 4430		50	
Intensive livestock	*5*	*230*		
Breeding pigs	5011		50	
Fattening pigs	5012		50	
Integrated pig farms	5013		40	
Laying hen	5021		30	
Poultry	5022		30	
Other intensive livestock	Other 5		30	
Combined	*6,7 and 8*	*120*		
		-------+		
Total		*1,500*		

5. Recruitment of farms 2005

5.1 Basic principles 2005

The recruitment for 2005 took place in two steps. At the end of 2005 farms which were recruited for the bookkeeping year 2006. The selection of farms was based on the agricultural census 2004. A substantial part of these farms were also used for the 2005 sample. In the summer of 2006 additional farms were recruited by a number of accounting offices to fill in some remaining gaps.[1] The goal of the recruitment was to increase the number of available farms in the bookkeeping system and apply a more strategic approach in the choice of types of farming in the EU variant and the CSP variant. The EU variant focuses on the financial economic indicators as required by the European Commission, the CSP (Corporate Social Performance) variant covers data on a wide range of topics, such as environment and animal welfare (see section 5.3 for a more detailed description of these variants).

5.2 Elaboration of selection plan

Table 5.1 gives a more detailed description of the selection plan as presented in table 4.1.

Table 5.1 Detailed selection plan 2005 per stratum

Type of farming	Size class			Total
	1	2	3	
Field crop farms				
- starch potatoes	10	10	10	30
- organic crops	10	10	10	30
- other field crop farms	45	51	54	150
Horticulture				
Vegetables under glass				
- paprika	10	10	10	30
- cucumber	10	10	10	30
- tomato	10	10	10	30
- other	10	10	10	30

[1] In the meantime slight changes in the selection plan were applied. This report presents the original selection plan for 2005. The changes will be documented in the report for 2006. The major change is the reduction of the number of farms in the type 'field vegetables'.

Table 5.1 Detailed selection plan 2005 per stratum (continued)

Type of farming	Esu size class			Total
	1	2	3	
Cut flowers under glass				
- rose	10	10	10	30
- chrysanthemum	10	10	10	30
- other	13	14	13	40
Fruit	12	14	14	40
Nurseries	13	13	14	40
Mushroom	10	10	10	30
Bulbs	13	13	14	40
Other open air	10	10	10	30
Grazing livestock				
Dairy				
- organic	10	10	10	30
- non-organic	103	104	103	310
Calf fattening	10	10	10	30
Other grazing livestock	17	16	17	50
Intensive livestock				
Breeding pigs	20	16	14	50
Fattening pigs	16	16	18	50
Integrated pig farms	14	12	14	40
Laying hen	10	10	10	30
Poultry	10	10	10	30
Other intensive livestock	10	10	10	30
Combined	37	41	42	120
Total				1,500

5.3 Recruitment of farms

Based on the available number of farms in the FADN sample and the expected number of farms ending their participation in 2004 an estimate is made of the number of farms to be recruited. Furthermore the variant of bookkeeping has been explicitly considered. An evaluation has been made of the policy and research relevance of sectors and based on this importance a decision has been made whether a type of farming is assigned to the EU variant, the CSP variant or a combination of both. This implied that some farms had to be switched to the other variant. In some cases this would result in the drop-out of the farm. This has been taken into consideration in the number of farms to be recruited.

Table 5.2 Number of farms to be recruited

Type of farming	Variant	Esu size class			Total
		1	2	3	
Field crop farms					
- starch potatoes	combi	4	0	0	4
- organic crops	csp	0	0	2	2
- other field crop farms	combi	9	12	17	38
Horticulture					
Vegetables under glass					
- paprika	csp	1	1	0	2
- cucumber	csp	0	0	0	0
- tomato	csp	4	0	0	4
- other	csp	4	1	7	12
Cut flowers under glass	csp				
- rose	csp	7	3	2	12
- chrysanthemum	csp	4	2	4	10
- other	csp	1	1	2	4
Plants	csp	0	3	0	3
Other glass	combi	8	4	7	19
Field vegetables	combi	10	0	10	20
Fruit	combi	5	0	0	5
Nurseries	eu	12	9	12	33
Mushroom	eu	9	4	7	20
Bulbs	combi	2	3	0	5
Other open air	eu	5	1	4	10
Grazing livestock					
Dairy					
- organic	combi	0	2	0	2
- non-organic	csp	16	0	3	19
Calf fattening	combi	4	0	0	4
Other grazing livestock	combi	12	4	5	21
Intensive livestock					
Breeding pigs	csp	0	0	2	2
Fattening pigs	csp	3	0	8	11
Integrated pig farms	csp	0	0	2	2
Laying hen	csp	0	0	0	0
Poultry	csp	0	0	0	0
Other intensive livestock	eu	5	5	5	15
Combined	combi	3	7	12	22
Total					301

Based on the number of farms to be recruited, as displayed in table 5.2, farms were randomly selected from the agricultural census of the year 2004. The random draw of farms took place per stratum. The number of drawn farms per stratum was 7 times higher than the required number of farms to be sure to have enough addresses even with a high non response rate in specific types of farming. The addresses were requested from an agency (Dienst Regelingen) of the Ministry of Agriculture. The farm identifiers of the randomly selected farms were sent to the Ministry who sent back the addresses of these farms (under the strict condition that this information was only used for the recruitment of farms for the FADN). Using these addresses farms were contacted and asked to participate in the FADN.

Farms are asked to participate in the system in order to compensate for attrition and to take structural changes in agriculture into account. Some of the farms approached during the recruitment phase refused to participate. These refusals do not cause problems if these farms do not differ from farms that participate in their place. In the case where farms that refuse to participate systematically differ from the participating farms, this could result in a bias. If for example older farmers are less inclined to participate, this will result in a different age distribution in the sample compared to the population. The representativeness of the data with respect to age will be called into question (whether this is a problem or not depends on the research goals and the extent to which the important variables correlate with age). The representativeness is analysed in chapter 6. Table 5.3 describes the response rate in the different types of farming. This table only includes those farms which were asked to participate in the CSP variant (this variant will be explained in more detail at the end of this section).

To develop a better understanding of the reasons for non-response a number of questions were asked to all farmers approached. Table 5.4 shows the results for the questions asked. In these questions the farmer had to indicate to which extend he/she agrees with a statement about his knowledge or his attitude. The table shows a clear difference between those farmers who are willing to cooperate and those who are not. The ones who are willing to participate are more informed about the activities of LEI and the existence of the FADN. The participants are also better informed about the use of the FADN data. Providing data is considered more useful by those who are willing to participate. The opinion about LEI with respect to the objectivity and the carefulness is better among the participants. The last question shows that non participants have a significant lower trust in the government.

Table 5.3 Response rate in different types of farming, recruitment for CSP variant

	Refusals	Recruited	Unsuitable	Total	Unsuitable (%)	Response (%)
Field crop farms						
- starch potatoes	7	2	0	9	0	22
- organic crops	0	0	0	0		
- other field crop farms	73	37	13	123	11	34
Horticulture						
Vegetables under glass						
- paprika	3	4	2	9	22	57
- cucumber	0	0	0	0		
- tomato	4	5	3	12	25	56
- other	21	6	14	41	34	22

Table 5.3 Response rate in different types of farming, recruitment for CSP variant (continued)

	Refusals	Recruited	Unsuitable	Total	Unsuitable (%)	Response (%)
Cut flowers under glass						
- rose	29	12	11	52	21	29
- chrysanthemum	29	14	4	47	9	33
- other	11	4	4	19	21	27
Plants	6	3	1	10	10	33
Other glass	1	0	0	1	0	0
Field vegetables	0	0	0	0		
Fruit	18	5	4	27	15	22
Nurseries	0	0	0	0		
Mushroom	0	0	0	0		
Bulbs	1	1	0	2	0	50
Other open air	1	0	1	2	50	0
Grazing livestock						
Dairy						
- organic	0	2	1	3	33	100
- non-organic	18	3	0	21	0	14
Calf fattening	0	0	0	0		
Other grazing livestock	65	13	14	92	15	17
Intensive livestock						
Breeding pigs	0	0	0	0		
Fattening pigs	2	0	1	3	33	0
Integrated pig farms	0	0	0	0		
Laying hen	0	0	0	0		
Poultry	0	0	0	0		
Other intensive livestock	1	0	0	1	0	0
Combined	8	2	3	13	23	20
Total	298	113	76	487		

Table 5.4 Attitude of farmers (-2 not agree till 2 agree)

	Non participant		Participant	
	average	SE	average	SE
1 Informed about LEI	1.49	0.08	2.23	0.06
2 Informed about the FADN system	0.50	0.09	1.39	0.14
3 Informed about the use of FADN data	0.20	0.08	1.11	0.14
4 Usefulness of FADN system	0.79	0.06	1.64	0.08
5 Usefulness of providing data	0.71	0.06	1.84	0.08
6 Carefulness of LEI	0.99	0.05	1.69	0.10
7 Objectivity of LEI	1.01	0.06	1.67	0.09
8 Trust in the government	-0.15	0.06	0.47	0.10

SE - standard error.

Using these same variables discriminant analysis was applied to find the factors that are most discriminating between farmers who are willing to participate and farmers who refuse to participate. The analyses of the attitude of farmers shows that 'usefulness of FADN system', 'usefulness of providing data' are the most important factors in predicting the participation of an individual farmer. This is a similar result compared to the previous recruitment (Vrolijk et al., 2006).

Table 5.5 describes the number of farms where accounts were completed for the first time for the bookkeeping year 2005. Due to several factors this is not exactly the same as the number of farms recruited. Firstly, farms can drop out during the first year of participation. Secondly, some farms were already recruited during a previous year, but due to capacity problems their bookkeeping was not completed for that year.

Table 5.5 Number of farms with 2005 as first year of completion of bookkeeping

Type of farming	Size class		
	1	2	3
Field crop farms			
- starch potatoes			
- organic crops			1
- other field crop farms	5	5	1
Horticulture			
Vegetables under glass			
- paprika		3	
- cucumber			
- tomato	1	3	2
- other	2	2	1
Cut flowers under glass			
- rose	3	4	3
- chrysanthemum	2	2	6
- other	1	3	1
Plants		2	2
Other glass	3	1	1
Field vegetables	2		
Fruit	2		1
Nurseries	2	1	2
Mushroom	6		
Bulbs	4	5	1
Other open air	2	2	3
Grazing livestock			
Dairy			
- organic	2		
- non-organic	15	5	1
Calf fattening			
Other grazing livestock	5	5	3

Table 5.5 Number of farms with 2005 as first year of completion of bookkeeping (continued)

Type of farming	Esu size class		
	1	2	3
Intensive livestock			
Breeding pigs		2	2
Fattening pigs			1
Integrated pig farms	1		
Laying hen		1	
Poultry		1	
Other intensive livestock	3	5	
Combined	2	6	11
Total	63	58	43

Tabel 5.6 Comparison of the field of observation (population) and the sample available for research purposes 2005 (agricultural census 2005)

Type of farming	Code	Number of farms		
		population	total	CPS
Field crop farms	*1*			
- starch potatoes		1,068	30	28
- organic crops		237	29	26
- other field crop farms		7,232	112	89
Horticulture	*2+3*			
Vegetables under glass	2012			
- paprika		418	35	30
- cucumber		268	34	29
- tomato		301	26	18
- other		803	29	18
Cut flowers under glass	2022			
- rose		397	32	20
- chrysanthemum		239	27	17
- other		1,838	50	38
Plants	2022	1,230	31	27
Other glass		787	27	11
Field vegetables	2011	839	40	9
Fruit	3210	1,480	37	30
Nurseries	3480	1,941	28	4
Mushroom	2033	295	24	0
Bulbs	2021	926	43	22
Other open air		1,335	30	11
Grazing livestock	*4*			
Dairy	4110+4120+4370			
- organic		332	32	30
- non-organic		20,471	299	238
Calf fattening	4380	1,030	29	14
Other grazing livestock	4410+4420+4430	7,564	46	25

Tabel 5.6 Comparison of the field of observation (population) and the sample available for research purposes 2005 (agricultural census 2005) (continued)

Type of farming	Code	Number of farms		
		population	total	CSP
Intensive livestock	*5*			
Breeding pigs	5011	1,499	54	43
Fattening pigs	5012	1,443	41	29
Integrated pig farms	5013	1,095	44	39
Laying hen	5021	866	37	33
Poultry	5022	451	31	24
Other intensive livestock	other 5	366	25	4
Combined	*6-8*	5,724	104	60
Total		62,475	1,406	966

In table 5.6 a distinction is made between CSP observations (corporate social performance) and the total number of observations. Poppe (2004) describes that the introduction of a new bookkeeping system and budget cuts have resulted in a large pressure on available capacity. To deal with this pressure, a flexible data collection system has been introduced with two main variants in the data collection; the EU variant and the CSP variant. In the EU farm income variant the most essential financial economic information is collected. This is the information that each member state is obliged to provide to Brussels. The information covered in this variant mainly focuses on family farm income, the balance sheet, a limited number of technical data (cropping pattern, livestock) and information on the EU subsidies. In the second variant, the CSP variant, a wide range of data is collected for EU and national purposes. It covers all the topics that are nowadays considered relevant in a report on the corporate social performance of a company or a farm. Therefore, besides the financial economic information as collected in the EU variant, a wide range of data is collected such as environmental data, other farm incomes, off farm income, animal welfare, animal health and the level of innovation of firms.

5.4 Supply of farm results to the European Commission 2005

The final delivery of 2005 data to EU has taken place in December 2006. Data of 1,458 farms have been provided to Brussels (table 5.7). This is the highest number of farms since many years and is close to the intended 1,500 farms.

Table 5.7 Comparison between the number of farms supplied to the Eu and those available for research

Bookkeeping year	Provided to the European Commission	Weighted farms available for research	Other available farms a)
1990/91	1,587	1,576	12
1991/92	1,505	1,547	8
1992/93	1,513	1,516	7
1993/94	1,525	1,520	7
1994/95	1,546	1,534	13
1995/96	1,536	1,530	6
1996/97	1,551	1,545	6
1997/98	1,529	1,522	7
1998/99	1,368	1,363	5
1999/00	1,341	1,334	7
2000 b)	N/A	N/A	N/A
2001	1,330	1,310	20
2002	1,358	1,344	14
2003	1,437	1,399	38
2004	1,420	1,392	28
2005	1,458	1,406	52

a) Other available farms are farms that are also available but without a weight. Reasons for not having a weight are: a farm is outside of the defined field of observation because a farm is too large or to small according to the information in the agricultural census. In alternative weighting systems (based on the characteristics of the farm these farms do get a weight; b) Bookkeeping year 1999/00 ended for arable farms and husbandry at April 30, 2000. Due to capacity problems related to IT problems, farm data for the period of April 30, 2000 to December 31, 2000 (respectively January 1, 2000 to December 31, 2000) are not processed but estimated based on data of 1999/00 and 2000/01.

6. Evaluation sample 2005

6.1 Introduction

In this chapter the FADN sample for the year 2005 is evaluated in a qualitative and quantitative way. Section 6.2 provides an evaluation of the methodology of stratification and weighting. A crucial element is the calculation of weights. Section 6.3 provides the quantitative evaluation of the year 2005. This section focuses on the quality of the estimations that can be made based on the sample.

6.2 Evaluation of stratification and weighting

6.2.1 Introduction

This section deals with some practical problems related to the estimation process. Weights of individual farms are used to make estimations of frequencies, totals and averages of groups of farms (aggregated results) based on the data from the agricultural census and the FADN data.

The method to calculate the weights of individual farms is crucial. The goal is to achieve unbiased estimates with a minimal variance. This enables the estimation of the confidence interval of the real population value and the minimisation of the total error. This is true for direct estimators. In case of ratio estimator this is not necessarily true, (Vrolijk et al. (2001) and Appendix3 for a more extensive description of ratio estimators and other estimators).

In the next section the method to calculate the weights of the farms is described in general terms. The method applied to calculate the weights is evaluated from a practical and theoretical perspective.

6.2.2 Method of calculation of weights

The objective of the Dutch FADN system is to give a representative view of the total population. The question is therefore how to draw conclusions on totals, averages and frequencies that are valid for the whole population based on individual farm data. For example how much is the average family farm income of all farms in agriculture and horticulture. The solution is found in weighting: the individual farm data are raised to the population level (for some variables the estimated values can be compared to the data that is available for the whole population, i.e. data which is included in the yearly agricultural census). A weight is assigned to every observed farm in the FADN system. The weight is defined as the ratio between the number of farms in a stratum according the agricultural census and the number of farms in the sample (in the FADN system). For the assignment of farms in the FADN system to strata the information from the year 2005 is used. This data can be different from the data when the farm was chosen in the system for the first time. This implies some kind of post-stratification. Weights can be calculated as soon as a substantial number of farms have been completed. During the year, when additional farms are completed, the weights are recalculated. The weights of the farms are recalculated until the accounts of all farms are completed and the final set of weights can be established. For preliminary estimations based on for example 50% of the farms, one should be aware of the fact that this 50% is not necessary representative for the whole population.

Table 6.1 Sampling fractions in different strata (sample 2005)

Type of farming	Size class		
	1	2	3
Field crop farms			
- starch potatoes	0.02	0.02	0.06
- organic crops	0.14	0.13	0.12
- other field crop farms	0.01	0.02	0.08
Horticulture			
Vegetables under glass			
- paprika	0.09	0.06	0.07
- cucumber	0.11	0.09	0.15
- tomato	0.15	0.08	0.10
- other	0.02	0.04	0.11
Cut flowers under glass			
- rose	0.10	0.08	0.06
- chrysanthemum	0.13	0.13	0.12
- other	0.02	0.02	0.04
Plants	0.02	0.02	0.04
Other glass	0.03	0.04	0.06
Field vegetables	0.04	0.07	0.26
Fruit	0.02	0.02	0.06
Nurseries	0.01	0.02	0.06
Mushroom	0.05	0.14	0.26
Bulbs	0.03	0.04	0.08
Other open air	0.01	0.03	0.08
Grazing livestock			
Dairy			
- organic	0.06	0.09	0.14
- non-organic	0.01	0.01	0.04
Calf fattening	0.03	0.02	0.06
Other grazing livestock	0.00	0.01	0.06
Intensive livestock			
Breeding pigs	0.02	0.03	0.14
Fattening pigs	0.02	0.04	0.19
Integrated pig farms	0.02	0.03	0.14
Laying hen	0.02	0.04	0.26
Poultry	0.05	0.05	0.16
Other intensive livestock	0.05	0.09	0.15
Combined	0.01	0.02	0.07

The (post) stratification of the farms is based on the agricultural census 2005. The population in a specific stratum is continuously changing; therefore the farms that belong to a stratum in 2004 are not exactly the same as the farms that belong to that stratum in 2005. Due to these changes farms included in one stratum could have had different inclusion probabilities at the time of recruitment. In theory, to achieve unbiased estimators these differences in inclusion probabilities should be taken into account in the estimation process. However, the consequence of this would be a very complicated system with many different substrata with different inclusion probabilities. Therefore this complicated procedure is not applied. As a result, the theoretical assumption of a strict a-select sample cannot be validated.

Although the calculation method applied in practice can lead to systematic distortions between estimated values and real values, the assumption of a random sample is made. This leads to several attractive consequences. The method to calculate weights is relatively easy, it involves a limited set of homogenous strata and it results in a more effective use of data.

Because of the applied sampling procedure (see section 2.1) the different strata have different sampling fractions. Strata with relatively homogenous units have a lower sampling fraction than very heterogeneous strata. This also implies that farms have very diverging weights. Farms from a homogenous cluster will have a larger weight (in principal the reciprocal of the sampling fraction) and therefore represent a larger number of farms. The differences in sampling fractions are shown in table 6.1. These percentages are calculated by dividing the required number of farms in the selection plan (table 5.1) by the number of population units (table 3.3).

6.2.3 Remarks on the weights of 2005

In the report on farm results 2005 the research population is defined as all farms in the agricultural census 2005 (between the lower and upper threshold). The weight per farm is calculated as the ratio between the number of farms in the census and the number of farms in the sample.

In the calculation of aggregated results (averages, frequencies and totals) for the year 2005 the agricultural census 2005 is the starting point. Because of the complete registration of farms in the population (almost all farms are registered in the agricultural census) the aggregated numbers of farms are exactly the same as the number of farms in the census. However, in using these numbers in the calculation of weights for estimations for 2005 two remarks should be made.

Every year all horticultural and agricultural farms are registered in the agricultural census, but this registration only represents the situation at a certain moment during the year. Therefore it is possible that farms are missing from this registration. Furthermore the trend is for number of farms to fall significantly (this trend is stronger for certain types of farms and less strong for others). As a consequence estimations for the year 2005 might be overestimations of reality. Distortions in the number of farms in the census can therefore cause incorrect estimations of aggregates.

Furthermore the typology of farms according to the agricultural census might differ from the typology according to the FADN data. The census reflects the situation at a certain point in time, while the FADN system describes the farm during a whole year. In order to take into account these differences two weighting methodology are available in the Dutch FADN system.

6.3 Quantitative evaluation of 2005

6.3.1 Introduction

This section focuses on the quality of the estimations based on the FADN sample 2005. Section 6.3.2 provides information on the coverage of the sample. Section 6.3.3 analyses the extent to which distortions might occur between the sample and the population due to over or under representation of farms with specific characteristics; for example due to non-response in relation to factors explaining the non-response and the applied weighting methodology. Section 6.3.4 provides information on the reliability of estimates.

6.3.2 Coverage

It is desirable to have a sample that represents the population as well as possible. A clear distinction should be made between the coverage and the representativeness. This section describes the coverage, section 6.3.3 deals with the representativeness. To get an idea about the extent to which the total population is covered by the sample it is relevant to distinguish several aspects. Farms that are too small or are not registered in time are not part of the agricultural census (b). The sampling frame (c) is the basis for the choice of sample farms and consists of farms registered in the agricultural census and have a size of more then 16 esu and less then 1,200 esu. From this sampling frame the sample is drawn (d).

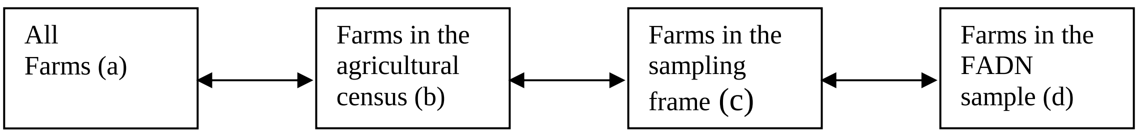


Figure 6.1 Relationship between FADN sample and all farms

Table 6.2 gives an indication to what extent the FADN sample covers the whole population. Therefore a comparison is made between the farms in the sampling framework (all the farms that have a chance of being included in the FADN sample) (c) and the total population as described by the agricultural census (b). Direct comparison with all farms (a) would be better but the unregistered farms are unknown, and the practical difference is very limited. The sampling framework covers the population to a large extent. For example with respect to the production, more than 91% is covered by the sample. Small farms are excluded from the sampling framework, this means that a substantial number of the farms and to a lesser extent also of labor are outside of the sampling frame. With respect to agricultural activities, the table shows that some activities are not well covered by the sample. This mainly concerns the activities that are commonly found on very small or on very large specialised farms.

In policy analysis and research it is essential to distinguish between farming types (for example specialised pig fattening farms) and agricultural activities (pig fattening). In the report on the redesign of the FADN sample it was illustrated that types of farming should not be the only focus of research (Vrolijk and Lodder, 2002). Agricultural activities are important in many research projects.

To give a complete picture of a certain agricultural activity it is therefore important to look at the activities on all farm types (table 6.3). For example, not only pig fattening farms will create added value from pig fattening, also other types of farms can be involved in this activity (although it is not their main business). The next table describes to which extent a certain activity can be found on certain types of farming. The figures in italic express that an activity belongs to that type of farming (based on the principal types of farming). For example, 83.2% of the agricultural activity fattening pigs can be found on the intensive livestock farms. This means that 16.8% of this activity can be found on farms that belong to other types of farming, for example arable farms. Looking in more detailed, the skewness is even larger. Type of farming 5011, the specialised pig fattening farms are responsible for 54.7% of the pig fattening activity. This implies that 45.3% of this activity takes place within other types. Production of mushrooms is a very specialised agricultural activity. More than 99% of this activity takes place on specialised mushroom farms.

Table 6.2 Coverage of the sample compared to agricultutal census (2005)

Variable-agricultural census	Number according to census	Not covered in sample (%)		Percentage covered by sample
		of which <16 esu	of which >1,200 esu	
Numbers				
Farms	81830	23.2	0.5	76.4
Farm managers	95682	11.3	0.7	88.0
Family labour	114619	11.1	0.6	88.2
Paid labour	45053	2.7	16.3	81.0
Total labour	159673	8.8	5.1	86.2
Size in hectares				
Agricultural area	1920811	5.4	1.3	93.3
Arable	823493	4.7	1.1	94.1
Grassland	980359	6.4	0.5	93.1
Horticulture under glass	10540	0.1	20.6	79.3
Vegetables in the open air	100964	1.3	7.9	90.8
Other agriculultural area	1593	9.7	5.1	85.2

Table 6.2 Coverage of the sample compared to agricultural census (2005) (continued)

Variable-agricultural census	Number according to census	Not covered in sample (%)		Percentage covered by sample
		of which <16 esu	of which >1,200 esu	
Number of animals				
Dairy cows	1433202	0.1	0.1	99.7
Fattening calves	828740	1.1	1.3	97.6
Ewes	676877	21.2	0.0	78.8
Fattening pigs	5504295	1.4	0.8	97.9
Breeding pigs	1244272	0.2	0.8	99.0
Laying hens	41047700	0.4	3.7	95.9
Poultry	44496116	0.1	1.8	98.1
Size in hectares				
Winter cereal	116040	4.5	1.1	94.4
Seed potatoes	39262	0.2	1.5	98.3
Consumption potatoes	65830	1.5	1.1	97.4
Starch potatoes	50692	0.9	1.9	97.2
Sugar beets	91313	2.7	1.0	96.2
Peas for canning	5091	1.8	5.3	93.0
Seed onions	16778	0.5	1.2	98.3
Grass seed	27639	2.8	1.3	95.9
Green maize	235088	7.0	0.2	92.7
Celeriac	1128	1.9	2.3	95.8
Brussel sprouts	3095	0.3	0.0	99.7
Cabbage all types	4867	1.2	0.6	98.3
Carrots	2551	2.0	3.2	94.8
Winter carrot	4700	0.5	4.6	94.9
Chicory	3423	0.2	0.1	99.7
Asparagus	2334	2.7	2.4	94.9
Horticultural sees	748	8.1	11.0	80.9
Tulips	10551	0.2	14.1	85.7
Hedges	2640	2.6	2.0	95.5
Trees	4992	0.9	18.5	80.6
Apples	9737	1.7	0.1	98.2
Pears	6692	1.8	0.0	98.2
Tomatoes under glass	1396	0.0	50.2	49.8
Cucumbers under glass	631	0.0	6.3	93.7
Paprika under glass	1236	0.0	24.1	75.9
Roses	780	0.0	24.1	75.9
Chrysanthemum	598	0.1	5.7	94.3
Fresia	167	0.0	2.9	97.1
Ornamentals leave	589	0.1	29.8	70.2
Ornamentals flower	788	0.0	29.9	70.0
Mushrooms	77	0.0	24.9	74.9

Table 6.3 Relationship between types of farming and agricultural activities - share of esu (farms between 16 and 1,200 esu) 2005

Type of farming	Dairy	Cattle	Sheep	Goat	Grassland	Fattening pig	Other pig	Laying hen	Poultry
Field crop farms									
- starch potatoes	0.00	0.31	0.15	0.02	0.22	0.00	0.21	0.13	0.79
- organic crops	0.01	0.13	0.09	0.01	0.25	0.00	0.05	0.07	0.09
- other field crop farms	0.07	2.59	3.60	0.08	4.35	0.15	0.94	0.75	2.35
Horticulture									
Vegetables under glass									
- paprika	0.00	0.00	0.02	0.00	0.01	0.00	0.00	0.00	0.00
- cucumber	0.00	0.00	0.01	0.00	0.02	0.01	0.00	0.00	0.00
- tomato	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00
- other	0.00	0.00	0.07	0.01	0.06	0.00	0.03	0.04	0.00
Cut flowers under glass									
- rose	0.00	0.00	0.04	0.01	0.02	0.00	0.00	0.00	0.00
- chrysanthemum	0.00	0.00	0.01	0.00	0.01	0.00	0.00	0.00	0.00
- other	0.00	0.00	0.11	0.01	0.09	0.00	0.01	0.00	0.00
Plants	0.00	0.00	0.05	0.01	0.09	0.00	0.00	0.00	0.00
Other glass	0.00	0.00	0.04	0.01	0.04	0.00	0.01	0.00	0.00
Field vegetables	0.01	0.07	0.09	0.00	0.15	0.03	0.09	0.03	0.11
Fruit	0.00	0.07	0.14	0.01	0.24	0.02	0.07	0.15	0.00
Nurseries	0.03	0.34	0.12	0.01	0.25	0.14	0.28	0.08	0.00
Mushroom	0.00	0.01	0.02	0.00	0.07	0.00	0.02	0.00	0.00
Bulbs	0.05	0.07	0.20	0.33	0.14	0.02	0.16	0.01	0.24
Other open air	0.01	0.14	0.10	0.03	0.19	0.05	0.19	0.09	0.27
Grazing livestock									
Dairy									
- organic	1.27	0.82	0.58	0.12	0.46	0.05	0.08	0.16	0.00
- non-organic	92.61	51.58	23.12	1.52	4.00	2.27	9.18	1.34	1.29
Calf fattening	0.02	0.57	0.53	0.07	0.05	0.00	0.28	0.15	0.13
Other grazing livestock	1.28	27.25	59.14	90.15	73.37	0.30	1.12	0.48	0.20

Table 6.3 Relationship between types of farming and agricultural activities - share of esu (farms between 16 and 1,200 esu) 2005 (continued)

Type of farming	Dairy	Cattle	Sheep	Goat	Grassland	Fattening pig	Other pig	Laying hen	Poultry
Intensive livestock									
Fattening pigs	0.04	0.40	0.92	0.10	1.37	54.70	3.84	0.19	0.14
Breeding pigs	0.00	0.29	0.59	0.03	1.02	0.38	35.08	0.03	0.00
Integrated pig farms	0.03	0.40	0.60	0.17	0.87	26.74	26.45	0.10	0.51
Laying hen	0.02	0.09	0.42	0.01	0.65	0.07	0.32	79.00	0.04
Poultry	0.01	0.06	0.24	0.02	0.29	0.03	0.19	0.11	66.85
Other intensive livestock	0.03	0.12	0.17	0.25	0.24	1.25	1.56	4.80	4.73
Mixed	4.49	14.68	8.83	7.03	11.50	13.78	19.82	12.28	22.25
Total	100	100	100	100	100	100	100	100	100

Table 6.3 (continued) Relationship between types of farming and agricultural activities - share of esu (farms between 16 and 1,200 esu) 2005

	Wheat	Root crops	Vegetable open air	Fruit	Tree	Mushroom	Bulbs	Vegetables glass	Cut flowers glass	Ornamentals glass
Field crop farms										
- Starch potatoes	6.29	14.20	0.41	0.06	0.00	0.00	0.00	0.00	0.00	0.00
- organic crops	1.87	1.05	5.34	0.33	0.07	0.00	0.08	0.01	0.00	0.00
- other field crop farms	55.19	61.50	69.29	0.82	0.11	0.00	0.62	0.01	0.01	0.00
Horticulture										
Vegetables under glass										
- paprika	0.02	0.00	0.01	0.03	0.02	0.00	0.00	28.07	0.09	0.09
- cucumber	0.03	0.00	0.00	0.00	0.00	0.00	0.00	10.45	0.03	0.00
- tomato	0.01	0.01	0.02	0.00	0.01	0.00	0.02	35.31	0.03	0.00
- other	0.18	0.01	0.28	0.27	0.04	0.00	0.00	22.48	0.22	0.00
Cut flowers under glass										
- rose	0.01	0.00	0.02	0.00	0.09	0.00	0.00	0.02	29.49	0.07
- chrysanthemum	0.00	0.00	0.00	0.09	0.00	0.00	0.00	0.02	9.74	0.07
- other	0.13	0.06	0.09	0.05	0.11	0.00	1.20	0.14	48.52	0.23

Table 6.3 (continued) Relationship between types of farming and agricultural activities - share of esu (farms between 16 and 1,200 esu) 2005 (continued)

	Wheat	Root crops	Vegetable open air	Fruit	Tree	Mushroom	Bulbs	Vegetables glass	Cut flowers glass	Ornamentals glass
Plants	0.03	0.00	0.00	0.05	0.23	0.00	0.08	0.15	0.71	95.78
Other glass	0.07	0.02	0.09	0.37	2.90	0.00	3.49	1.84	5.79	2.69
Field vegetables	0.54	0.42	1.77	0.21	0.06	0.00	0.02	0.31	0.01	0.00
Fruit	0.38	0.20	0.24	84.82	0.08	0.00	0.00	0.03	0.00	0.00
Nurseries	0.47	0.18	0.20	0.30	84.88	0.00	0.13	0.02	0.07	0.03
Mushroom	0.03	0.00	0.00	0.13	0.00	99.28	0.00	0.00	0.00	0.00
Bulbs	0.87	0.98	1.50	0.01	0.06	0.00	76.30	0.01	1.17	0.00
Other open air	0.66	0.52	0.75	1.87	4.36	0.03	10.03	0.61	3.39	0.16
Grazing livestock										
Dairy										
- organic	0.24	0.01	0.03	0.01	0.00	0.00	0.00	0.00	0.00	0.00
- non-organic	3.94	3.95	2.34	0.48	0.28	0.01	0.24	0.01	0.00	0.00
Calf fattening	0.22	0.07	0.11	0.02	0.04	0.00	0.00	0.00	0.00	0.00
Other grazing livestock	4.07	0.57	0.18	0.35	0.04	0.00	0.03	0.00	0.00	0.00
Intensive livestock										
Fattening pigs	1.97	0.36	0.37	0.03	0.05	0.00	0.00	0.00	0.00	0.00
Breeding pigs	1.34	0.20	0.19	0.02	0.05	0.00	0.00	0.00	0.00	0.00
Integrated pig farms	2.07	0.64	0.64	0.05	0.08	0.00	0.00	0.00	0.00	0.00
Laying hen	0.56	0.17	0.12	0.01	0.04	0.00	0.01	0.00	0.00	0.00
Poultry	0.38	0.13	0.03	0.01	0.02	0.00	0.00	0.00	0.00	0.00
Other intensive livestock	0.29	0.09	0.06	0.01	0.01	0.00	0.00	0.00	0.00	0.00
Mixed	18.14	14.64	15.93	9.61	6.37	0.68	7.74	0.51	0.73	0.90
Total	100	100	100	100	100	100	100	100	100	100

6.3.3 Representativeness

Because of the stratification scheme the sample will provide a good representation of the population on the main characteristics (stratification variables) at the beginning of a year. During the year farms might drop out of the sample and changes might occur in the population. Despite these changes the representativeness is maintained by applying post-stratification on the resulting sample and the changed population. Representativeness with respect to the stratification variables does not necessary imply that the sample is representative for all variables. Such a full representativeness is impossible unless the sample size approximates the whole population. Table 6.4 shows to what extent the sample is representative for a number of variables in the agricultural census.

The following guideline can help in the interpretation of the table: a relative difference which is close to the relative standard error can not be regarded as proof of systematic differences between the sample and the population. If the relative difference is more than two times the relative standard error then it is less likely that these differences can be explained by sampling errors. It is very unlikely that the difference is caused by coincidence if the relative difference is more than 3 times the relative standard error.

An example can illustrate how the table should be interpreted. The average number of dsu (dutch size units) of pigs as measured in the agricultural census 2005 is 7.51 (i.e. the average of all farms within the field of observation). If the same variable is estimated based on the FADN sample an average of 7.99 is calculated. It might seem that the number of pigs is slightly overestimated in the sample. However, the relative standard error of the estimate is 3.3%. When this standard error is compared to the relative difference between both values (6%) than the conclusion, that there is a significant difference, cannot be supported.

Table 6.4 Comparison of farms in the agricultural census (16-1,200 esu) and farms in the Dutch FADN (agricultural census 2005)

Variable	Average calculated based on		Relative standard error (FADN)	Ratio Census and FADN		
				all farms	farms with value > 0	
	census (1)	FADN (2)		average 1/2	number	average per farm
Size						
dsu	102.02	107.57	0.8	94.8	100.0	94.8
Activities (dsu)						
Field crops	12.00	13.05	2.9	91.9	89.5	102.7
Grassland	1.90	1.91	13.0	99.2	101.4	97.8
Fallow land	0.00	0.00	28.3	231.7	105.6	219.4
Horticulture in the open	14.76	15.43	3.2	95.7	98.5	97.1
Horticulture under glass	22.77	24.07	2.0	94.6	98.0	96.6
Cattle	34.04	36.04	1.6	94.4	98.4	95.9

Source: Agricultural census 2005.

Table 6.4 Comparison of farms in the agricultural census (12-1,200 esu) and farms in the Dutch FADN (agricultural census 2005) (continued)

Variable	Average calculated based on		Relative standard error (FADN)	Ratio Census and FADN		
				all farms	farms with value > 0	
	census (1)	FADN (2)		average 1/2	number	average per farm
Dairy cows	27.61	29.21	1.7	94.5	93.8	100.7
Fattening cattle	0.99	1.25	17.8	79.2	96.8	81.7
Veal	1.67	1.73	7.7	96.4	110.6	87.1
Horses	2.47	0.84	19.4	293.1	137.4	213.4
Sheep	0.47	0.73	24.2	63.5	93.1	68.2
Goats	0.30	1.09	30.1	27.5	80.5	34.1
Pigs	7.51	7.99	3.3	94.0	98.4	95.6
Fattening pigs	3.39	3.35	4.8	101.1	105.7	95.7
Breeding pigs	4.11	4.58	4.4	89.7	93.2	96.2
Poultry	3.11	3.59	6.7	86.7	81.0	107.0
Fattening peepers	0.95	0.97	13.3	97.5	100.0	97.5
Laying hen	1.47	2.03	8.8	72.5	68.1	106.4
Dugs	0.05	0.02				
Turkey	0.11	0.21	32.2	53.3	73.5	72.5
Rabbits	0.04	0.05	49.6	75.7	172.3	43.9
Fur animals	0.50	0.37	31.9	135.6	192.1	70.6
Sizes (ha)						
UAA	28.69	30.33	1.7	94.6	99.6	94.9
Field crops	12.41	13.62	2.8	91.1	91.8	99.2
Horticulture open air	1.47	1.59	4.7	92.4	98.5	93.8
Horticulture glass	0.13	0.14	2.3	92.9	98.0	94.9
Permanent grass	11.43	11.61	4.3	98.5	101.4	97.1
Temporary grassland	3.18	3.34	7.6	95.2	91.5	104.1
Fallow	0.02	0.01	28.3	231.7	105.6	219.4
Other	6.21	1.35	12.1	461.3	96.0	480.7
Forest	0.50	0.02	38.1	2,630.6	141.1	1,863.8
Acreages field crops						
Grains	3.18	3.40	5.8	93.6	89.6	104.5
Leguminous plants	0.06	0.05	28.9	105.0	88.3	118.8
Commercial crops	0.11	0.08	28.3	137.1	127.3	107.7
Seeds	0.42	0.58	13.4	73.0	65.5	111.5
Tuberous and carrots	3.91	4.28	3.6	91.6	87.1	105.1
Green fodder	3.66	3.91	4.7	93.7	93.7	99.9
Green fertilizer	0.47	0.79	19.1	59.7	75.0	79.6

Source: Agricultural census 2005.

Table 6.4 Comparison of farms in the agricultural census (16-1,200 esu) and farms in the Dutch FADN (agricultural census 2005) (continued)

Variable	Average calculated based on		Relative standard error (FADN)	Ratio Census and FADN		
				all farms	farms with value > 0	
	census (1)	FADN (2)		average 1/2	number	average per farm
Horticulture in the open air						
Vegetables (market garden)	0.63	0.72	9.1	87.5	84.4	103.7
Vegetables (field scale)	0.34	0.37	14.9	93.5	93.3	100.2
Stone fruit	0.27	0.28	8.2	95.2	105.9	89.9
Small fruits	0.02	0.04	47.5	53.3	63.5	84.0
Flower nursery	0.04	0.04	22.2	102.7	102.7	100.0
Tree nursery	0.18	0.15	13.6	114.3	119.9	95.3
Flower bulbs	0.30	0.34	5.9	87.9	91.4	96.2
Glas houses						
Vegetables	0.05	0.05	4.1	93.1	90.8	102.5
Tomatoes	0.01	0.01	6.1	95.3	68.8	138.6
Cucumbers	0.01	0.01	7.8	90.0	87.4	102.9
Paprika	0.02	0.02	4.6	90.6	74.5	121.5
Fruit	0.00	0.00	57.5	34.3	37.5	91.6
Cut flowers	0.05	0.06	3.0	86.6	99.5	87.1
Roses	0.01	0.01	5.7	93.6	99.9	93.7
Chrysanthemum	0.01	0.01	9.1	86.7	69.7	124.5
Plants	0.02	0.02	6.7	96.7	115.7	83.6
Tree nursery	0.01	0.01	27.7	93.5	107.8	86.7
Flat glass	0.00	0.00				
Standing glass	0.13	0.14	2.3	93.8	99.4	94.3
Mushrooms						
Cell	0.03	0.03	8.4	105.1	103.5	101.5
Size (are)	0.00	0.00	10.3	91.3	103.5	88.2
Chicory						
Size (are)	0.03	0.07	37.3	49.7	54.3	91.5
Bulbs						
Tulips (pieces)	18.62	21.83	16.0	85.3	88.7	96.2
Narcissus (kg)	0.05	0.04	54.1	133.2	116.7	114.1
Substrate growing (are)						
Vegetable	0.03	0.04	6.0	87.6	82.9	105.7
Flowers	0.01	0.02	12.0	76.7	80.4	95.4

Source: Agricultural census 2005.

Table 6.4 Comparison of farms in the agricultural census (16-1,200 esu) and farms in the Dutch FADN (agricultural census 2005) (continued)

Variable	Average calculated based on		Relative standard error (FADN)	Ratio Census and FADN		
				all farms	farms with value > 0	
	census (1)	FADN (2)		average 1/2	number	average per farm
Stable capacity (number of animals)						
Fattening calves	15.79	20.82	17.5	75.9	97.5	77.8
Fattening pigs	109.60	110.15	5.0	99.5	106.1	93.8
Peepers	826.54	882.39	15.2	93.7	100.4	93.3
Laying hen	605.38	1038.48	11.5	58.3	62.8	92.8
Characteristics firm and entrepreneur						
Main occupation (%)	1.13	1.08	1.1	105.0	100.0	105.0
Legal entity (%)	5.36	2.50	12.8	214.6	214.7	100.0
Age	52.24	50.35	0.8	103.8	100.0	103.7
Labour						
Total	3.64	3.68	3.9	98.9	100.0	98.9
Male	2.32	2.23	2.7	103.6	99.1	104.6
Female	1.32	1.44	7.3	91.6	94.0	97.4
Paid labour	1.42	1.36	10.3	104.8	100.6	104.2

Source: Agricultural census 2005.

The information in table 6.4 gives an indication for which variables and thus for which research projects it might be wise to perform post-stratification or use alternative estimation techniques to take into account the differences between the sample and the population. For example, in studies in which the age of the farmer plays an important role it might be useful to apply alternative estimation techniques.

The last two columns of table 6.4 provide more detailed information on the difference between the population and the sample. These differences can be explained on one hand by differences in the number of farms on which a certain activity occurs (a value larger than zero) and on the other by the average of this activity on farms which are in this activity. For example: the number of dsu dairy cows in the FADN is higher than in the agricultural census. This difference is partly explained by a higher estimation of the number of farms with dairy cows and partly by a 0.6% lower estimation of esu of dairy cows on farms with dairy cows (94.8 = 93.8% * 100.7).

A comparison between the sample and the population as registered in the agricultural census does not fully answer the question whether estimations of financial, economic and technical characteristics are bias free. It is for example possible that farms with relatively good or bad management skills and therefore performance are over represented in the sample.

6.3.4 Reliability

The previous subsection provides some indicators whether there are systematic differences between the sample and the population (representativeness of sample). This section focuses on the reliability of the estimates.

The calculation of averages of groups based on sampling units implies that there can be differences between the estimated value and true population value. These differences can occur due to the random selection of units to be included in the sample. Table 6.5 provides an indication of the level of precision of the estimates for a set of important goal variables.

The precision of estimates can be measured by the standard error of the estimate of a variable. The standard error is used to calculate the confidence interval. This confidence interval describes the range in which the true population value will be given a certain level of certainty. The confidence interval ranges from the calculated average minus two times the standard error to the calculated average plus two times the standard error. The calculated averages of two groups are significantly different (with a 95% certainty) if the difference is larger than two times the square root of the sum of squares of the standard errors of the two group averages.

This section provides the reliability of estimates for a number of important goal variables for different types of farming. This calculation is based on the available CSP observations (see section 5.3).

Table 6.5 Reliability of estimates of important goal variables per type of farming, based on FADN sample (2005)

Type of farming	Goal variable					
	family farm income	total revenues	return a)	savings	income farm	net farm result
Field crop farms						
- starch potatoes	6,895	25,241	2.2	84,411	8,329	4,869
- organic crops	30,232	90,294	4.5	31,203	32,692	35,094
- other field crop farms	4,525	17,495	2.8	6,972	4,319	4,387
Horticulture						
Vegetables under glass						
- paprika	21,961	34,776	1.8	24,338	25,182	19,228
- cucumber	23,802	51,911	2.6	22,145	23,992	20,411
- tomato	59,559	106,505	3.4	55,028	55,487	44,649
- other	17,969	49,188	4.1	17,073	18,167	17,841
Cut flowers under glass						
- rose	42,258	117,639	3.8	38,299	41,599	41,354
- chrysanthemum	27,023	204,871	2.8	18,657	27,032	28,377
- other	13,892	52,846	2.6	11,701	14,627	13,109
Plants	21,585	95,200	3.7	16,904	21,437	22,088
Other glass	6,209	29,803	4.5	5,942	6,626	13,041
Field vegetables	34,952	83,872	13.9	25,680	32,742	32,004
Fruit	8,881	18,238	4.2	13,836	9,493	9,902

a) Revenues per 100 euro costs; *insufficient number of observation in CSP variant.

Table 6.5 Reliability of estimates of important goal variables per type of farming, based on FADN sample (2005) (continued)

Type of farming	Goal variable					
	family farm income	total revenues	return a)	savings	income farm	net farm result
Nurseries	*	*	*	*	*	*
Mushroom	*	*	*	*	*	*
Bulbs	81,818	258,265	5.7	79,164	81,519	73,805
Other open air	28,271	58,666	6.4	31,483	29,029	25,343
Grazing livestock						
Dairy						
- organic	4,934	7,895	2.6	4,886	4,694	5,616
- non-organic	2,380	4,504	0.8	2,581	2,680	2,093
Calf fattening	8,696	18,308	3.4	9,432	8,159	6,921
Other grazing livestock	27,561	30,822	8.4	33,882	32,390	28,394
Intensive livestock						
Breeding pigs	10,045	32,400	2.1	9,330	10,040	8,920
Fattening pigs	13,778	71,697	3.3	13,292	13,475	8,956
Integrated pig farms	13,724	31,158	1.7	13,032	15,085	9,792
Laying hen	14,151	33,620	2.5	12,862	13,418	11,610
Poultry	10,384	29,847	2.2	9,905	10,320	12,049
Other intensive livestock	*	*	*	*	*	*
Mixed	5,536	19,514	2.2	6,301	6,699	5,277

a) Revenues per 100 euro costs; *insufficient number of observation in CSP variant.

Table 6.6 Reliability of estimates of important goal variables per main type of farming, base d on FADN sample (2005)

Type of farming	Goal variable					
	family farm income	total revenues	return	savings	income farm	net farm result
Field crops	4,018	15,359	2.4	12,131	3,911	3,890
Vegetables under glass	14,293	30,561	2.0	13,696	14,176	12,245
Cut flowers under glass	12,622	47,848	2.1	10,798	13,018	12,099
Pigs	7,213	29,546	1.5	6,861	7,338	5,317
Poultry	9,961	24,355	1.8	9,112	9,504	8,678
Grazing livestock	7,289	8,553	2.2	8,908	8,545	7,454
All farms	4,028	7,624	1.2	5,171	4,561	3,996

There are clear differences in the significance of estimates between different types of farming. The estimates for the dairy sector are the most reliable because of the large number of farms included in the sample, which reflects the importance of the dairy sector in Dutch agriculture. The decision on the number of farms is described in Vrolijk and Lodder (2002).

Table 6.7 Coefficient of variation of estimates of important goal variables per main type of farming, based on FADN sample (2005)

Type of farming	Goal variable					
	family farm income	total revenues	return	savings	income farm	net farm result
Field crops	0.112	0.075	0.029	-1.300	0.081	-0.162
Vegetables under glass	0.497	0.045	0.023	-0.426	0.394	-0.181
Cut flowers under glass	0.281	0.066	0.024	-0.430	0.255	-0.204
Pigs	0.082	0.059	0.015	0.130	0.074	0.398
Poultry	0.430	0.045	0.020	-0.788	0.282	-0.199
Grazing livestock	0.124	0.042	0.029	0.324	0.116	-0.201

Table 6.8 Coefficient of variation of estimates of important goal variables per type of farming, based on FADN sample (2005)

Type of farming	Goal variable					
	family farm income	total revenues	return	savings	income farm	net farm result
Field crop farms						
- starch potatoes	0.16	0.12	0.03	-1.15	0.14	-0.23
- organic crops	0.69	0.26	0.05	1.47	0.46	-3.33
- other field crop farms	0.13	0.09	0.03	-7.80	0.09	-0.18
Horticulture						
Vegetables under glass						
- paprika	-0.33	0.05	0.02	-0.18	-0.48	-0.12
- cucumber	0.19	0.05	0.03	0.61	0.19	0.56
- tomato	-6.59	0.10	0.04	-0.67	-19.45	-0.50
- other	0.30	0.11	0.05	1.04	0.27	-0.41
Cut flowers under glass						
- rose	-0.90	0.10	0.04	-0.35	-0.98	-0.36
- chrysanthemum	0.18	0.15	0.03	0.30	0.18	0.43
- other	0.27	0.09	0.03	-0.64	0.25	-0.21
Plants	0.33	0.13	0.04	1.28	0.29	-1.83
Other glass	1.92	0.12	0.08	-0.48	0.27	-0.29
Field vegetables	0.69	0.24	0.19	-4.62	0.56	-0.94
Fruit	1.31	0.09	0.06	-1.12	0.62	-0.14
Nurseries	*	*	*	*	*	*
Mushroom	*	*	*	*	*	*
Bulbs	0.53	0.34	0.06	0.75	0.51	1.31
Other open air	0.40	0.22	0.08	0.94	0.39	-1.38
Grazing livestock						
Dairy						
- organic	0.10	0.04	0.03	0.28	0.08	-0.14
- non-organic	0.04	0.02	0.01	0.07	0.04	-0.05
Calf fattening	0.15	0.12	0.04	0.68	0.13	-1.02
Other grazing livestock	0.57	0.20	0.12	4.45	0.47	-1.15

Table 6.8 Coefficient of variation of estimates of important goal variables per type of farming, based on FADN sample (2005) (continued)

Type of farming	Goal variable					
	family farm income	total revenues	return	savings	income farm	net farm result
Breeding pigs	0.10	0.06	0.02	0.14	0.09	0.40
Fattening pigs	0.27	0.18	0.03	0.46	0.22	-1.90
Integrated pig farms	0.12	0.05	0.02	0.20	0.12	0.40
Laying hen	-6.23	0.07	0.03	-0.32	1.54	-0.17
Poultry	0.14	0.04	0.02	0.23	0.13	7.21
Other intensive livestock	*	*	*	*	*	*
Mixed	0.13	0.08	0.03	0.34	0.11	-0.14

Tables 6.7 and 6.8 describe the relative standard error (coefficient of variance). This is the standard error divided by the group average. A higher relative standard error implies less reliable estimates, but the value is strongly affected by the absolute value of the average. If the average value approaches zero, the relative standard error can become very large. A meaningful evaluation of the standard error requires a simultaneous use of tables 6.5 and 6.6 on one hand and tables 6.7 and 6.8 on the other.

References

Bont, C.J.A.M. de, W.H. van Everdingen and B. Koole, *Standard Gross Margins in the Netherlands*. LEI, Den Haag, 2003.

Dijk, J., *De steekproef gewogen*. Onderzoeksverslag 53. LEI, Den Haag, 1989.

Dijk, J., B. Kortekaas, K. Lodder and J. Luijt, *Netjes over de drempel; voorstel tot harmonisatie van de steekproeven voor de boekhoudnetten land-, tuin- en bosbouw*. Interne Nota 422. LEI-DLO, Den Haag, 1994.

Dijk, J., K. Lodder, J. Luyt and H.C. Pruis, *Voorstel voor de indeling van de populatie land- en tuinbouwbedrijven in groepen en bijbehorend keuzeplan*. Interne Nota 437. LEI-DLO, Den Haag, 1995a.

Dijk, J., J. Schering and K. Lodder, *Improvement of the weighting system of the FADN*. LEI-DLO, Den Haag, 1995b.

Dijk, J., *Nonresponse in the Dutch Farm Account Data Network*. Paper presented at the Essex 1996 Fourth International Conference on Social Science Methodology, 1-5 July 1996. LEI-DLO, Den Haag, 1996.

Dijk van, J.P.M., B.E. Douma and A.L.J. van Vliet, *Bedrijfsuitkomsten in de landbouw (BUL) Boekjaren 1992/93 t/m 1995/96*. Periodieke Rapportage 11-95/96. LEI-DLO, Den Haag, 1997a.

Dijk van, J.P.M., B.E. Douma and A.L.J. van Vliet, *De financiële positie van de landbouw (FIP) Boekjaar 1995/96 en vergelijkingen met voorgaande jaren*. Periodieke Rapportage 12-95/96. LEI-DLO, Den Haag, 1997b.

Dijk, van J.P.M., J.J.P. Groot, K. Lodder and H.C.J. Vrolijk, *De steekproef voor het Bedrijven-Informatienet van het LEI; Bedrijfskeuze 1999 en selectieplan 2000*. Report 6.00.94. LEI, Den Haag, 2000.

Dijk, van J.P.M., J.J.P. Groot, K. Lodder, L.C. van Staalduinen and H.C.J. Vrolijk, *De steekproef voor het Bedrijven-Informatienet van het LEI; Bedrijfskeuze 1998 en selectieplan 1999*. Report 6.99.94. LEI, Den Haag, 1999.

Dijk, J.P.M. van, B.E. Douma and A.L.J. van Vliet, *BUL: Bedrijfsuitkomsten in de landbouw (bedrijfsresultaten (verlies- en winstrekeningen) voor akkerbouw- en veehouderijbedrijven*. PR 11. LEI-DLO, Den Haag, 1996.

Dijk, van J.P.M., K. Lodder and H.C.J. Vrolijk, *De steekproef voor het Bedrijven-Informatienet van het LEI; Bedrijfskeuze 2000 en selectieplan 2001*. Report 1.01.02. LEI, Den Haag, 2002.

Poppe, K.J., *Het bedrijven informatienet van A tot Z*. Report 1.03.06. LEI, Den Haag, 2004.

Vrolijk, H.C.J., 'Working Procedures for the selection of farms in the FADN'. In: G. Beers and K. Poppe (eds.), *PACIOLI 9; Innovations in the FADN*. Report 8.02.02. LEI, Den Haag, 2002.

Vrolijk, H.C.J. and K. Lodder, *Voorstel tot vernieuwing van het steekproefplan voor het Bedrijven-Informatienet*. Report 1.02.02. LEI, Den Haag, 2002.

Vrolijk, H.C.J., W. Dol and G. Cotteleer, *Schatten van kenmerken van kleine deelgebieden*. LEI, Den Haag, 2002.

Vrolijk H.C.J. and G.C. Cotteleer, *Non-respons en rotatie in het Bedrijven-Informatienet; Kwantitatieve en kwalitatieve analyse van de effecten*. Report 1.05.01. LEI, Den Haag, 2005.

Vrolijk H.C.J., 'Non-response in the Dutch FADN; Qualitative reasons and quantitative impacts'. In: K.J. Poppe, *PACIOLI 13; Microeconomic Data on Farm Diversification, Rural Businesses and the Intra-generational Transfer* (2005) pp. 130-139. Report 8.05.04. LEI, Den Haag, 2005.

Vrolijk, H.C.J., H.B. van der Veen and J.P.M. van Dijk, *Sample of Dutch FADN 2004; design principles and quality of the sample of agricultural and horticultural holdings*. Report 1.06.03. LEI, Den Haag, 2006.

Vrolijk, H.C.J., 'Sampling of organic farms in the Dutch FADN: lessons learned'. In: M. Rippin, H. Willer, N. Lampkin, A. Vaughan, *Towards a European Framework for Organic Market Information* (2006) pp. 87-90.

Welten, J.P.P.J., *Berekening en toepassing van Nederlandse grootte-eenheden en standaardbedrijfseenheden* (nge 1992 en sbe 1993). Periodieke Rapportage 63-92. LEI, Den Haag, 1996.

Appendix 1. FADN: 2005 Selection Plan

FADN: 2005 Selection Plan

Country: The Netherlands
Region name: Region code:
Sub-region name: Sub-region code:

Type of farm		Economic size class				
1. National code	2. FADN Code 1999/725 (EC)	3. National code	4. FADN Code	5. Description (ESU)	6. No.of farms to be selected	7. No.of farms in the population
akkerbouw-bedrijven	1		7	16-40	34.7	2,744
			8	40-100	60.6	3,327
			9	100-250	79.4	2,073
			10	250 en meer	23.9	302
opengronds-groente-bedrijven	2011		7	16-40	11.2	252
			8	40-100	14.7	305
			9	100-250	16.6	223
			10	250 en meer	20.5	83
glasgroente-bedrijven	2012		7	16-40	3.1	123
			8	40-100	10.5	352
			9	100-250	37.5	501
			10	250 en meer	69.0	815

Type of farm		Economic size class				
1. National code	2. FADN Code 1999/725 (EC)	3. National code	4. FADN Code	5. Description (ESU)	6. No.of farms to be selected	7. No.of farms in the population
opengronds-bloem(bollen)-bedrijven	2021		7	16-40	5.1	284
			8	40-100	8.2	427
			9	100-250	13.7	469
			10	250 en meer	27.8	469
glasbloemen-bedrijven	2022		7	16-40	5.9	297
			8	40-100	13.6	657
			9	100-250	42.1	1,361
			10	250 en meer	74.5	1,546
champignon-bedrijven	2033		7	16-40	1.3	24
			8	40-100	4.4	81
			9	100-250	8.2	108
			10	250 en meer	16.2	82
(overig tuinbouw)	2013+2023+ 2039		7	16-40	2.5	131
			8	40-100	4.9	257
			9	100-250	7.3	242
			10	250 en meer	10.0	172
fruitbedrijven	321		7	16-40	6.8	362
			8	40-100	13.4	644
			9	100-250	17.3	436
			10	250 en meer	2.4	38
overige blijvende-teelt-bedrijven	340		7	16-40	7.2	525
			8	40-100	12.8	856
			9	100-250	14.7	660
			10	250 en meer	19.8	340
melkvee-bedrijven	4110+4120+4370		7	16-40	18.6	1,429
			8	40-100	122.0	9,107
			9	100-250	180.9	9,763
			10	250 en meer	18.5	504

Type of farm		Economic size class				
1. National code	2. FADN Code 1999/725 (EC)	3. National code	4. FADN Code	5. Description (ESU)	6. No.of farms to be selected	7. No.of farms in the population
kalvermesterij	4380		7	16-40	5.7	196
			8	40-100	9.4	411
			9	100-250	13.1	393
			10	250 en meer	1.8	30
overig rundvee- en grasland-bedrijven	other 43+44		7	16-40	15.3	4,703
			8	40-100	14.7	2,176
			9	100-250	15.0	596
			10	250 en meer	5.0	89
fokvarkens-bedrijven	5011		7	16-40	1.5	67
			8	40-100	14.7	668
			9	100-250	19.4	649
			10	250 en meer	14.5	115
vleesvarkens	5012		7	16-40	10.8	604
			8	40-100	15.8	590
			9	100-250	16.8	214
			10	250 en meer	6.6	35
gesloten varkens	5013		7	16-40	0.6	24
			8	40-100	8.1	329
			9	100-250	17.2	635
			10	250 en meer	14.2	107
legkippen-bedrijven	5021		7	16-40	1.8	106
			8	40-100	5.4	309
			9	100-250	10.6	356
			10	250 en meer	12.2	95
vleespluimvee	5022		7	16-40	2.2	41
			8	40-100	7.8	148
			9	100-250	14.2	226
			10	250 en meer	5.8	36

Appendix 2. Farm type classification Rules

(((<zetmeelaardappelbedrijven> als (landbouwtellingsbedrijf.zetmeelaardappelen [NGE] / Landbouwtellingsbedrijf.bedrijfsgrootte [NGE] > 0,33)

anders <biologische gewassen>
als (Landbouwtellingsbedrijf.biologisch [x1] = 1)

anders <akkerbouwbedrijven>) als Landbouwtellingsbedrijf.mei_neg [x1] < 2000

anders (<biologische melkveebedrijven> als Landbouwtellingsbedrijf.biologisch [x1] = 1

anders <melkveebedrijven>) als (Landbouwtellingsbedrijf.mei_neg [x1] = 4110
or (Landbouwtellingsbedrijf.mei_neg [x1] = 4120 or

anders <kalvermesterijbedrijven> als Landbouwtellingsbedrijf.mei_neg [x1] = 4380

anders <andere graasdierbedrijven> als (Landbouwtellingsbedrijf.mei_neg [x1] = 4390 or (Landbouwtellingsbedrijf.mei_neg [x1] = 4410 or (Landbouwtellings-bedrijf.mei_neg [x1] = 4420 or (Landbouwtellingsbedrijf.mei_neg [x1] = 4448 or (Landbouwtellingsbedrijf.mei_neg [x1] = 4449 or (Landbouwtellings-bedrijf.mei_neg [x1] = 4430)))))))

anders <fokvarkensbedrijven> als Landbouwtellingsbedrijf.mei_neg [x1] = 5011

anders <vleesvarkensbedrijven> als Landbouwtellingsbedrijf.mei_neg [x1] = 5012

anders <gesloten varkensbedrijven> als Landbouwtellingsbedrijf.mei_neg [x1] = 5013

anders <legkippenbedrijven> als Landbouwtellingsbedrijf.mei_neg [x1] = 5021

anders <vleespluimveebedrijven> als Landbouwtellingsbedrijf.mei_neg [x1] = 5022

anders <andere hokdierbedrijven> als (Landbouwtellingsbedrijf.mei_neg [x1] >= 5023 and (Landbouwtellingsbedrijf.mei_neg [x1] <= 5032))

anders <biologische gewassen> als (Landbouwtellingsbedrijf.mei_neg [x1] >= 6000 and (Landbouwtellingsbedrijf.mei_neg [x1] < 7000 and (Landbouwtellings-bedrijf.biologisch [x1] = 1))) anders <andere combinatiebedrijven>) als (Landbouwtellingsbedrijf.mei_neg [x1] < 2000 or (Landbouwtellingsbedrijf.mei_neg [x1] >= 4000))

anders ((<paprikabedrijven> als (Landbouwtellingsbedrijf.paprika [NGE] / Landbouwtellingsbedrijf.bedrijfsgrootte [NGE] > 0,67)

anders <komkommerbedrijven> als (Landbouwtellingsbedrijf.komkommer [NGE] / Landbouwtellingsbedrijf.bedrijfsgrootte [NGE] > 0,67)

anders <tomatenbedrijven> als (Landbouwtellingsbedrijf.tomaten [NGE] / Landbouwtellingsbedrijf.bedrijfsgrootte [NGE] > 0,67)

anders <overige glasgroentebedrijven>) als Landbouwtellingsbedrijf.mei_neg [x1] = 2012

anders (<chrysantenbedrijven> als (Landbouwtellingsbedrijf.chrysanten [NGE] / Landbouwtellingsbedrijf.bedrijfsgrootte [NGE] > 0,67)

anders <rozenbedrijven> als (Landbouwtellingsbedrijf.rozen [NGE] / Landbouwtellingsbedrijf.bedrijfsgrootte [NGE] > 0,67)

anders <overige snijbloembedrijven>) als (Landbouwtellingsbedrijf.mei_neg [x1] = 2022 and (Landbouwtellingsbedrijf.snijbloemen [NGE] / Landbouwtellingsbedrijf.bedrijfsgrootte [NGE] > 0,67))

anders <plantenbedrijven> als (Landbouwtellingsbedrijf.mei_neg [x1] = 2022 and (Landbouwtellingsbedrijf.planten [NGE] / Landbouwtellingsbedrijf.bedrijfsgrootte [NGE] > 0,67))

anders (<biologische gewassen> als (Landbouwtellingsbedrijf.biologisch [x1] = 1)

anders <opengrondsgroentebedrijven>)
als (Landbouwtellingsbedrijf.mei_neg [x1] = 2011)

anders <fruitbedrijven> als (Landbouwtellingsbedrijf.mei_neg [x1] = 3210)

anders <boomkwekerijbedrijven> als (Landbouwtellingsbedrijf.mei_neg [x1] = 3480)

anders <paddestoelbedrijven> als (Landbouwtellingsbedrijf.mei_neg [x1] = 2033)

anders <bloembollenbedrijven> als (Landbouwtellingsbedrijf.mei_neg [x1] = 2021 and (Landbouwtellingsbedrijf.bol [NGE] / landbouwtellingsbedrijf.bedrijfsgrootte [NGE] > 0,67)) anders <overige opengrondsbedrijven> als (Landbouwtellingsbedrijf.glas [NGE] / Landbouwtellingsbedrijf.bedrijfsgrootte [NGE] <= 0,50)

anders <overige glasbedrijven>))

Appendix 3. Ratio estimation

A3.1 Introduction

Estimates of averages and totals of a population are usually based on the values of the variable as observed in the sample. Total milk production can be estimated by raising milk production as observed in sample to the population by means of the weights available in the Informatienet.

However, in some cases the use of an auxiliary variable (for example the number of cows) can result in more precise estimates. Milk production on a farm will be highly correlated with the number of dairy cows. This auxiliary variable, number of cows, can be used to make more reliable estimates of the total milk production. The reason why these indirect estimates can be more reliable is that the ratio of the goal and the auxiliary variable is more stable than the separate variables. The total milk production can strongly diverge from farm to farm. A direct estimate would consequently show a high variance. The ratio milk production per cow will be much more stable. The application of this ratio estimate requires the availability of data on the auxiliary variable. In this example the total number of dairy cows in the country should be known.

In other cases we are not interested in the estimation of a population total (e.g. total milk production) but it in the ratio itself. Examples of research variables, which should be considered as ratios, are the use of antibiotics per animal, the average yield per hectare, or the average milk production per cow.

A3.2 Calculating ratio estimates an their variances

Ratio estimate

Ratio estimates can be calculated based on a stratified or unstratified sample. In case of an unstratified sample, the ratio estimate is calculated as follows:

$$\hat{r} = \frac{\sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} x_i} \tag{1}$$

Where:

$\hat{r}$ = ratio estimate

y_i = value of variable y for observation i

x_i = value of variable x for observation i

In case of a stratified sample $\hat{r}$ is calculated with[1]:

$$\hat{r} = \frac{\sum_{h=1}^{H}\left(\frac{N_h}{n_h}\sum_{i=1}^{n_h} y_{hi}\right)}{\sum_{h=1}^{H}\left(\frac{N_h}{n_h}\sum_{i=1}^{n_h} x_{hi}\right)} \tag{2}$$

Where:
n_h = sample of stratum h
N_h = population of stratum h

Variance of ratio estimate

Subsequently the variance of the ratio estimate can be calculated. In case of an unstratified sample the variance is calculated in the following way:

$$\mathrm{var}(\hat{r}) = \frac{1}{n(n-1)\bar{x}^2}\left(\sum_{i=1}^{n} y_i^{\,2} - 2\hat{r}\sum_{i=1}^{n} y_i x_i + \hat{r}^2 \sum_{i=1}^{n} x_i^2\right) \tag{3}$$

Where:
x_i = variable x for element i
y_i = variable y for element i
$\hat{r}$ = ratio (y / x)
N = number of sample elements
N = number of population elements

Calculating the variance of the ratio estimate using the strata can be done with the following 2 equations:

$$\mathrm{var}(\hat{r}_h) = \frac{1}{n(n-1)\bar{x}_h^2}\left(\sum_{i=1}^{n} y_{hi}^{\,2} - 2\hat{r}_h\sum_{i=1}^{n} y_{hi} x_{hi} + \hat{r}_h^2 \sum_{i=1}^{n} x_{hi}^2\right) \tag{4}$$

$$\mathrm{var}_s(\hat{r}) = \sum_{h=1}^{H}\left(\frac{N_h}{N}\right)^2 \mathrm{var}(\hat{r}_h) \tag{5}$$

Where:
H = stratum
N_h = number of farms in population in stratum h
N = total number of farms in population

The standard error of $\hat{r}$ is:

---

[1] Note that $\frac{N_h}{n_h}$ equals the weight of stratum h.

s.e. $(\hat{r}) = \sqrt{\mathrm{var}(\hat{r})}$ (6)

And the confidence interval of 95% is calculated as $\hat{r} \pm 1.96 * s.e.(\hat{r})$.

A3.3 Calculating variance of difference of ratios

In case we are especially interested in the difference between the ratio estimates of two periods, the variance of the difference has to be calculated. First we show the calculation in case of an unstratified sample.

$$\mathrm{var}(\hat{r}_1 - \hat{r}_2) = \mathrm{var}(\hat{r}_1) + \mathrm{var}(\hat{r}_2) - 2\,\mathrm{cov}(\hat{r}_1 \hat{r}_2) \tag{7}$$

$$\mathrm{var}(\hat{r}_.) = \frac{1}{n(n-1)\bar{x}^2} \left(\sum_{i=1}^{n_t} y_{.i}{}^2 - 2\hat{r}_. \sum_{i=1}^{n_t} y_{.i} x_{.i} + \hat{r}_.^2 \sum_{i=1}^{n_t} x_{.i}^2\right) \tag{8}$$

$$\mathrm{cov}(\hat{r}_1 \hat{r}_2) = \frac{1}{n(n-1)\bar{x}_1 \bar{x}_2} * \sum_{i=1}^{n} (y_{1i} y_{2i} - \hat{r}_1 y_{2i} x_{1i} - \hat{r}_2 y_{1i} x_{2i} + \hat{r}_1 \hat{r}_2 x_{i1} x_{2i}) \tag{9}$$

Where:

$x_{.i}$ = variable x in period t for element i

$y_{.i}$ = variable y in period t for element i

$\hat{r}_.$ = ratio period t

n = number of elements in sample in period t

N = number of elements in population in period t

Calculating stratified variance

To calculate the stratified variance, we first calculate the unweighted variance per stratum.

$$\mathrm{var}(\hat{r}_{h1} - \hat{r}_{h2}) = \mathrm{var}(\hat{r}_{h1}) + \mathrm{var}(\hat{r}_{h2}) - 2\,\mathrm{cov}(\hat{r}_{h1} \hat{r}_{h2}) \tag{10}$$

$$\mathrm{var}(\hat{r}_{h.}) = \frac{1}{n(n-1)\bar{x}_{h.}^2} \left(\sum_{i=1}^{n} y_{h.i}{}^2 - 2\hat{r}_h \sum_{i=1}^{n} y_{h.i} x_{h.i} + \hat{r}_h^2 \sum_{i=1}^{n} x_{h.i}^2\right) \tag{11}$$

$$\mathrm{cov}(\hat{r}_{h1} \hat{r}_{h2}) = \frac{1}{n(n-1)\bar{x}_{h1} \bar{x}_{h2}} * \left(\sum_{i=1}^{n_h} y_{h1i} y_{h2i} - r_{h1} \sum_{i=1}^{n_h} y_{h2i} x_{h1i} - \hat{r}_{h2} \sum_{i=1}^{n_h} y_{h1i} x_{h2i} + \hat{r}_{h1} \hat{r}_{h2} \sum_{i=1}^{n_h} x_{h1i} x_{h2i}\right) \tag{12}$$

$$\text{var}_s(\hat{r}_1 - \hat{r}_2) = \sum_{h=1}^{H}\left(\frac{N_h}{N}\right)^2 \text{var}(\hat{r}_{h1} - \hat{r}_{h2}) \tag{13}$$

Where:
h = stratum
N_h = number of farms in population in stratum h
N = total number of farms in population

Significant differences
The standard error of the difference in $\hat{r}_1 - \hat{r}_2$ is:

$$\text{s.e.}(\hat{r}_1 - \hat{r}_2) = \sqrt{\text{var}(\hat{r}_1 - \hat{r}_2)} \tag{14}$$

We use the t-test with n-1 degrees of freedom to test whether the difference is significant. And the confidence interval of 95% is calculated as $(\hat{r}_1 - \hat{r}_2) \pm 1.96 * s.e.(\hat{r}_1 - \hat{r}_2)$.

A3.4 Example

A1.4.1 Unstratified Ratio estimate

Table A3.1 shows the data of the example for calculating the ratio estimate and its' variance. First of all we calculate the ratio r (equation 1).

$$\hat{r}_1 = \frac{\sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} x_i} = \frac{16,527}{2,438} = 6.78$$

The variance of this ratio estimate can be calculated in the following way (according to equation3):

$$\text{var}(\hat{r}_1) = \frac{1}{n(n-1)\bar{x}^2}\left(\sum_{i=1}^{n} y_i^{\,2} - 2\hat{r}_1\sum_{i=1}^{n} y_i x_i + \hat{r}_1^2 \sum_{i=1}^{n} x_i^2\right)$$
$$= \frac{1}{31(31-1)*78.65^2}(13,685,067 - 2*6.78*1,612,356 + 6.78^2*239,262)$$
$$= 0.490$$

This results in the standard error of the ratio estimate:

s.e.($\hat{r}_1$) = 0.700

Table A3.1 Data example ratio estimate for an unstratified sample a)

Farm	Number of animals period	Measurement of dose antibiotics period
1	116	185
2	200	1,000
3	88	460
4	120	567
5	115	450
6	100	520
7	125	1,100
8	130	1,600
9	95	1,200
10	145	1,200
11	58	120
12	50	265
13	70	125
14	55	170
15	90	280
16	60	412
17	71	225
18	58	380
19	72	450
20	55	430
21	54	340
22	62	625
23	88	900
24	92	870
25	80	1,200
26	42	110
27	27	128
28	20	130
29	23	125
30	40	470
31	37	490

a) Although the figures are realistic, they are no real data.

A3.4.1 Stratified Ratio estimate

For the illustration of the ratio estimate for a stratified sample and the test of significance of a difference over two periods the dataset is extended as shown in table A3.2.

Table A3.2 Data example ratio estimate for a stratified sample

Farm	Stratum	Number of animals period 1	Measurement of dose antibiotics period 1	Number of animals period 2	Measurement of dose antibiotics period 2
1	1	116	185	126	110
2	1	200	1,000	205	1,000
3	1	88	460	90	470
4	1	120	567	121	490
5	1	115	450	105	590
6	1	100	520	130	840
7	1	125	1,100	128	1,125
8	1	130	1,600	97	1,215
9	1	95	1,200	160	1,425
10	1	145	1,200	101	220
11	2	58	120	110	470
12	2	50	265	56	60
13	2	70	125	73	125
14	2	55	170	60	290
15	2	90	280	75	350
16	2	60	412	63	400
17	2	71	225	70	425
18	2	58	380	52	560
19	2	72	450	76	544
20	2	55	430	88	630
21	2	54	340	95	782
22	2	62	625	80	900
23	2	88	900	46	83
24	2	92	870	53	170
25	2	80	1,200	57	375
26	3	42	110	44	62
27	3	27	128	29	78
28	3	20	130	20	82
29	3	23	125	24	90
30	3	40	470	42	400
31	3	37	490	36	480

The stratified ratio estimate for period 1 is calculated using equation 2 (see table A3.3):

$$\hat{r}_1 = \frac{\sum_{h=1}^{H}\left(\frac{N_h}{n_h}\sum_{i=1}^{n_h} y_{hi}\right)}{\sum_{h=1}^{H}\left(\frac{N_h}{n_h}\sum_{i=1}^{n_h} x_{hi}\right)} = \frac{1{,}383{,}900}{201{,}000} = 6.89$$

Table A3.3 Stratified ratio estimate period 1

	Stratum 1	Stratum 2	*Stratum 3*
$\sum_{i=1}^{n_h} y_{hi}$	8,282	6,792	1,453
$\sum_{i=1}^{n_h} x_{hi}$	1,234	1,015	189
N_h	500	1,500	1,200
n_h	10	15	6
Ratio estimate	6.89		

In a similar way $\hat{r}_2$ can be calculated, the resulting value is 5.91.

To calculate the stratified variance and standard error of the ratio estimate, we use equations 4 and 5. The results are shown in table A3.4.

$$\text{var}(\hat{r}_h) = \frac{1}{n(n-1)\bar{x}_h^2}(\sum_{i=1}^{n} y_{hi}{}^2 - 2\hat{r}_h \sum_{i=1}^{n} y_{hi} x_{hi} + \hat{r}_h^2 \sum_{i=1}^{n} x_{hi}^2)$$

$$\text{var}_s(\hat{r}) = \sum_{h=1}^{H} \left(\frac{N_h}{N}\right)^2 \text{var}(\hat{r}_h)$$

Table A3.4 Sum of values of variables per stratum in period 1

Sum of values of:	Stratum 1	Stratum 2	Stratum 3
x_1^2	161,400	71,471	6,391
x_1*y_1	1,067,230	494,645	50,481
y_1^2	8,690,214	4,472,844	522,009
N	400	2,250	1,200
n	10	15	6
Variance	1.19	1.10	4.15
(Nh/N)^2	0.011	0.342	0.097
Stratified variance	0.79		
Stratified standard error	0.89		

A3.4.2 Estimation of difference of ratios

In case we want to test the difference of ratios between two periods, we can apply the following approach.

The difference between the ratios (based on a stratified sample) is:

$\hat{r}_1 - \hat{r}_2$ = 6.89 - 5.91 = 0.98

The stratified variance of the difference in the ratio can be calculated using equations 10 to 13.

$$\mathrm{var}(\hat{r}_{h1} - \hat{r}_{h2}) = \mathrm{var}(\hat{r}_{h1}) + \mathrm{var}(\hat{r}_{h2}) - 2\,\mathrm{cov}(\hat{r}_{h1}\hat{r}_{h2}) \qquad (10)$$

$$\mathrm{var}(\hat{r}_{h.}) = \frac{1}{n(n-1)\bar{x}_{h.}^2}\left(\sum_{i=1}^{n} y_{h.i}{}^2 - 2\hat{r}_h \sum_{i=1}^{n} y_{h.i}x_{h.i} + \hat{r}_h^2 \sum_{i=1}^{n} x_{h.i}^2\right) \qquad (11)$$

$$\mathrm{cov}(r_{h1}r_{h2}) = \frac{1}{n(n-1)\bar{x}_{h1}\bar{x}_{h2}} * \left(\sum_{i=1}^{n_h} y_{h1i}y_{h2i} - r_{h1}\sum_{i=1}^{n_h} y_{h2i}x_{h1i} - \hat{r}_{h2}\sum_{i=1}^{n_h} y_{h1i}x_{h2i} + \hat{r}_{h1}\hat{r}_{h2}\sum_{i=1}^{n_h} x_{h1i}x_{h2i}\right) \qquad (12)$$

$$\mathrm{var}_s(\hat{r}_1 - \hat{r}_2) = \sum_{h=1}^{H}\left(\frac{N_h}{N}\right)^2 \mathrm{var}(\hat{r}_{h1} - \hat{r}_{h2}) \qquad (13)$$

Table A3.5 shows the sum of the values relevant for the calculation of the variances and covariances given in table A3.6

Table A3.5 Input for the calculation of the variance of the difference

Sum of values of:	Stratum 1	Stratum 2	Stratum 3
x_1^2	161,400	71,471	6,391
x_1*y_1	1,067,230	494,645	50,481
y_1^2	8,690,214	4,472,844	522,009
x_2^2	170,161	78,482	6,813
x_2*y_2	1,003,675	469,182	42,870
y_2^2	7,347,675	3,391,724	415,152
y_1*y_2	7,372,180	2,725,130	461,914
y_2*x_1	930,620	397,905	42,180
y_1*x_2	1,062,367	440,081	51,532
x_1*x_2	161,586	69,992	6,595
$\hat{r}_1$	6.71	6.69	7.69
$\hat{r}_2$	5.93	5.85	6.11
N	400	2,250	1,200
n	10	15	6

Table A3.6 Stratified variance

	Stratum 1	Stratum 2	Stratum 3
Variance period 1	1.19	1.10	4.15
Variance period 2	0.99	0.57	4.60
Covariance	0.90	0.23	4.32
Variance	0.39	1.21	0.11
(Nh/N)^2	0.011	0.342	0.097
Total stratified var($\hat{r}_1 - \hat{r}_2$)	0.427		

The stratified variance of the estimator ($\hat{r}_1 - \hat{r}_2$) equals 0.427, leading to a stratified standard error of 0.654. The 95% confidence interval:

$$(\hat{r}_1 - \hat{r}_2) \pm 1.96 * s.e.(\hat{r}_1 - \hat{r}_2) = 0.98 \pm 1.96 * 0.654$$

The lower limit is -0.28 and the upper limit is 2.24. Therefore the conclusion is that the difference between the two periods is not significant.