

J. Dijk

Onderzoekverslag 53

DE STEEKPROEF GEWOGEN

EEN EVALUATIE VAN HET LEI-BOEKHOUDNET
VAN LANDBOUWBEDRIJVEN



SIGN: L 28-53
EX. NO: A
MLV: 8941511

September 1989

Landbouw-Economisch Instituut
Afdeling Landbouw

REFERAAT

DE STEEKPROEF GEWOGEN; EEN EVALUATIE VAN HET LEI-BOEKHOUDNET VAN
LANDBOUWBEDRIJVEN

Dijk, J.

Den Haag, Landbouw-Economisch Instituut, 1989

Onderzoekverslag 53

ISBN 90-5242-038-6

258 p., tab., fig.

Dit rapport geeft een analyse van de kwaliteit van de uitkomsten van de LEI-steekproef van landbouwbedrijven. Centraal staat de vraag of de LEI-steekproef in staat is om een representatief en betrouwbaar beeld te geven van de bedrijfsuitkomsten van de landbouwbedrijven in Nederland en van de financiële positie van de ondernemers op die bedrijven. Met het oog daarop worden de steekproeven voor de boekjaren 1975-1987 doorgelicht.

Geconcludeerd wordt dat de steekproef op een groot aantal punten een goede weergave geeft van de situatie in de Nederlandse landbouw. Systematische en toevallige afwijkingen tussen steekproef en populatie zijn voor de belangrijkste steekproefuitkomsten relatief klein. De toepassing van stratificatie bij de steekproefopzet en het gebruik van wegingsfactoren bij de berekening van uitkomsten zorgen ervoor dat de relatief hoge non-respons niet leidt tot vertekende resultaten.

Het rapport begint met een korte schets van de gegevensverzameling in de landbouw in de periode 1923-1975. Aan het eind worden de lijnen doorgetrokken naar de periode vanaf 1988. Met ingang van dit boekjaar is de steekproefopzet licht gewijzigd.

Steekproef/Stratificatie/Non-respons/Landbouwbedrijven/Bedrijfs-
uitkomsten/Financiële positie/Representativiteit/Betrouwbaarheid

CIP-GEGEVENS KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Dijk, J.

De steekproef gewogen : een evaluatie van het
LEI-boekhoudnet van landbouwbedrijven / J. Dijk. - Den Haag
: Landbouw-Economisch Instituut. - Ill., fig., tab. -
(Onderzoekverslag / Landbouw-Economisch Instituut ; 53)
ISBN 90-5242-038-6
SISO 631.5 UDC 631.15/.16(492) NUGI 835
Trefw.: landbouweconomie.

Overname van de inhoud toegestaan, mits met duidelijke bronvermelding.

Inhoud

	Blz.
WOORD VOORAF	11
SAMENVATTING	13
DEEL 1: INLEIDING EN HISTORISCH OVERZICHT	19
1. INLEIDING	21
1.1 Inleiding	21
1.2 De historie	21
1.3 Doel van het onderzoek	22
1.4 De aanpak van het onderzoek	23
1.5 De relevantie van het onderzoek	24
1.6 De opzet van dit verslag	24
2. GEGEVENSVERZAMELING IN DE LANDBOUW VANAF 1923	27
2.1 Inleiding	27
2.2 Boekhoudingen van boekhoudbureaus	27
2.3 De indeling van bedrijven in groepen	28
2.4 De situatie in de veertiger jaren	29
2.5 Twee bedrijfsuitkomsten-statistieken	31
3. GEGEVENSVERZAMELING DOOR HET LEI VANAF 1940	33
3.1 Inleiding	33
3.2 De bedrijfskeuze tot 1966	33
3.3 De overgang op bedrijfskeuze via een steekproef	36
3.4 Veranderende doelstellingen van het boekhoudnet	37
3.5 De steekproefopzet vanaf 1966	37
3.6 De steekproeven voor de boekjaren 1968-1974	38
3.7 Representativiteit	42
3.8 Verandering van de opzet in 1975	43
DEEL 2: DE STEEKPROEF IN DE BOEKJAREN 1975 TOT EN MET 1987	45
4. DOELSTELLINGEN EN MIDDELEN	47
4.1 Inleiding	47
4.3 De doelpopulatie	47
4.3 De doelvariabelen	48
4.4 Kwaliteit van de informatie	49
4.4.1 Inleiding	49
4.4.2 Representativiteit	49
4.4.3 Betrouwbaarheid	50
4.4.4 Concretisering	50
4.5 Middelen	51

INHOUD (1e vervolg)

	Blz.
5. OPZET VAN DE STEEKPROEF	52
5.1 Inleiding	52
5.2 Stratificatie	52
5.3 Indeling in grootte-klassen	53
5.4 Indeling in bedrijfstypen	53
5.5 32 basisstrata	54
5.6 Verdergaande stratificatie	54
5.6.1 Vorming van substrata	54
5.6.2 Een voorbeeld van substratificatie	55
5.7 Bewaking van de representativiteit	56
5.7.1 Inleiding	56
5.7.2 Non-respons bij de keuze	56
5.7.3 Non-respons in de uitwerkingsfase	57
5.7.4 Het roterende karakter van de steekproef	57
5.7.5 Non-respons voor financiële gegevens	58
6. STATISTISCHE ACHTERGRONDEN	59
6.1 Inleiding	59
6.2 Het meten van betrouwbaarheid	59
6.3 Berekening van standaardfouten	60
6.4 Standaardfouten bij een gestratificeerde steekproef	62
6.5 Verdeling van de steekproefbedrijven over de strata	64
6.5.1 Optimale allocatie	64
6.5.2 Sbe per bedrijf als hulpvariabele	65
6.6 Vorming van strata	66
6.6.1 Optimale stratificatie	66
6.6.2 Een voorbeeld	67
6.6.3 Optimale allocatie binnen optimale strata	67
6.6.4 Allocatie in de praktijk	68
6.7 Van 32 basisstrata naar 275 strata	68
6.8 De indeling in strata nader bekeken	69
7. BIJSTELLING VAN DE STEEKPROEF OP GROND VAN NEVENDOELSTELLINGEN	71
7.1 Inleiding	71
7.2 BUL-groepen	71
7.3 FIP-groepen	72
7.4 Regionale aanvullingen	72
7.4.1 Vijftig bedrijven	72
7.4.2 Aanvulling voor akkerbouwgroepen	72
7.4.3 Afstemming bedrijfskeuze op BUL-groepen	73

INHOUD (2e vervolg)

	Blz.
8. DE KEUZE VAN DE BEDRIJVEN	75
8.1 Inleiding	75
8.2 Blijvende bedrijven	75
8.3 Nieuwe bedrijven	76
8.4 Vervangende keuze bij non-respons	77
9. DE BEREKENING VAN STEEKPROEFUITKOMSTEN	79
9.1 Inleiding	79
9.2 Stratificatie achteraf	79
9.3 Meitelling per begin boekjaar als basis	80
9.4 Verschil tussen doelpopulatie en keuze-populatie	81
9.5 Wegingsfactoren	83
9.6 Samenvoeging van strata	84
9.7 Post-stratificatie	84
10. NON-RESPONS	87
10.1 Inleiding	87
10.2 Respons	87
10.3 Non-respons bij de bedrijfskeuze	89
10.3.1 Gangbare berekening van non-respons	89
10.3.2 De keuze in de boekjaren 1985-1987	90
10.3.3 Non-respons en bedrijfstype	91
10.3.4 Non-respons en bedrijfsgrootte	92
10.3.5 Non-respons en leeftijd van de ondernemer	93
10.3.6 Non-respons en oppervlakte	95
10.3.7 Non-respons en regio	96
10.3.8 Non-respons en opleiding van de bedrijfsleider	97
10.3.9 Geschiktheid en bereidheid tot deelname	98
10.3.10 Realisatie keuze per 1 mei	101
10.3.11 Non-respons aan het begin van het boekjaar	102
10.3.12 Conclusies	103
10.4 Non-respons bij de uitwerking van boekhoudingen (BUL)	103
10.4.1 Inleiding	103
10.4.2 Samenhang met stratificatie-variabelen	104
10.4.3 De noodzaak tot stratificatie	105
10.5 Non-respons bij de uitwerking van boekhoudingen (FIP)	105
10.5.1 Inleiding	105
10.5.2 Bereidheid tot het leveren van een financieringsboekhouding	106
10.5.3 Uitsluiting van bedrijven met een onvolledige boekhouding	107

INHOUD (3e vervolg)

	Blz.
11. REPRESENTATIVITEIT	108
11.1 Inleiding	108
11.2 Vergelijking steekproef met metelling (BUL)	109
11.2.1 Inleiding	109
11.2.2 Kleine afwijkingen	109
11.2.3 Arbeidskrachten	110
11.2.4 Oppervlakten	110
11.2.5 Intensieve veehouderij	111
11.2.6 Overige dieren	115
11.2.7 Gewassen	115
11.2.8 Overige metellingsgegevens	116
11.3 Vergelijking steekproef met metelling (FIP)	117
11.3.1 Inleiding	117
11.3.2 Vergelijking voor de totale populatie	117
11.3.3 Uitsluiting van bedrijven met een onvolledige boekhouding	118
11.4 Vergelijking steekproef met externe bronnen	121
11.4.1 Inleiding	121
11.4.2 Rundveehouderij	121
11.4.3 Akkerbouw	125
11.4.4 Intensieve veehouderij	127
11.5 Vergelijking 32 en 275 wegingsfactoren	129
11.6 Conclusies	133
12. BETROUWBAARHEID OP NATIONAAL NIVEAU	135
12.1 Inleiding	135
12.2 De nationale boerderij	135
12.3 Betrouwbaarheid	136
12.4 De doelvariabelen	136
12.5 Bedrijfseconomische grootheden	138
12.5.1 Standaardfouten	138
12.5.2 De invloed van de stratificatie	139
12.6 Financiële grootheden	141
12.6.1 Standaardfouten	141
12.6.2 De invloed van de stratificatie	143
12.7 Quotiënten	144
12.9 Evaluatie van het allocatie-mechanisme	148
12.9.1 Inleiding	148
12.9.2 Een voorbeeld	149
12.9.3 De boekjaren 1981-1986	150
12.9.4 Enkele kanttekeningen	154
12.10 Bedrijfstypering en betrouwbaarheid	154

INHOUD (4e vervolg)

	Blz.
13. BETROUWBAARHEID OP GROEPS-NIVEAU	156
13.1 Inleiding	156
13.2 Indeling op basis van boekhoudnetgegevens	156
13.3 Van strata naar BUL-groepen	157
13.4 Betrouwbaarheid van uitkomsten van BUL-groepen	160
13.4.1 Geschatte wegingsfactoren	160
13.4.2 Vaste wegingsfactoren	162
13.5 Betrouwbaarheid van totalen van BUL-groepen	164
13.6 Betrouwbaarheid van uitkomsten van FIP-groepen	164
13.6.1 Geschatte wegingsfactoren	164
13.6.2 Vaste wegingsfactoren	167
13.7 Betrouwbaarheid van totalen van FIP-groepen	170
13.8 Betrouwbaarheid van schattingen van aantallen bedrijven	170
13.9 Betrouwbaarheid van verschillen tussen groepen	172
14. ALTERNATIEVE SCHATTERS	174
14.1 Inleiding	174
14.2 Quotiëntschatters	174
14.2.1 Inleiding	174
14.2.2 Enkelvoudige steekproeven	174
14.2.3 Gestratificeerde steekproeven	175
14.2.4 Quotiëntschattingen voor de nationale boerderij	176
14.2.5 Quotiëntschatters en representativiteit	176
14.2.6 Quotiëntschatters als alternatief voor stratificatie	178
14.3 Regressieschatters	179
14.3.1 Inleiding	179
14.3.2 Enkelvoudige steekproeven	179
14.3.3 Gestratificeerde steekproeven	180
14.4 Verschilschatters	182
14.4.1 Inleiding	182
14.4.2 Enkelvoudige steekproeven	182
14.4.3 Gestratificeerde steekproeven	182
14.5 Algemene regressieschatters	183
14.6 Schatters voor kleine deelpopulaties	184
15. HET BOEKHOUDNET ALS ROTEREND PANEL	185
15.1 Inleiding	185
15.2 Bedrijven niet langer dan zes jaar in admini- stratie	185
15.3 Jaarlijkse vervanging afhankelijk van doelstel- lingen	185
15.3.1 Inleiding	185

INHOUD (5e vervolg)

	Blz.
15.3.2 Het schatten van veranderingen	186
15.3.3 Het schatten van een meerjarig gemiddelde	187
15.3.4 Het schatten van gemiddelden voor een be- paald jaar	188
15.4 Voorlopige uitkomsten op basis van blijvende be- drijven	189
15.4.1 Inleiding	189
15.4.2 Oude wegingsfactoren	190
15.4.3 Voorlopige wegingsfactoren	191
15.4.4 Verschillen met definitieve uitkomsten	191
15.5 Prognoses voor een lopend jaar	193
16. CONCLUSIES PERIODE 1975-1987	195
16.1 Inleiding	195
16.2 Representativiteit	195
16.3 Betrouwbaarheid	197
DEEL 3: DE STEEKPROEF VANAF BOEKJAAR 1988	201
17. OVERGANG OP DE NEG-TYPERING	203
17.1 Inleiding	203
17.2 De NEG-typering	203
17.3 BUL-groepen op NEG-basis	204
17.4 Afbakening steekproefpopulatie	204
17.5 Naar een uniforme typering	205
17.6 BUL-typering: homogeniteit of herkenbaarheid?	206
18. VAN DOELSTELLINGEN NAAR STEEKPROEF	208
18.1 Inleiding	208
18.2 Uitwerking hoofd- en nevendoelestellingen 1975- 1987	208
18.3 Afstemming stratificatie op groepsindeling bij publikatie	209
18.4 Explicitering van doelstellingen	210
19. CONCRETISERING VAN DE DOELSTELLINGEN	212
19.1 Inleiding	212
19.2 De nationale boerderij	212
19.3 BUL-groepen	213
20. STEEKPROEFOPZET VANAF BOEKJAAR 1988	217
20.1 Inleiding	217
20.2 Steekproefopzet op basis van hoofddoelstelling	217
20.3 Confrontatie hoofd- en nevendoelestellingen	220

INHOUD (6e vervolg)

	Blz.
20.4 Kleine inkrimping boekhoudnet	222
20.5 Te verwachten betrouwbaarheid in boekjaar 1988	222
21. VERGELIJKBAARHEID OUDE EN NIEUWE OPZET	226
21.1 Inleiding	226
21.2 Populatie en bedrijfstypen bij de stratificatie	226
21.3 De oude en de nieuwe nationale boerderij	227
21.4 Oude en nieuwe BUL-groepen	230
22. CONCLUSIES	231
LITERATUUR	234
BIJLAGEN	240
1. Stratificatie bij de steekproefopzet 1975-1987	240
2. Steekproefpercentages per basisstratum (1987)	243
3. De opbouw van het boekhoudnet	244
4. Overzicht van standaardfouten voor FIP-grootheden	245
5. (Sbe) ^k als hulpvariabele	247
6. Stratumherkomst van grotere akkerbouwbedrijven	248
7. Keuze-aantallen en steekproefpercentages bij de steekproefopzet voor boekjaar 1988	249
8. Vergelijking BUL-wensen voor en na boekjaar 1988	252
9. Standaardfouten voor deelpopulaties	254

Woord vooraf

In 1987 verscheen bij het LEI een Mededeling waarin een beschrijving werd gegeven van het boekhoudnet van landbouwbedrijven. Deze Mededeling gaf de achtergronden van de LEI-steekproef van ruim 1.000 landbouwbedrijven zoals die sinds 1975 jaarlijks wordt opgezet. Dit onderzoekverslag vormt in zekere zin een vervolg op die Mededeling. Het gaat in op de vraag in hoeverre de opzet van de steekproef in de praktijk leidt tot kwalitatief goede uitkomsten. Daarbij wordt niet alleen de huidige steekproef in ogeschouw genomen, maar wordt ook aandacht besteed aan de verzameling van boekhouddgegevens voor 1975 en zelfs aan de gang van zaken in de periode voor de oprichting van het LEI in 1940. In enkele slothoofdstukken wordt de lijn doorgetrokken naar de toekomst.

Het rapport is bedoeld als handreiking aan de gebruikers van gegevens uit het LEI-boekhoudnet van landbouwbedrijven. Kennis van de mogelijkheden en de beperkingen van de steekproefuitkomsten kan leiden tot een beter gebruik van de jaarlijkse cijferstroom.

Dank is verschuldigd aan iedereen, binnen en buiten de afdeling landbouw van het LEI, die de auteur heeft ondersteund bij de uitvoering van het onderzoek.

De directeur



J. de Veer

Den Haag, september 1989

Samenvatting

De LEI-steekproef is een belangrijke bron van informatie over de bedrijfsuitkomsten in de Nederlandse landbouw. De kwaliteit van die informatie is het centrale onderwerp in dit rapport. Onderzocht is hoe representatief en betrouwbaar de steekproefuitkomsten waren in de periode 1975-1987.

Het rapport haalt eerst de periode voor 1975 voor het voetlicht. Al meer dan 65 jaar worden in Nederland statistieken van bedrijfsuitkomsten in de landbouw gepubliceerd. Aanvankelijk gebeurde dat alleen op basis van boekhoudingen die voor fiscale doeleinden waren opgesteld. Na de Tweede Wereldoorlog werden ook bedrijfsuitkomsten op basis van bedrijfseconomische boekhoudingen gepubliceerd.

Vanaf 1923 verzorgde de *Directie van den Landbouw* de bedrijfsuitkomstenstatistieken. De gegevens van enkele duizenden bedrijven werden verzameld via de landbouw-boekhoudbureaus. Een vertaling van de gemiddelde bedrijfsuitkomsten naar de populatie van alle landbouwbedrijven was niet eenvoudig te maken omdat niet alle delen van de populatie even sterk vertegenwoordigd waren in wat genoemd werd de "gecontroleerde bedrijven". De uitkomsten werden dan ook vooral in termen van kosten en opbrengsten per hectare gepresenteerd.

In 1946 nam het LEI de statistiek over van de Directie van den Landbouw. Na enkele jaren werd deze activiteit echter weer afgestoten omdat het LEI inmiddels ook bedrijfsuitkomsten op bedrijfseconomische basis was gaan verzorgen. Publikatie van twee verschillende statistieken met verschillende uitgangspunten werd door het LEI-bestuur niet gewenst geacht. Het LEI koos voor publikatie op basis van de boekhoudingen die door het instituut zelf werden bijgehouden. Het Centraal Bureau voor de Statistiek (CBS) nam de bedrijfsuitkomstenstatistiek op fiscale basis over.

In de periode 1947-1962 publiceerde het LEI jaarlijks statistische overzichten van de uitkomsten van landbouwbedrijven. De gegevens hiervoor waren afkomstig van bedrijven die daarvoor werden aangezocht door regionale keuze-commissies. In deze commissies hadden vertegenwoordigers van het Landbouwschap, de Landbouwvoorlichting en het LEI zitting. Geprobeerd werd om "vakbekwame" boeren voor een periode van vijf jaar tot deelname aan het LEI-boekhoudnet te bewegen. De bedrijven moesten qua produktieomstandigheden en bedrijfsvoering "typisch" zijn voor de regio waarin ze lagen.

In de vijftiger en zestiger jaren werd onderzocht in hoeverre deze manier van kiezen van bedrijven leidde tot bedrijfsuitkomsten die een goede afspiegeling gaven van de situatie in de populatie. Geconstateerd werd dat de LEI-bedrijven gemiddeld ge-

nomen weliswaar goede representanten waren voor de bedrijven in de populatie, maar dat de spreiding in de bedrijfsuitkomsten geringer was dan in de populatie. Daarom werd in 1966 overgeschakeld op een bedrijfskeuze via een steekproef. Daardoor konden ook de extreem slechte en de extreem goede bedrijven in het boekhoudnet worden opgenomen. De naam die het LEI zich in 25 jaar had opgebouwd gaf de mogelijkheid om zonder tussenkomst van commissies toch een goede respons te realiseren.

De overgang op een ander systeem van bedrijfskeuze hield ook verband met EG-verplichtingen. Ten behoeve van het EG-informatienet van landbouwbedrijven moesten jaarlijks 750 boekhoudingen geleverd worden die samen een beeld moesten kunnen geven van de situatie in de Nederlandse landbouw. De nadruk werd verlegd van bepaalde "typische" groepen binnen de populatie naar de gehele Nederlandse landbouw.

Er werd gewerkt met een gestratificeerde steekproef met verschillen in steekproefpercentage tussen de strata. Eerst werd gestratificeerd naar EEG-bedrijfstype en hectare-klasse. Langzamerhand deden ook de sbe en de regio hun intrede bij de stratificatie. Verschillen in respons en blijvende aandacht voor afzonderlijke (regionale) groepen binnen de populatie noopten daartoe.

Vanaf 1975 werd een nieuw stratificatie-schema gebruikt waarbij de populatie van alle landbouwbedrijven die een bepaalde minimale grootte hadden in acht bedrijfstypen werden ingedeeld. Per bedrijfstype werden vier sbe-klassen onderscheiden. Binnen de zo gevormde 32 basisstrata werd verder gestratificeerd met behulp van de variabelen oppervlakte, leeftijd van de ondernemer en regio. Zo ontstonden ieder jaar ongeveer 275 strata. De stratificatie was er op gericht om zowel systematische als toevallige fouten terug te dringen. Hoofddoelstelling was om schattingen voor de "totale nationale boerderij" zo betrouwbaar mogelijk te doen zijn. Daarnaast werd de doelstelling gehandhaafd dat voor ongeveer twintig groepen bedrijven binnen de populatie afzonderlijke schattingen met een voldoende betrouwbaarheid gemaakt moesten kunnen worden.

De 1.150 bedrijven die in het kader van de steekproef in administratie gehouden konden worden, werden op een zodanige manier over de strata verdeeld dat er een redelijk evenwicht tussen de hoofddoelstelling en de nevendoelestellingen gerealiseerd werd. Bedrijven werden aangezocht om zes jaar aan het boekhoudnet mee te doen. Ieder jaar werd een kwart van de bedrijven vervangen door nieuwe bedrijven. Dit betrof zowel de vervanging van bedrijven die zes jaar hadden deelgenomen als de aanvulling voor bedrijven die tussentijds hun deelname opzegden.

Bij de bedrijfskeuze werden aangezochte bedrijven die niet in staat waren of weigerden deel te nemen aan de steekproef vervangen door "gelijkende" bedrijven om er zo voor te zorgen dat bepaalde soorten bedrijven niet ondervertegenwoordigd zouden raken in de steekproef door selectieve respons. Bij het berekenen

van steekproefuitkomsten na afloop van een boekjaar werd aan ieder bedrijf een wegings- of ophoogfactor toegekend die aangaf welk aantal bedrijven in de populatie door een specifiek steekproefbedrijf werd vertegenwoordigd. Daarmee werd gecorrigeerd voor verschillen in steekproef- en responspercentages.

Een belangrijke vraag is of de gehanteerde steekproefmethode ondanks verschillen in respons tussen delen van de populatie leidt tot representatieve steekproefuitkomsten. Bij de bedrijfskeuze is 17% van de bedrijven niet geschikt voor of niet in staat tot deelname. Van de resterende bedrijven is ongeveer 50% bereid om deel te nemen aan de steekproef. Daarbij zijn echter grote verschillen in respons tussen bedrijfstypen, sbe-, oppervlakte- en leeftijds- klassen en regio's te constateren.

Van de bedrijven die zich bereid verklaren om deel te nemen valt nog eens 10% af als het er aan het begin van het boekjaar daadwerkelijk op aan komt om gegevens te leveren. Daarnaast zijn er strata waar het zo moeilijk is om een bedrijf tot deelname te bewegen dat er voor een deel van de te kiezen bedrijven nog geen bereidverklaring gevonden is aan het begin van het boekjaar. Gemiddeld gaat het daarbij om 10% van het aantal te kiezen bedrijven.

Van de 1.150 bedrijven die in administratie worden genomen, leveren er gemiddeld ruim 100 geen uitgewerkte boekhouding aan het eind van het boekjaar. Ook hier zijn weer verschillen tussen delen van de populatie aanwezig. Niet alle resterende 1.000 bedrijven die worden opgenomen in de "Bedrijfsuitkomsten in de Landbouw" (BUL) zijn bereid om ook een financieringsboekhouding te leveren. Voor de samenstelling van "De Financiële positie van de landbouw" (FIP) zijn daarom maar ongeveer 900 boekhoudingen per jaar beschikbaar. Hiervan worden er nog eens ongeveer 60 uitgesloten omdat ze geen volledige financieringsboekhouding kunnen leveren (NV's en BV's).

In de praktijk blijkt dat ondanks de vrij forse non-respons in de verschillende stadia van de gegevensverzameling en ondanks het verband dat er is tussen de non-respons en de kenmerken van bedrijven, de steekproefuitkomsten nauwelijks systematisch afwijken van de uitkomsten in de populatie. Dat geldt zeker bij vergelijking van structuur-kenmerken van de bedrijven (de metelling). Maar ook bij vergelijking van steekproefgegevens met gegevens uit andere bronnen zijn er nauwelijks aanwijzingen dat de steekproefuitkomsten niet representatief zouden zijn voor de uitkomsten in de populatie. Fysiske opbrengsten per hectare of per dier verschillen weinig met die in andere bronnen.

Een sluitende controle voor de representativiteit van de bedrijfsuitkomsten kan niet gerealiseerd worden omdat dan van alle bedrijven in de populatie gegevens over die bedrijfsuitkomsten beschikbaar zouden moeten zijn. Maar op punten waar wel gegevens beschikbaar zijn is de representativiteit vrij goed. De manier

van kiezen van steekproefbedrijven en het gebruik van wegingsfactoren bij het berekenen van steekproefuitkomsten zorgen daarvoor.

Naast mogelijke systematische fouten bevatten steekproefuitkomsten altijd toevallige "fouten". Door de gevolgde steekproefmethode en het vrij grote aantal steekproefbedrijven bezitten de schattingen voor de nationale boerderij een grote mate van betrouwbaarheid. De kans dat de steekproefuitkomsten door het steekproeftoeval meer dan enkele procenten afwijken van de werkelijke uitkomsten is voor de meeste geschatte grootheden klein. Relatieve standaardfouten van 1 tot 2% zijn geen uitzondering. Door de stratificatie worden de standaardfouten in veel gevallen met tientallen procenten gereduceerd. Deze reductie is vooral toe te schrijven aan de stratificatievariabelen bedrijfstype en sbe-klasse.

Voor groepen binnen de populatie liggen de (relatieve) standaardfouten in het algemeen op een hoger niveau. De kleinere aantallen steekproefbedrijven dragen daaraan bij, maar ook het feit dat de groepsgrenzen niet samenvallen met de stratumgrenzen zorgt voor een vergroting van de standaardfouten, met name bij schattingen voor aantallen bedrijven en groepstotalen. De betrouwbaarheid van de schattingen is echter in vrijwel alle gevallen alleszins acceptabel.

Door gebruik te maken van quotiënt- en regressieschatters kan in sommige gevallen de betrouwbaarheid nog vergroot worden.

Verschillen in uitkomsten tussen twee boekjaren zijn door het roterende karakter van de steekproef vrij betrouwbaar te schatten. Datzelfde geldt voor voorlopige uitkomsten op basis van een deel van de steekproefbedrijven.

Met ingang van boekjaar 1988 zal het LEI bij de publikatie van bedrijfsuitkomsten gebruik maken van de NEG-typering. Deze nieuwe bedrijfstypering die in 1986 in Nederland is ingevoerd, sluit aan bij de al langer bestaande EG-typering.

Voor de nieuwe BUL- en FIP-groepen zijn concrete doelstellingen geformuleerd voor wat betreft de betrouwbaarheid van de steekproefuitkomsten. Vanuit deze doelstellingen is de tot 1987 gebruikte steekproefopzet enigszins aangepast. Hoofd- en neven-doelstellingen van de steekproef worden daarbij explicieter tegen elkaar afgewogen dan in het verleden. Daarnaast laat de nieuwe opzet de indeling bij de stratificatie meer aansluiten bij de indeling in groepen bij de publikatie van gegevens. Daardoor kan een grotere betrouwbaarheid van schattingen voor deze groepen gerealiseerd worden.

De populatie waarop de steekproef betrekking heeft wordt vanaf 1988 ook afgebakend met behulp van de NEG-typering. Het boekhoudnet van landbouwbedrijven heeft in het vervolg betrekking op alle bedrijven die groter zijn dan 20 nge (nederlandse grootte eenheden) en kleiner dan 500 nge en niet behoren tot de bedrijfstypen "tuinbouwbedrijven", "blijvende teeltbedrijven" of "tuin-

bouw/ blijvende teeltbedrijven". Het boekhoudnet van tuinbouwbedrijven richt zich op laatstgenoemde groepen.

Deel 1: inleiding en historisch overzicht

"Eveneens moet ervoor gewaarschuwd worden om uit de betrekkelijk kleine groep bedrijven gevolgtrekkingen te maken ten aanzien van den gang van zaken in het Nederlandsch Landbouwbedrijf in het algemeen. Daarvoor is de representatie der verschillende bedrijfspvormen en bedrijfsgrootten in het geheel te aanvechtbaar. Immers het voor Nederland zoo belangrijke kleinbedrijf is te zwak vertegenwoordigd. Niettemin wordt het dezerzijds gewenscht geacht om, vooral nu het landbouwbedrijf van zoovele zijden kritisch wordt beschouwd, aan te toonen, dat het hier ook niet overal "botertje is tot aan den boom".

(Ir. C.H.J. Maliepaard, Voorbericht van "Bedrijfsuitkomsten van den Landbouw 1942-43")

1. Inleiding

1.1 Inleiding

Het verzamelen en publiceren van bedrijfseconomische en financiële gegevens van landbouwbedrijven is een activiteit van het LEI die al vanaf de oprichting van het instituut verricht wordt. Daarvoor wordt een zogenaamd boekhoudnet van bedrijven bijgehouden. Deze bedrijven leveren gedurende een aantal achtereenvolgende jaren een groot aantal gegevens die betrekking hebben op het bedrijf en de ondernemer. Het aantal bedrijven dat in het boekhoudnet is opgenomen, is maar een fractie van het aantal bedrijven dat in de jaarlijkse CBS-meitelling geregistreerd wordt. De toepassing van geschikte steekproefmethoden en het gebruik van achtergrondinformatie uit de metellingen, maken het mogelijk om met een relatief klein aantal bedrijven een goede beschrijving te geven van wat zich op de landbouwbedrijven in Nederland afspeelt.

Het boekhoudnet van landbouwbedrijven waar dit rapport betrekking op heeft, richt zich op akkerbouw- en veehouderijbedrijven. Een beschrijving van het LEI-boekhoudnet van tuinbouwbedrijven is te vinden in van der Plas (1985).

1.2 De historie

De manier waarop verzameling en publikatie van bedrijfsuitkomsten in de landbouw plaatsvinden, is in de loop van de tijd verschillende malen aan verandering onderhevig geweest. Dat hing onder meer samen met de doelstelling van deze activiteiten. In de beginjaren van het LEI waren kostprijsberekeningen en andere bedrijfseconomische analyses sterk bepalend voor de informatiebehoefte. Voor dat doel werden heel gericht bedrijven benaderd door keuze-commissies. Voor tientallen specifieke groepen binnen de landbouw werden cijfers gepubliceerd. Vanaf de zestiger jaren kwam het accent meer te liggen op een beschrijving van de gehele Nederlandse landbouw en werd overgegaan op de keuze van bedrijven door middel van een steekproef. In 1975 werd de steekproefopzet op grond van de ervaringen tot dan toe enigszins bijgesteld. Lodder (1987) heeft een beschrijving gegeven van de gestratificeerde steekproef die sinds 1975 wordt getrokken en van de manier waarop steekproefuitkomsten worden berekend.

De voortdurende ontwikkeling in de opzet van het LEI-boekhoudnet had niet alleen te maken met veranderingen in doelstellingen. Ook het groeiende inzicht in de factoren die samenhangen met het al dan niet terecht komen van bedrijven in het boekhoudnet, droeg daar aan bij. De zorg voor representativiteit van de cijfers die het LEI op grond van het boekhoudnet naar buiten

bracht, leidde enkele malen tot een aanpassing in de keuze van bedrijven voor het boekhoudnet. De overgang op een bedrijfskeuze via een steekproef in 1966 was de meest fundamentele aanpassing.

1.3 Doel van het onderzoek

Aan een klein deel van alle landbouwbedrijven wordt dus gevraagd om een uitvoerige boekhouding bij te houden ten behoeve van het LEI. De cijfers van deze boekhoudnetbedrijven worden geacht een goed beeld te geven van het economische en financiële reilen en zeilen van alle landbouwbedrijven. Het gebruik van een steekproef brengt echter met zich mee dat er door het toeval verschillen kunnen bestaan tussen de uitkomsten die op grond van het boekhoudnet worden berekend en de werkelijke uitkomsten op alle bedrijven. Deze verschillen zullen in het algemeen klein zijn als het aantal steekproefbedrijven groot is. Bij berekeningen op basis van kleine aantallen steekproefbedrijven kan het toeval zich echter sterk doen gelden.

Daarnaast is er het gevaar dat in de steekproef bepaalde soorten bedrijven onder- of oververtegenwoordigd zijn, waardoor systematische verschillen tussen steekproef en populatie ontstaan. Indien bijvoorbeeld de "betere" bedrijven een grotere bereidheid zouden hebben om mee te werken aan de steekproef, waardoor ze te sterk vertegenwoordigd zouden worden in de steekproef, zouden de steekproefuitkomsten een te positief beeld geven van de situatie zoals die gemiddeld in de populatie is.

Het doel van dit onderzoek is om na te gaan of de steekproef die het LEI sinds 1975 jaarlijks opzet, in staat is om een goede beschrijving te geven van de populatie waar de steekproef betrekking op heeft. Bij de opzet in 1975 werd gesteld dat de steekproef zo betrouwbaar mogelijke totaalcijfers voor de Nederlandse landbouw moest opleveren. "Betrouwbaar" wil in dit verband zeggen dat de invloed van het toeval op de steekproefschattingen binnen acceptabele grenzen blijft. Een andere eis die gesteld werd, was dat er sprake moest zijn van "een zo getrouw mogelijke afspiegeling" van de Nederlandse landbouw "in zijn volle verscheidenheid" (Lodder, 1987:8). Deze eis van representativiteit betekende dat voorkomen moest worden dat bepaalde delen van de populatie onder- of oververtegenwoordigd zouden zijn in de steekproef.

Dit onderzoek gaat na of de steekproef voldoet aan de eisen die daar aan werden en worden gesteld. Daarbij draait het om de volgende twee hoofdvragen:

- 1) *hoe representatief zijn de steekproefuitkomsten? Zijn er systematische afwijkingen tussen steekproef en populatie?*
- 2) *hoe betrouwbaar zijn de steekproefuitkomsten? In welke orde van grootte liggen de toevallige afwijkingen tussen steekproef en populatie?*

Deze vragen worden zowel op de gehele populatie van Nederlandse landbouwbedrijven als op groepen binnen die populatie gericht. Nagegaan wordt of de steekproefopzet optimaal is in verband met de eisen die aan de steekproefuitkomsten gesteld worden. Ook wordt onderzocht of structurele veranderingen in de landbouw sinds 1975 en/of veranderende eisen ten aanzien van de te publiceren gegevens - welke groepen en welke variabelen zijn belangrijk - noodzaken tot eventuele aanpassingen in de opzet van de steekproef.

Tenslotte wordt aandacht gegeven aan de gevolgen van de komende introductie van de NEG-typing bij de publikatie van boekhoudnetcijfers, voor de opzet van de steekproef.

1.4 De aanpak van het onderzoek

Voor het onderzoek wordt gebruik gemaakt van de steekproeven in de boekjaren 1975-1986. De representativiteit van de steekproefuitkomsten wordt op twee manieren onderzocht. In de eerste plaats worden de meetingsgegevens van steekproef en populatie naast elkaar gelegd. Dat geeft een indruk van de mate waarin de steekproefbedrijven ten aanzien van allerlei structuurkenmerken overeenkomen met de bedrijven in de populatie. In de tweede plaats worden de uitkomsten van de steekproef op een aantal punten vergeleken met gegevens uit andere bronnen. Daardoor kan duidelijk worden of de bedrijven in de steekproef op deze punten al dan niet systematisch afwijken van de bedrijven in de populatie.

De betrouwbaarheid van steekproefschattingen wordt gemeten met de standaardfout van die schattingen. Deze geven aan welke marges rond een schatting in acht moeten worden genomen in verband met de rol van het toeval. De theorie van de gestratificeerde steekproeven wordt toegepast bij het berekenen van standaardfouten. Aangegeven wordt welke winst - in termen van betrouwbaarheid - stratificatie oplevert.

Het onderzoek richt zich voornamelijk op de groepen bedrijven die in de jaarlijkse LEI-publikatie "Bedrijfsuitkomsten in de Landbouw" centraal staan. Een groot aantal variabelen wordt onder de loep genomen. De verschillen in kwaliteit en bruikbaarheid van schattingen voor verschillende variabelen komen daarmee goed in beeld.

Steekproefuitkomsten resulteren uit een combinatie van de steekproefmethode bij de opzet van de steekproef en de schattingsmethode bij de berekening van de uitkomsten. Nagegaan wordt of aanpassingen in de stratificatie en het gebruik van meer informatie uit de metellingen bij de schattingen, kan leiden tot kwalitatief betere uitkomsten.

Voor boekjaar 1988 wordt een iets aangepaste steekproefopzet gepresenteerd. Daarbij worden de doelstellingen van de LEI-steekproef vrij concreet gemaakt in termen van doelpopulaties en doelvariabelen.

1.5 De relevantie van het onderzoek

Het LEI-boekhoudnet levert een grote stroom van gegevens over de landbouwbedrijven in Nederland. Deze gegevens worden onder andere gebruikt door en verzameld in opdracht van de EG, de nationale overheid en het landbouwbedrijfsleven. Daarnaast wordt veel landbouw-economisch onderzoek in Nederland verricht op basis van dit boekhoud- of informatienet.

De kwaliteit van uitspraken die gedaan worden op basis van gegevens uit het boekhoudnet, hangt in sterke mate af van de kwaliteit van die gegevens. Daarom is het van belang om inzicht te hebben in de totstandkoming er van. De waarde die aan de steekproefuitkomsten moet worden toegekend, wordt in belangrijke mate bepaald door de gebruikte methoden bij het verzamelen van bedrijfsgegevens en bij het berekenen van gemiddelden, totaalcijfers en andere grootheden. Daarom wordt in dit rapport eerst een beschrijving gegeven van die methoden. Deze beschrijving sluit nauw aan bij Lodder (1987). Het inzicht in de kwaliteit van de gegevens kan echter sterk vergroot worden indien op basis van jarenlange ervaring met de steekproef, gekwantificeerd kan worden wat de kwaliteit van bepaalde cijfers is. Goede steekproefmethoden zijn immers wel een noodzakelijke, maar niet altijd een voldoende voorwaarde voor het leveren van bruikbare uitkomsten.

Dit onderzoek wil bijdragen aan een verantwoord gebruik van de uitkomsten van het LEI-boekhoudnet. Sterke en minder sterke kanten van de steekproef zullen aan de orde komen. Niet elk van de duizenden gegevens die per bedrijf worden vastgelegd, is even representatief en betrouwbaar. Dat wil niet zeggen dat sommige gegevens daarom niet bruikbaar zouden zijn. Informatie over representativiteit en betrouwbaarheid kan een gebruiker aanwijzingen geven over de wijze waarop hij gegevens het beste kan gebruiken.

Een kort historisch overzicht van de periode voor 1975 aan het begin van het verslag, geeft enig zicht op de mogelijkheden en beperktheden van het cijfermateriaal dat in die periode op grond van het boekhoudnet door het LEI naar buiten is gebracht.

1.6 De opzet van dit verslag

Dit onderzoekverslag bestaat uit drie delen. De kern van het onderzoek wordt beschreven in deel 2. Daaraan voorafgaand wordt in deel 1 een korte schets gegeven van de gegevensverzameling in de landbouw in de periode tot 1975. Deel 2 beslaat de periode 1975-1987, terwijl deel 3 de gegevensverzameling door het LEI vanaf boekjaar 1988 beschrijft 1).

1) In dit onderzoekverslag wordt onder boekjaar verstaan: de periode die loopt van 1 mei van het aangegeven jaar tot en met 30 april van het volgende jaar.

Deel 1 bespreekt in hoofdstuk 2 de bedrijfsuitkomsten-statistiek van de *Directie van den Landbouw* over de periode 1923-1941. Deze statistiek was gebaseerd op fiscale boekhoudingen die afkomstig waren van landbouwboekhoudbureaus. Na de oorlog is deze statistiek een aantal jaren door het LEI gepubliceerd. Vanaf 1950 werd dit een CBS-activiteit.

In hoofdstuk 3 wordt ingegaan op de statistieken die het LEI in de periode 1946-1963 publiceerde op basis van gegevens van door keuze-commissies gekozen bedrijven. In dit zelfde hoofdstuk wordt de overgang in 1966 naar een steekproefsgewijze bedrijfskeuze beschreven. De ervaringen met het LEI-boekhoudnet tot 1975 vormden de aanleiding tot een aantal wijzigingen in de steekproefopzet in 1975.

Dit eerste deel geeft aan in welk licht gegevens uit het LEI-boekhoudnet in de periode voor 1975 geïnterpreteerd moeten worden 1). Tegelijkertijd wordt duidelijk gemaakt dat een bepaalde wijze van verzamelen en publiceren van gegevens sterk samenhangt met de doelstellingen en de middelen van de informatieverzameling. Verder laat dit deel zien waarom in 1975 voor een vernieuwde steekproefopzet is gekozen.

Deel 2 vormt de kern van dit verslag. Daarin wordt eerst een beschrijving gegeven van de steekproefopzet in de periode 1975-1987. Hoofdstuk 4 bespreekt de doelstellingen en middelen van het LEI-boekhoudnet. Het volgende hoofdstuk geeft in grote lijnen de steekproefopzet voor het boekhoudnet weer, zoals die voortvloeit uit de hoofddoelstelling van het boekhoudnet. De statistische achtergronden van deze opzet - waarbij stratificatie een belangrijke rol speelt - worden in hoofdstuk 6 besproken; in hoofdstuk 5 worden ze slechts kort aangeduid. Hoofdstuk 7 geeft aan in hoeverre de opzet uit hoofdstuk 5 wordt bijgesteld op basis van nevendoelstellingen. De concrete keuze van bedrijven voor het boekhoudnet komt aan de orde in hoofdstuk 8.

Na deze beschrijving van de doelstellingen en de opzet van de steekproef in de hoofdstukken 4 tot en met 8 wordt in hoofdstuk 9 ingegaan op de berekening van steekproefuitkomsten op basis van de gegevens van de deelnemende bedrijven. Het berekenen van wegingsfactoren staat daarbij centraal.

De hoofdstukken 10 tot en met 16 geven een evaluatie van de steekproef in de boekjaren 1975-1987. Centrale begrippen in deze hoofdstukken zijn representativiteit en (statistische) betrouwbaarheid.

1) Het bij sommige gebruikers van de steekproefuitkomsten nog steeds aanwezige misverstand dat het LEI-boekhoudnet de "betere" bedrijven zou bevatten, berust voor een deel op het idee van de gerichte bedrijfskeuze zoals die in de vijftiger jaren werd uitgevoerd.

In hoofdstuk 10 komen de factoren die de representativiteit bedreigen aan de orde. Verschillende vormen van non-respons passeren de revue. Hoofdstuk 11 laat zien in hoeverre de steekproefopzet en de wijze van berekening van resultaten ondanks de non-respons toch tot representatieve steekproefuitkomsten heeft geleid.

In de resterende hoofdstukken van deel 2 wordt de rol van het toeval in de steekproefuitkomsten belicht. Standaardfouten worden gepresenteerd waarmee betrouwbaarheidsintervallen voor schattingen kunnen worden geconstrueerd. In hoofdstuk 12 wordt dat gedaan voor de gehele Nederlandse landbouw en voor het gemiddelde landbouwbedrijf. Hoofdstuk 13 geeft de betrouwbaarheid van schattingen voor bepaalde groepen binnen de populatie. De slot-hoofdstukken van deel 2 geven aan hoe de betrouwbaarheid van schattingen vergroot kan worden door gebruik te maken van hulp-informatie die voor alle bedrijven in de populatie beschikbaar is, en van de panel-structuur van het boekhoudnet. Deel 2 mondt uit in een aantal conclusies (hoofdstuk 16).

In deel 3 wordt verder gewerkt met deze conclusies. Hoofdstuk 17 beschrijft een verandering in de populatie waarop het boekhoudnet zich richt en een verandering in de groeps-samenstellingen bij publikatie van gegevens. Door deze veranderingen per 1 mei 1988 zal de nationale en internationale vergelijkbaarheid van gegevens worden vergroot. De hoofdstukken 18 en 19 maken de doelstellingen van het boekhoudnet en de afweging van conflicterende doelstellingen iets explicieter dan in het verleden. Enige accentverschuivingen in die afweging worden besproken. Ook wordt ingegaan op de afstemming van de steekproefopzet op de opnieuw geformuleerde doelstellingen.

Hoofdstuk 20 geeft aan hoe de in de voorgaande hoofdstukken besproken wijzigingen leiden tot een nieuwe steekproefopzet. De opzet voor boekjaar 1988 wordt vrij gedetailleerd in beeld gebracht. In dit hoofdstuk wordt ook aangegeven wat de te verwachten betrouwbaarheid van de cijfers is in de nieuwe opzet. Uiteraard wordt ook informatie gegeven over de verschillen in populatie en groeps-samenstelling bij stratificatie en publikatie tussen de oude en de nieuwe opzet (hoofdstuk 21). Een concluderend hoofdstuk besluit het verslag.

2. Gegevensverzameling in de landbouw vanaf 1923

2.1 Inleiding

In de jaarlijkse LEI-publikatie "Bedrijfsuitkomsten in de landbouw" (Bedrijfsuitkomsten, 1988) staat in bijlage 1 een "Overzicht van de door de Directie van den Landbouw, Centraal Bureau voor de Statistiek en Landbouw-Economisch Instituut gepubliceerde statistieken van bedrijfsuitkomsten". Daaruit blijkt dat al vanaf 1923 jaarlijks gepubliceerd wordt over deze uitkomsten. De *Directie van den Landbouw* maakte enige tijd na afloop van een boekjaar de "Bedrijfsuitkomsten van den Landbouw" bekend. Deze werden gepresenteerd en besproken in een publikatie die als titel had: "De economische toestand van den Landbouw in 19../.. en andere onderwerpen".

In dit hoofdstuk wordt globaal aangegeven hoe de verzameling en publikatie van bedrijfsgegevens plaats vond in de tijd dat het LEI nog niet bestond.

2.2 Boekhoudingen van boekhoudbureaus

De bedrijfsuitkomsten in de vooroorlogse periode werden samengesteld aan de hand van boekhoudingen die afkomstig waren van landbouw-boekhoudbureaus. "De belangstelling voor dezen eenvoudigen vorm van bedrijfscontrole neemt overal in den lande toe, al moet hierbij wel worden opgemerkt, dat de groei van het aantal aangeslotenen bij het eene boekhoudbureau aanzienlijk groter is dan bij het andere", staat te lezen in de inleiding van "De bedrijfsuitkomsten van den Landbouw 1938/1939" (Directie van den Landbouw, 1940). De cijfers over dit boekjaar verschenen in juli 1940. "In sommige landbouwgebieden is evenwel de gecontroleerde oppervlakte zoo gestegen, dat men zich niet meer de vraag behoeft te stellen of de hier vermelde uitkomsten wel als representatief voor de betreffende streek kunnen worden beschouwd. Dit is zonder eenigen twijfel het geval indien de gecontroleerde oppervlakte 8 pct. of meer van de totale oppervlakte bouw- en grasland van het landbouwgebied bedraagt", zegt dezelfde inleiding. Met de groei van het aantal deelnemende bedrijven groeide dus het vertrouwen dat een goede weergave van de bedrijfsuitkomsten werd gegeven. Voor boekjaar 1938 werden de bedrijfsuitkomsten van ongeveer 8.000 bedrijven verwerkt. In 1930 waren de uitkomsten nog gebaseerd op minder dan 4.000 bedrijven. Dit aantal bleef over na een voorselectie door de boekhoudbureaus en een zekere uitval tijdens de verwerking van de resultaten.

"Er blijft dan nog de vraag of de bedrijven, die in het onderzoek zijn opgenomen, als representatief voor het betreffende

gebied kunnen gelden (...). Het is zeer wel mogelijk dat in meer achterlijke gebieden de meer vooruitstrevende landbouwers er voor voelen zich aan te sluiten bij een boekhoudbureau. Zulke verschijnselen zijn in enkele gebieden wel aan te wijzen" (Directie van den Landbouw, 1937:8).

De berekeningswijze van bedrijfsresultaten verschilde op een aantal punten tussen de boekhoudbureaus. Grotere uniformiteit werd maar heel geleidelijk aan bereikt.

2.3 De indeling van bedrijven in groepen

De bedrijfsuitkomsten werden voor een groot aantal groepen bedrijven berekend. Tot boekjaar 1936 werden voornamelijk cijfers voor 66 afzonderlijke landbouwgebieden - en voor vijf of zes groepen van landbouwgebieden - gepresenteerd. Vanaf 1936 werden ook tabellen gegeven die per groep van landbouwgebieden onderverdelingen gaven naar oppervlakte-klassen en bedrijfssoorten. Veel tabellen vermeldden opbrengsten en uitgaven per hectare in plaats van opbrengsten en uitgaven in absolute zin. Dat was ook min of meer noodzakelijk. Want hoewel er voor de meeste gebieden een vrij groot aantal bedrijven beschikbaar was voor de berekening van bedrijfsuitkomsten, en op grond daarvan geconcludeerd werd dat de "representativiteit" wel redelijk was (zie paragraaf 2.2), weken de deelnemende bedrijven toch enigszins af van alle bedrijven die gerepresenteerd moesten worden. De verdeling van de bedrijven over de grootte-klassen was nogal scheef. "In dit opzicht kan men nog niet van een juiste representatie spreken" (Directie van den Landbouw, 1940:30). Tabel 2.1 geeft inzicht in de verdeling van de bedrijven over de grootte-klassen. Uit deze tabel blijkt dat de kleinere bedrijven sterk ondervertegenwoordigd waren. Ook de veenkoloniale bedrijven waren sterk ondervertegenwoordigd in de vooroorlogse jaren (Directie van den Landbouw, 1940:30).

Met de toename van het aantal bedrijven dat een boekhouding leverde voor de bedrijfsuitkomstenstatistiek, daalde de gemiddelde oppervlakte per deelnemend bedrijf. In 1930 bedroeg de gemiddelde oppervlakte bijna 27 hectare. In de loop van de dertiger jaren kwam dit cijfer steeds dichterbij 20 hectare te liggen.

Bij de indeling van bedrijven in groepen in de jaarlijkse statistieken werden vijf "bedrijfssoorten" onderscheiden: zuiver akkerbouw (> 90% bouwland), overwegend akkerbouw (70-90% bouwland), gemengd bedrijf (30-70% bouwland), overwegend veeteelt (70-90% grasland) en zuiver veeteelt (> 90% grasland).

Vanaf boekjaar 1937 werd ook inzicht gegeven in de variatie in bedrijfsuitkomsten in de landbouw (Directie van den Landbouw, 1939:82). Er werd vastgesteld dat die variatie vrij groot was. Met het oog op de bedrijven die ver onder het gemiddelde zaten

Tabel 2.1 Het aantal "gecontroleerde" bedrijven (> 1 ha) in boekjaar 1938 en het aantal bedrijven waar het bedrijfshoofd als hoofdberoep "landbouwer" opgeeft in de telling van 1930, ingedeeld naar bedrijfsoppervlakte

Oppervlakte- klasse in ha	Gecontroleerde bedrijven in 1938		Bedrijven met "land- bouwers" in 1930	
	aantal	%	aantal	%
1 - 5	276	4,1	58.295	33,3
5 - 10	950	14,1	50.832	29,0
10 - 20	2.530	37,6	39.814	22,7
20 - 50	2.574	38,3	23.572	13,5
>= 50	398	5,9	2.512	1,5
Alle bedrijven	6.728	100,0 *)	175.025	100,0

*) Het betreft hier alleen bedrijven die in groepen vallen met meer dan tien verwerkte bedrijven; daarom is het totaal kleiner dan het aantal van 8.000 dat in paragraaf 2.2 genoemd wordt.

Bron: Directie van den Landbouw, 1940.

werd gesteld dat het duidelijk was dat hier voor de voorlichting nog een dankbare taak was weggelegd.

Vanaf boekjaar 1936 werd op ruim honderd zogenaamde "type-bedrijven" begonnen met een uitvoerige boekhouding. Deze bedrijven lagen verspreid over het gehele land en ze moesten bijdragen aan de systematische vergroting van het bedrijfseconomisch inzicht (Directie van den Landbouw, 1937:5-6). Het zijn de voorlopers van de latere LEI-bedrijven (zie hoofdstuk 3). Op deze bedrijven werden de kosten uitgeplitst naar de produkten waarvoor ze direct of indirect werden aangewend. Daardoor werd kostprijsberekening per produkt mogelijk. In de jaren vlak voor de oorlog zijn enkele tientallen boekhoudingen van "type-bedrijven" uitgewerkt (Horring, 1940).

2.4 De situatie in de veertiger jaren

In het voorlopig verslag over boekjaar 1942 wordt geconstateerd: "Het is inderdaad een feit, dat de bedrijfsuitkomsten, zooals deze worden besproken in "De Economische Toestand van den Landbouw" (...) in verschillende gevallen eenigszins als monsterd na den maaltijd komen, terwijl in de laatste jaren tengevolge van den oorlogstoestand ook in het verschijnen van deze publicatie een niet onaanzienlijke vertraging is ontstaan. Een op de lippen van alle aandachtige lezers liggende vraag is: "Zou het dan geen

aanbeveling verdienen de verwerking van de cijfers zoo te bespoe- digen, dat op het tijdstip der publicatie van de verkregen cij- fers deze laatsten nog voldoende recent zijn?" (Directie van den Landbouw, 1944:3).

Verwijzend naar de "goede medewerking der Landbouwboekhoud- bureaux" wordt geconstateerd dat het "eerder publiceeren der ver- kregen bedrijfsresultaten alleen mogelijk zal zijn, wanneer ge- werkt kan worden onder optimale omstandigheden, iets wat zelden voorkomt. Bij de samenstelling der prognose van de bedrijfsuit- komsten 1942-1943 is dan ook een geheel ander systeem gevolgd. Voor een betrekkelijk gering aantal bedrijven zijn de gemiddelde bedrijfsuitkomsten berekend per ha over de boekjaren 1941-1942 en 1942-1943. Om het gevaar van onjuistheden door het geringe aantal betrokken bedrijven te omzeilen, is steeds uitgegaan van dezelfde bedrijven". In het Voorbericht van deze publikatie wordt wel ge- waarschuwd om uit de cijfers niet te snel conclusies te trekken over de gang van zaken op het Nederlandse landbouwbedrijf in het algemeen, zeker niet waar het gaat om absolute grootheden. Met de gegevens van 925 bedrijven werden cijfers gegeven voor 43 groepen bedrijven.

Na de oorlog wordt de publikatie van bedrijfsuitkomsten voortgezet, maar zijn er wel enige organisatorische veranderingen doorgevoerd. In het "Verslag over de landbouw in Nederland" van het Ministerie dat in 1951 verschijnt en waarin een toelichting wordt gegeven op de "Bedrijfsuitkomsten over de jaren 1946, 1947 en 1948", wordt gezegd dat de verwerking van gegevens voor de be- drijfsuitkomstenstatistiek na 1941 is gestaakt en in 1946 weer is voortgezet. De verzameling, bewerking en publikatie van de gege- vens van de landbouwboekhoudbureaus wordt dan aan het inmiddels opgerichte LEI overgedragen. Onder de titel "Statistiek van de bedrijfsuitkomsten in de landbouw" publiceerde het LEI vanaf boekjaar 1946 jaarlijks de cijfers, terwijl het Ministerie in haar jaarlijkse verslag een beschouwing over die cijfers gaf.

Vanaf 1946 is er ook een wijziging in de manier waarop de bedrijfsuitkomsten worden bepaald. "De verdeling der in de sta- tistiek verwerkte bedrijven over de landbouwgebieden en grootte- klassen wordt thans niet meer bij toeval bepaald door de bij de boekhoudbureaux aangesloten aantallen bedrijven, maar van te vo- ren vastgesteld, teneinde een zo goed mogelijke evenredige verte- genwoordiging van de werkelijke aantallen bedrijven tegenover het geheel te verkrijgen. Voor 1946 en 1947 is daartoe een schema van 3.000 bedrijven opgesteld; voor 1948 en volgende jaren is dit aantal tot 5.000 uitgebreid. Ook bij deze opzet zijn de zeer kleine bedrijven nog onvoldoende vertegenwoordigd, daar deze slechts zelden bij een boekhoudbureau zijn aangesloten" (Ministe- rie van Landbouw, 1951:4; LEI, 1951:5-10). Verder werd een nieuwe indeling in landbouwgebieden, grootte-klassen en bedrijfstypen geïntroduceerd en werden de grootheden "gezinsinkomen" en "netto

overschot" centraal gesteld in de presentatie. Ook werd van de gelegenheid gebruik gemaakt om een aantal andere boekhoudkundige aanpassingen door te voeren. De vergelijkbaarheid van voor- en naoorlogse gegevens werd daardoor moeilijk. Koppejan (1949) heeft een poging gedaan om te laten zien hoe de gegevens op elkaar aansluiten.

Bij de samenstelling van groepen bedrijven werd uitgegaan van een minimum van vijftien bedrijven per bedrijfsgroep (een grootteklasse in een landbouwgebied) om enigszins betrouwbare cijfers per groep te krijgen.

Een vergelijking van de cijfers van de boekhoudingen met die van de landbouw-inventarisatie van 1947 leerde dat de gemiddelde bedrijfsgrootte 40-90% boven het gemiddelde lag. De kleinere bedrijven bleken nog sterk ondervertegenwoordigd te zijn. De meeste bedrijfsuitkomsten werden daarom nog steeds in gulden per hectare gegeven. Deze wijze van presenteren had echter ook te maken met de tijdrovendheid van het berekenen van individuele bedrijfsgegevens die vervolgens weer gemiddeld moesten worden per groep.

2.5 Twee bedrijfsuitkomsten-statistieken

Al vrij snel na 1946 stopte het LEI met het publiceren van de "Statistiek van de bedrijfsuitkomsten in de landbouw". In 1951 verscheen de laatste statistiek, die betrekking had op het boekjaar 1949. In het "Woord vooraf" staat te lezen dat met ingang van boekjaar 1950 de publikatie van de bedrijfsuitkomstenstatistiek uit gegevens van de boekhoudbureaus zou worden stopgezet. De oorzaak daarvan was dat het LEI jaarlijks twee overzichten van bedrijfsuitkomsten publiceerde. Het ene was de hierboven besproken statistiek, het andere was een eigen statistiek waarin gemiddelde uitkomsten van een aantal groepen bedrijven werden gepubliceerd, gebaseerd op boekhoudingen met een zuiver economische grondslag. Het LEI had daarvoor zelf bedrijven in administratie (zie hoofdstuk 3). De bedoeling was dat de boekhoudingen uit beide bronnen, in elk geval op het punt van de aard van de gegevens, vergelijkbaar zouden zijn. Al snel bleek dat er nogal wat verschillen bestonden in de wijze waarop de uitkomsten berekend werden, onder andere ten aanzien van de posten onderhoud, afschrijvingen, overige kosten en voorraad-waardering. "Bij gehouden overleg bleek verder, dat deze verschillen in principe wel zouden kunnen worden opgelost. Dit kon echter, in verband met de omvangrijke benodigde informatie per bedrijf, slechts met zo hoge extra kosten geschieden, dat een bedrijfsuitkomstenstatistiek van enigszins behoorlijke omvang niet meer bereikbaar zou zijn.

Er bleef dus slechts de keuze over tussen handhaven van de huidige, als weergave der bedrijfsuitkomsten in de landbouw, niet bevredigende statistiek en het beëindigen daarvan. Het handhaven van twee naast elkaar lopende reeksen van bedrijfsuitkomsten - welke in beginsel vergelijkbaar moeten zijn maar dit in feite

niet zijn - opgesteld door hetzelfde Instituut werd ongewenst geacht" (LEI, 1951). Het Bestuur van het LEI besloot daarom te stoppen met de samenstelling en publikatie van de bedrijfsuitkomstenstatistiek. "In overweging is evenwel, in hoeverre door een andere instantie de verwerking en publikatie van de - ongewijzigde - gegevens der landbouwboekhoudingen kan worden voortgezet. Een dergelijke statistiek zou dan uitkomsten, naar fiscaal doel berekend, weergeven" (LEI, 1951).

Het CBS heeft vanaf 1950 deze statistiek voortgezet en publiceert nog steeds bedrijfsuitkomsten (de exploitatierekeningen op fiscale basis) op basis van boekhoudingen van landbouwboekhoudbureaus (Bedrijfsuitkomsten, 1988:172).

3. Gegevensverzameling door het LEI vanaf 1940

3.1 Inleiding

In dit hoofdstuk wordt beschreven hoe het LEI tot 1975 bedrijfsgegevens verzamelde en daarover publiceerde. In deel 2 komt de periode na 1975 aan de orde. De periode 1940-1974 is te verdelen in twee perioden. Tot 1966 was er sprake van een selecte keuze van bedrijven via keuze-commissies, terwijl daarna de keuze aselect plaats vond door middel van een steekproef.

"Gedurende de boekjaren 1947-1962 werd door het LEI jaarlijks een overzicht van bedrijfsuitkomsten gepubliceerd op basis van door het LEI bijgehouden boekhoudingen. Deze reeks werd gestaakt toen door het Centraal Bureau voor de Statistiek in plaats van een bedrijfsuitkomstenstatistiek op basis van fiscale boekhoudingen, werd begonnen met een jaarlijkse bedrijfsuitkomstenstatistiek op basis van bedrijfseconomische boekhoudingen van zowel landbouwboekhoudbureaus als het LEI" (Bedrijfsuitkomsten, 1970). Het CBS verzorgde deze statistiek van 1963-1966. Daarna nam het LEI weer de jaarlijkse publikatie van bedrijfsuitkomsten ter hand met behulp van de informatie van de bedrijven die het LEI in eigen administratie had.

3.2 De bedrijfskeuze tot 1966

Het doel van de Statistiek van de bedrijfsuitkomsten in de landbouw - waarover het vorige hoofdstuk ging - was het geven van een overzicht van de Nederlands landbouw in al zijn verscheidenheid (LEI, 1951:7). Daarom werd een uitgebreide indeling van bedrijven naar landbouwgebieden, grootte-klassen en bedrijfstypen gepresenteerd.

Het doel van het bijhouden van boekhoudingen door het LEI zelf, was van een andere aard. Daarbij stond het verkrijgen van basismateriaal voor bedrijfseconomische studie (kostprijsberekening en bedrijfsanalyse) en voorlichting voorop (LEI, 1954:10; Koppejan, 1949:12). Daarom was de bedrijfskeuze voor de LEI-bedrijven gericht op het redelijk goed geleide bedrijf (Ministerie van Landbouw, 1951:7-8). "Bij de landbouwboekhoudbureaux zijn de bedrijven hoofdzakelijk aangesloten ter verlichting van de administratieve taak voor belastingdoeleinden. De deelnemers aan de bedrijfsboekhouding van het LEI getroosten zich evenwel veel extra moeite uit belangstelling voor de bedrijfsstudie. Het is mogelijk, dat daardoor de deelnemers aan de LEI-boekhouding gemiddeld betere bedrijven en dus ook betere uitkomsten hebben, dan de bedrijven, die bij de boekhoudbureaux zijn aangesloten. Ten slotte zijn de bedrijven van de LEI-boekhouding doorgaans groter en

is de verdeling binnen de landbouwgebieden minder representatief dan die van de bedrijven die in de statistiek der bedrijfsuitkomsten zijn opgenomen" (LEI, 1949:13).

Het LEI-boekhoudnet kwam in 1941 van de grond. In het eerste jaar werden ruim 500 bedrijven in administratie genomen en het lag in de bedoeling om dit aantal in de volgende jaren te vergroten. Door de oorlogsomstandigheden was er in 1945 nog maar een klein deel van de oorspronkelijk in administratie genomen bedrijven over. De uitwerking van de boekhoudingen liep grote achterstand op. Bij een bombardement in 1945 ging alle documentatie verloren (LEI, 1950). In 1946 werd een nieuwe start gemaakt met het boekhoudnet. De opgedane ervaring zorgde er voor dat relatief snel een goed boekhoudnet kon worden opgezet. In 1948 werden al meer dan 1.000 boekhoudingen bijgehouden en twee jaar later waren dat er ruim 1.700.

Voor de eerste naoorlogse jaren van het LEI-boekhoudnet werden eerst overzichten per bedrijfstype gepubliceerd, waarna een samenvattend verslag over alle groepen gezamenlijk werd opgesteld. Met ingang van boekjaar 1952 werd vrij snel na afloop van een boekjaar het "Statistisch overzicht van de gemiddelde uitkomsten van landbouwbedrijven" gepubliceerd, waarbij de toelichtende tekst tot een minimum beperkt bleef. Het overzicht over 1952 kon in april 1954 gepresenteerd worden (LEI, 1954). "Het cijfermateriaal in de tabellen is geheel afkomstig uit de door ons in samenwerking met de Rijkslandbouwconsulenten bijgehouden boekhoudingen" (LEI, 1954:7). Er werd op gewezen dat de gepubliceerde gemiddelde cijfers niet in alle opzichten representatief geacht konden worden voor het gemiddelde van alle bedrijven in de betrokken gebieden. Dat gold vooral voor de absolute hoogte van de cijfers en in veel mindere mate voor ontwikkelingen in cijfers.

Voor ongeveer veertig groepen bedrijven - ingedeeld naar regio en bedrijfstype - werden cijfers over kosten en opbrengsten per hectare cultuurgrond weergegeven voor het betreffende boekjaar en de drie daaraan voorafgaande boekjaren. Per groep bedrijven werden die bedrijven buiten beschouwing gelaten, waarvan de bedrijfsvoering te veel van het normale in een bepaald gebied afweek. Maar "ondanks het buiten beschouwing laten van sterk afwijkende bedrijven moet gezegd worden dat de cijfers niet geheel representatief zijn voor een gebied" (LEI, 1954:9-10). Uit onderzoek (LEI, 1952a, 1952b) was gebleken dat de LEI-bedrijven op enkele punten afweken van de populatie die ze moesten beschrijven. Niet duidelijk was welke invloed dit had op de bedrijfsuitkomsten.

Voor een groep bedrijven werden alleen uitkomsten gegeven indien daarvoor meer dan vijftien boekhoudingen beschikbaar waren. Voor boekjaar 1952 werden 1.739 bedrijven in administratie genomen, waarvan er 1.092 een uitgewerkte boekhouding leverden. De watersnoodramp in 1953 was mede debet aan het grote aantal niet-uitgewerkte boekhoudingen.

In het "Woord vooraf" van het "Statistisch overzicht 1960/'61 van de uitkomsten van landbouwbedrijven" - dat in december 1961 (!) verscheen - blijkt dat de doelstelling van het boekhoudnet dan geformuleerd wordt als het verschaffen van gegevens voor de LEI-rapporten ten behoeve van het landbouwbeleid. Voor 52 groepen bedrijven werden cijfers gepresenteerd met behulp van 1.430 boekhoudnet-bedrijven. Er werden naast de kosten en opbrengsten per hectare of per koe ook meer absolute cijfers per bedrijf gegeven.

De bedrijven die het LEI in administratie nam werden aange-
wezen door regionale keuze-commissies, waarin vertegenwoordigers van het Landbouwschap, de Rijkslandbouwvoorlichting en het LEI zitting hadden. Elk van deze drie partijen moest instemmen met de opname van een bepaald bedrijf in het boekhoudnet (Snoek, 1963). Voor iedere groep bedrijven die onderscheiden werd, werden een aantal "representatieve" of "typische" bedrijven gekozen. Extreem goede of extreem slechte bedrijven werden niet uitgekozen. Alleen bedrijven die qua produktieomstandigheden en bedrijfsvoering "typisch" ("normaal") werden geacht voor een bepaalde regio, werden gekozen. Bovendien werd geprobeerd om bedrijven te kiezen die geleid werden door "vakbekwame" boeren (LEI, 1961:5). Bedrijven werden voor een periode van vijf jaar in administratie genomen.

Hoewel de vrij subjectieve manier van kiezen in de vijftiger en zestiger jaren een aantal nadelen met zich meebracht - waarop in de volgende paragraaf zal worden ingegaan - waren er ook voordelen aan verbonden. "Bij de bedrijfskeuze, die steunt op lokale kennis van de leden van de keuze-commissies zijn misschien de mogelijkheden om vertekening als gevolg van het optreden van weigeraars te voorkomen, groter. Bovendien wordt door het inschakelen van keuze-commissies de bereidheid tot deelneming veelal bevorderd" (Snoek, 1963:8).

"Vandaar dan ook dat in het verleden - toen het aantal weigeraars te groot werd geacht om een steekproef te doen slagen - de LEI-procedure wel als de enig bruikbare methode moest worden beschouwd. Nu echter het LEI in de loop der jaren meer bekendheid heeft gekregen en althans in sommige gebieden daardoor de kans op een groot percentage "weigeraars" aanmerkelijk is verminderd (...) zou in verschillende gebieden ook de steekproefmethode in aanmerking kunnen komen" (Snoek, 1963:9).

Het doel van het verzamelen en bestuderen van de bedrijfsgegevens van landbouwbedrijven was volgens Snoek tweeledig. In de eerste plaats werden de deelnemende bedrijven in staat gesteld om door middel van bedrijfsvergelijking zich een beeld te vormen omtrent hun relatieve positie in de desbetreffende groep bedrijven om op die wijze verbeteringen in de bedrijfsvoering te kunnen aanbrengen. In de tweede plaats konden sommige uitkomsten - onder andere de berekende kostprijzen - door de overheid en de georganiseerde landbouw, gebruikt worden als basis bij de onderhande-

lingen over het in de landbouw te voeren (prijis-)beleid (Snoek, 1963:5).

3.3 De overgang op bedrijfskeuze via een steekproef

Voor het boekjaar 1961 werd parallel aan de gangbare keuze-procedure een steekproef getrokken voor een groep weidebedrijven in het consulentenschap Sneek. Op deze wijze kon worden nagegaan of er tussen de groep keuzebedrijven en de groep steekproefbedrijven significante verschillen bestonden ten aanzien van het gemiddelde van en de spreiding in de waarden van een aantal grootheden.

Conclusie van dit onderzoek was dat - gegeven de beperkte hoeveelheid onderzoeksmateriaal - er geen significante verschillen in de gemiddelde waarden van diverse grootheden tussen de LEI-groep en de steekproefgroep konden worden aangetoond. Wel waren er significante verschillen in variantie tussen de groepen, bijvoorbeeld ten aanzien van netto overschot per hectare, dichtheid van de veebezetting en kostprijs van de melk. Zoals verwacht kon worden, lag de spreiding in uitkomsten op de steekproefbedrijven duidelijk hoger dan op LEI-bedrijven. "Dit hoeft ons overigens ook niet te verbazen wanneer men bedenkt, dat de keuze-commissies er naar streven de extreem goede en extreem slechte bedrijven niet in de te formeren groep op te nemen; het lijkt er dus op, dat zij in deze opzet ook inderdaad zijn geslaagd" (Snoek, 1963:21).

Voor de jaren 1962-1964 publiceerde het LEI geen bedrijfsuitkomsten van landbouwbedrijven. Het CBS, dat jaarlijks een exploitatierekening op fiscale basis publiceerde aan de hand van de boekhoudingen van landbouwboekhoudbureaus - zie hoofdstuk 2 -, ging tijdelijk over op publikatie van deze gegevens op bedrijfs-economische basis. Dat betekende echter niet dat het LEI-boekhoudnet verdwenen was. In die jaren werd een nieuwe opzet van de bedrijfskeuze uitgewerkt. In 1966 werd begonnen met het kiezen van bedrijven op basis van een steekproef. Bedrijven die gekozen waren volgens de oude LEI-keuze-procedure, verdwenen in de jaren daarna geleidelijk uit het boekhoudnet. Van de bedrijven waarover werd gepubliceerd over het boekjaar 1969 was al 90% volgens een steekproef gekozen (Bedrijfsuitkomsten, 1971). In het boekjaar 1971 was de overgang naar de steekproefprocedure voltooid.

Doordat er niet meer werd gelet op de kwaliteit van het bedrijf en de doelmatigheid van de bedrijfsvoering, ontstond er een grotere spreiding in de samenstelling van het bedrijfsplan, de produktieomstandigheden, de bedrijfsgrootte en het bedrijfsresultaat. Het gemiddelde niveau van de cijfers in vergelijkbare groepen werd echter nauwelijks beïnvloed. De door Snoek (1963) getrokken conclusies voor een klein deel van de Nederlandse landbouw, bleken een vrij algemene geldigheid te bezitten.

3.4 Veranderende doelstellingen van het boekhoudnet

Tegelijk met de opzet van de bedrijfskeuze veranderde ook de doelstelling van het bijhouden van het boekhoudnet. De bedrijfskeuze werd gericht op representativiteit voor de Nederlandse landbouw als geheel en niet meer in de eerste plaats op bepaalde typen, grootte-klassen, grondsoorten of gebieden.

De gerichtheid op de Nederlandse landbouw als geheel werd ingegeven door de eisen van het EG-informatienet. Daarnaast bleef echter de doelstelling van het jaarlijks weergeven van gemiddelde cijfers van groepen bedrijven. Deze groepen bedrijven werden achteraf uit de voor het EG-informatienet gekozen steekproefbedrijven gevormd. De aantallen per groep schoten voor sommige groepen echter tekort, iets wat zich in het verleden niet voordeed, omdat toen eenvoudig een zodanig aantal bedrijven per groep werd gekozen dat een redelijke betrouwbaarheid van de cijfers per groep werd bereikt. In de nieuwe opzet werd dit bezwaar ondervangen door gerichte aanvullingen per groep op de coörspronkelijke EG-steekproef. Daardoor kon voor de groepen waarvoor dat - om welke reden dan ook - gewenst was, jaarlijks in de BUL een overzicht van de bedrijfsuitkomsten worden gegeven.

De BUL '70/'71 formuleert expliciet waarop het boekhoudnet gericht is: het beoogt "een inzicht te geven in de ontwikkeling van de rentabiliteit van de Nederlandse landbouw zowel ten behoeve van de overheid en het bedrijfsleven als van de EEG-Commissie in Brussel (...). De gewenste omvang en samenstelling van het LEI-boekhoudnet wordt in eerste instantie bepaald door de eisen die hieraan door de EEG worden gesteld. Alle bedrijven in de EEG-landen zijn ingedeeld in groepen naar bedrijfstype en oppervlakteklasse (...). Nederland dient aan de EEG minimaal 750 bedrijfs-economische boekhoudingen te leveren, die betrekking hebben op landbouwbedrijven van 5 ha en groter en waarvan de ondernemer van hoofdberoep landbouwer is. Deze bedrijven zijn gekozen uit groepen, waarvan landelijk meer dan 2.000 bedrijven voorkomen. Per groep dienen minimaal 25 boekhoudingen beschikbaar te komen.

Voor het verkrijgen van inzicht in de rentabiliteit van de Nederlandse landbouw ten behoeve van de overheid en het bedrijfsleven in Nederland worden ter aanvulling van het EEG-informatienet in het LEI-boekhoudnet bovendien boekhoudingen bijgehouden van bedrijven waarvan, op basis van de EEG-classificatie, er minder dan 2.000 per groep in Nederland voorkomen" (Bedrijfsuitkomsten, 1972:171).

3.5 De steekproefopzet vanaf 1966

Vanaf het begin van de steekproefsgewijze opzet werd een gestratificeerde steekproef getrokken, waarbij de afgegrensde populatie werd ingedeeld in acht groepen of strata. De akkerbouwbe-

drijven en de veehouderijbedrijven werden elk in vier grootte-
klassen ingedeeld (zie tabel 3.1). Het steekproefpercentage in
een bepaald stratum was afhankelijk van het bedrijfstype en van
de oppervlakteklasse, om zo te komen tot meer betrouwbare gemid-
delde bedrijfsuitkomsten. Omdat de spreiding in bedrijfsuitkom-
sten op akkerbouwbedrijven groter was dan die op veehouderijbe-
drijven, was het steekproefpercentage in akkerbouwstrata 1.5 maal
zo groot als in veehouderijstrata, uitgaande van dezelfde opper-
vlakteklaas. Daarnaast werd het steekproefpercentage groter ge-
nomen naarmate de bedrijfsoppervlakte groter was. Als reden hier-
voor werd gegeven dat de spreiding in uitkomsten binnen een be-
paald bedrijfstype toeneemt met de bedrijfsgrootte. De gehanteer-
de steekproefpercentages zijn vermeld in tabel 3.1, die is over-
genomen uit de BUL over boekjaar 1968 (Bedrijfsuitkomsten, 1970).
De ongelijke trekkingspercentages impliceerden wel dat bij de
weergave van de gemiddelde bedrijfsuitkomsten gewerkt moest wor-
den met wegingsfactoren.

*Tabel 3.1 De gehanteerde steekproefpercentages bij de keuze voor
het boekhoudnet naar bedrijfstype en grootte-klasse*

EEG-bedrijfstype	Grootte-klasse			
	5-10 ha	10-20 ha	20-50 ha	> 50 ha
Akkerbouw	0,84	0,92	1,68	2,52
Veehouderij	0,56	0,62	1,12	1,68

In boekjaar 1971 werden de in 1966 vastgestelde steekproef-
percentages per stratum iets aangepast, om aan de vereiste 750
steekproefbedrijven voor het EEG-informatienet te komen.

Met ingang van boekjaar 1972 werden de bedrijven met een om-
vang van 10-40 sbe niet meer in het boekhoudnet opgenomen - ook
niet indien ze meer dan 5 hectare groot waren. Daardoor werd het
aantal bedrijven in de populatie met ruim 8.000 verminderd.

3.6 De steekproeven voor de boekjaren 1968-1974

Aanvankelijk werden de bedrijven niet rechtstreeks vanuit de
meitelling in de steekproef getrokken. Voorafgaande aan de eigen-
lijke trekking werd namelijk een steekproef uit de "mei-inventa-
risatie-formulieren" getrokken. Hieruit werden dan - na indeling
van deze steekproef in strata - bedrijven gekozen voor deelname
aan het boekhoudnet. Op basis van een steekproef uit de meitel-
ling 1966 van 3.430 bedrijven, werden 1.974 bedrijven bezocht
voor deelname aan het boekhoudnet. Dit resulteerde uiteindelijk

in 822 uitgewerkte bedrijfsboekhoudingen in het boekjaar 1968. Van de 1.974 aangewezen bedrijven bleek 19% niet voor deelname geschikt of tot deelname in staat te zijn. De belangrijkste reden hiervoor was dat de betreffende bedrijven inmiddels opgeheven waren of binnen enkele jaren opgeheven zouden worden. Van de overblijvende 1.601 bedrijven weigerde 30% mee te doen aan het boekhoudnet, met name vanwege het extra werk dat aan de boekhouding verbonden zou zijn en het ontbreken van belangstelling voor deelname. Er resteerden 1.119 bedrijven die bereid waren aan de boekhouding deel te nemen. Hiervan vielen gedurende de boekjaren 1967 en 1968 nog 297 bedrijven af wegens onvoldoende belangstelling of het niet insturen van de nodige gegevens.

Om te voorkomen dat er te weinig uitgewerkte boekhoudingen zouden zijn aan het eind van een boekjaar, werd na enkele jaren de regel gehanteerd dat minstens 120% van het gewenste aantal bedrijven in administratie werd genomen. Als basis voor de keuze per 1 mei van een boekjaar diende een steekproef van 7.000 à 8.000 bedrijven uit de metelling van een of twee jaar geleden.

De BUL '68/'69 geeft een overzicht van de bedrijfsuitkomsten in de boekjaren 1965-1968. Hoewel in boekjaar 1967 al een aantal bedrijven steekproefsgewijs werd gekozen, berusten de cijfers t/m boekjaar 1967 nog op ongewogen gemiddelden. Vanaf boekjaar 1968 werd met wegingsfactoren gewerkt. In de BUL werden gemiddelde uitkomsten voor 22 groepen bedrijven gepubliceerd. Er werd gestreefd naar een minimum aantal van veertig steekproefbedrijven per groep, "waarbij op grond van de ervaringen een redelijke mate van betrouwbaarheid mag worden verwacht" (Bedrijfsuitkomsten, 1970:10). Vanaf boekjaar 1970 werd gestreefd naar een minimum-aantal van 35 bedrijven per groep" (Bedrijfsuitkomsten, 1972:13).

De indeling in groepen gebeurde op grond van criteria "waarvan het sterkste onderscheidend vermogen ten aanzien van verschillen in de bedrijfsvoering mag worden verwacht". Verder werd gelet op homogeniteit ten aanzien van de verschillende bedrijfs-onderdelen binnen een groep. "Op grond hiervan is besloten relatief veel ruimte te besteden aan de indeling naar bedrijfstype en minder aan een vergaande regionale indeling. Bij gegeven opbouw van het bedrijfsplan blijken de gebiedsverschillen namelijk relatief onbelangrijk te zijn", behalve voor de akkerbouw.

Naast een indeling naar bedrijfstype en regio werd bij de publikatie van groepscijfers een indeling gemaakt naar grotere en kleinere bedrijven, waarbij de grens tussen groot en klein lag bij een bedrijfs grootte van 90 sbe 1) in 1968.

1) sbe- standaardbedrijfseenheden. De sbe is een verhoudingsgetal voor de omvang van een agrarische activiteit op basis van de toegevoegde waarde onder genormaliseerde omstandigheden bij een doelmatige bedrijfsvoering in een basisperiode (Dröge, Poppe en Prins, 1983).

Tabel 3.2 Normen voor het sbe-aandeel per produktierichting bij de indeling van bedrijven naar bedrijfstype in de BUL

Bedrijfstype	Produktierichting		
	akkerbouw (*)	rundveehouderij (incl. grasland)	veredeling (**)
Akkerbouwbedrijven	>= 80%		
Weidebedrijven		>= 80%	
Gemengde bedrijven			
a) met overwegend akkerbouw	40-80%		< 25%
b) met overwegend rundvee	< 40%	< 80%	< 25%
c) met relatief veel veredeling			>= 25%

*) In verband met het beperkte aantal bedrijven werd in afwijking van dit schema de grens van 80% later op 67% gesteld.

***) Vanaf boekjaar 1970: 30% in plaats van 25%.

"De indeling naar bedrijfstype werd gebaseerd op de procentuele verdeling van het aantal sbe over de bedrijfsonderdelen akkerbouw, rundveehouderij en veredeling (varkens- en pluimveehouderij) in het betreffende boekjaar. Het percentage in overige produktierichtingen (bijvoorbeeld fruitteelt of werk voor derden) is daarbij niet in aanmerking genomen". Tabel 3.2 geeft de indeling weer.

Vanaf boekjaar 1970 werd bij de bedrijfstypering bij de "gemengde bedrijven met aanmerkelijke veredeling" een grens van 30% in plaats van 25% gehanteerd. "Daarmee wordt de ontwikkeling naar grotere eenheden in varkenshouderij en pluimveehouderij op het gemengde bedrijf gevolgd" (Bedrijfsuitkomsten, 1972:13-16).

De indeling in regio's was gebaseerd op de LEI-indeling in 13 gebieden, die weer was gebaseerd op de indeling in 121 landbouwgebieden uit 1957.

In boekjaar 1970 werden 22 BUL-groepen gevormd. In de daarop volgende jaren moesten de twee regionale groepen gemengde bedrijven met overwegend rundveehouderij tot één groep worden samengevoegd, omdat ze afzonderlijk te weinig bedrijven bevatten. Vanaf 1973 kon onder invloed van structurele ontwikkelingen de groep grotere weidebedrijven in het Zuidelijk Zandgebied worden afgesplitst van die in het Oostelijk en Centraal Zandgebied. Voor sommige groepen bedrijven konden alleen cijfers worden gegeven dankzij de al genoemde aanvullingen op de voor de EEG noodzakelijke aantallen bedrijven. Aanvullingsbedrijven werden onder andere gekozen in het Noordelijk Veenweidegebied (weidebedrijven), het Noordelijk Kleigebied (akkerbouwbedrijven), en later ook in het Westelijk Weidegebied (weidebedrijven). Vanaf boekjaar 1974

warden bovendien extra bedrijven gekozen in de groep akkerbouwbedrijven in de Veenkoloniën en het Noordelijk Zandgebied.

In de BUL werd jaarlijks een indicatie gegeven van de spreiding in bedrijfsuitkomsten. Per groep werd de standaardafwijking weergegeven van het netto overschot (op pachtbasis) in guldens per hectare cultuurgrond, van het aantal sbe per vak en van het arbeidsinkomen van de ondernemer (op pachtbasis) in guldens per bedrijf. Met ingang van 1974 werden in de bijlage van de BUL geen *standaardafwijkingen* voor variabelen meer gegeven - om de spreiding in de waarden van een variabele aan te geven - , maar werden *standaardfouten* gepresenteerd, om een indruk te geven van de betrouwbaarheid van de geschatte waarden.

Over de opbouw van het boekhoudnet zegt de BUL over boekjaar 1974 het volgende: "Wat betreft het aantal jaren van deelname wordt er gestreefd naar een homogene opbouw van het deelnemersbestand. Op grond van ervaringscijfers is gebleken dat, indien met 100 gekozen bedrijven wordt begonnen, gemiddeld na verloop van één jaar als gevolg van opzegging door de deelnemers zelf (geen belangstelling meer enz.), er nog 77 bedrijven over zijn. Na twee jaar is het aantal teruggelopen tot 68, na drie jaar tot 61, na vier jaar tot 55, na vijf jaar tot 49 en na zes jaar tot 44 bedrijven. Indien nu de gemiddelde duur van deelneming aan de LEI-administratie wordt gesteld op zes jaar betekent dit dat bij een evenwichtige opbouw van het deelnemersbestand elk boekjaar moet worden begonnen met circa 23% eerste jaars deelnemers" (Bedrijfsuitkomsten, 1976:172).

Met ingang van 1974 verviel de eis dat het hoofdberoep van de ondernemer landbouwer moest zijn. Een landbouwbedrijf moest vanaf dat jaar voldoen aan de eis dat er minder dan 50% tuinbouw - in sbe's gemeten - voorkwam. De steekproef uit de meitelling waarop de eigenlijke steekproef werd gebaseerd, veranderde in 1974 enigszins van opzet. Aanvankelijk werd 5% van de bedrijven in de oppervlakteklasse 5-20 ha. in de eerste steekproef getrokken en 10% van de bedrijven in oppervlakteklasse 20-100 ha en verder alle bedrijven groter dan 100 ha. In 1974 werd uit de grootteklassen 40-110 sbe, 110-250 sbe, 250-500 sbe en meer dan 500 sbe respectievelijk 4,5%, 9%, 18% en 100% van de bedrijven in Nederland in de steekproef opgenomen. "Hiermede is dus als basis voor de steekproef de voorheen toegepaste indeling in oppervlakteklassen vervangen door een in sbe-klassen omdat met name ook de bedrijfsomvang van bedrijven met intensieve veehouderij aldus is te bepalen" (Bedrijfsuitkomsten, 1976:166). De trekking van de steekproefbedrijven uit dit steekproefbestand van circa 7.500 bedrijven, gebeurde met de in 1966 geformuleerde verhoudingsgetallen tussen de steekproefpercentages in de 8 EEG-strata die onderverdeeld werden in sbe- en gebieds-sub-strata.

Achteraf werden alle bedrijven opnieuw ingedeeld in strata waarbij als criteria het bedrijfstype, de bedrijfsomvang in hectares, de bedrijfsomvang in sbe en het gebied werden gehanteerd. Dit laatste criterium was noodzakelijk vanwege de aanvullingskeuze die vooral op bepaalde regio's was gericht.

Naast de BUL verscheen jaarlijks "De financiële positie van landbouwbedrijven" (FIP). De cijfers in deze publikatie waren gebaseerd op dat deel van de steekproefbedrijven dat in staat en bereid was een zogenaamde financieringsboekhouding bij te houden. De FIP sloot qua populatie en groepsindeling nauw aan bij de BUL (Aukema en Donkelaar, 1972; Aukema, 1974).

3.7 Representativiteit

In de BUL werd geconstateerd: "In hoeverre de matige response, die overigens praktisch onontkoombaar lijkt, de representativiteit verstoort, is onzeker" (Bedrijfsuitkomsten, 1970:141). Er werd op gewezen dat de leeftijdsopbouw in het boekhoudnet iets afweek van die in de populatie; oudere ondernemers waren ondervertegenwoordigd. Ten aanzien van de bedrijfsoppervlakte, het bedrijfstype en de regionale ligging werd de representativiteit zo goed mogelijk gehandhaafd door bij het niet kunnen of willen meewerken van een bedrijfshoofd een vervangend bedrijf te kiezen dat zoveel mogelijk leek op het bedrijf dat vervangen werd.

Bovendien werden de wegingsfactoren berekend op grond van de aantallen in de populatie bij de metelling (aan het begin van het betreffende boekjaar) en de aantallen *uitgewerkte* boekhoudingen per stratum. Daardoor werd ook het effect van tussentijdse beëindigingen - voorzover die samenhangen met de factoren die bij de stratificatie een rol speelden - gecompenseerd. Dit had verder als voordeel dat in de overgangsfase naar steekproeftrekking, ook bedrijven die op grond van de oude LEI-procedure waren gekozen, konden worden meegenomen in de berekeningen.

"Een inzicht in eventuele verstoringen als gevolg van samenhangen tussen enerzijds de redenen tot weigering van deelneming of oorzaken van voortijdige beëindiging, en anderzijds productieomstandigheden en kwaliteit van de bedrijfsvoering, die niet gevangen worden in de aangehouden EEG-klassenindeling naar bedrijfstype en regionale ligging, kan niet worden gegeven" (Bedrijfsuitkomsten, 1970:141).

In de BUL over boekjaar 1970 werd geconstateerd dat per oppervlakteklasse de bereidheid om aan het boekhoudnet deel te nemen groter was naarmate de bedrijfsvoering intensiever was. Daarom werd achteraf naar bedrijfstype, naar grootteklasse en naar intensiteitsklasse gestratificeerd en gewogen. Bij deze weging werden steekproefbedrijven ingedeeld in strata op grond van hun kenmerken gedurende het boekjaar en populatiebedrijven op grond van hun kenmerken aan het begin van het boekjaar. Hoewel dit praktisch gezien een werkbare procedure was, was het theoretisch niet geheel correct. Bij de bespreking van de steekproefopzet vanaf 1975 (deel 2) zal blijken dat vanaf dat jaar een betere wegingsprocedure wordt toegepast, waarbij de metelling ook de basis voor de stratificatie van de steekproefbedrijven is.

De BUL over boekjaar 1972 geeft iets meer informatie over de representativiteit van de bedrijven in het boekhoudnet voor de bedrijven in de populatie: "Op de in totaal bezochte 755 bedrijven is de mediaan van de leeftijd van het bedrijfshoofd 51 jaar en op de 281 bedrijven die bereid waren deel te nemen 46 jaar" (Bedrijfsuitkomsten, 1974:197). Omdat er bij vervangende keuze geen rekening werd gehouden met de leeftijd, betekende dit dat de oudere ondernemers ondervertegenwoordigd werden. In de BUL '73/'74 en '74/'75 werd iets uitvoeriger ingegaan op de representativiteit van de gekozen bedrijven, met name ten aanzien van de leeftijd van het bedrijfshoofd en de intensiteit van de bedrijfsvoering. Geconcludeerd werd dat bedrijven die niet geschikt waren voor of niet in staat waren tot deelname, een bedrijfshoofd hadden dat gemiddeld ouder was en een bedrijfsplan dat gemiddeld minder intensief was (minder sbe per ha) dan alle bezochte bedrijven. "Dit sluit dus aan bij de redenen van het niet geschikt of niet in staat zijn, te weten hoge leeftijd, ziekte en/of geen opvolger of het tevens nog uitoefenen van een ander beroep". Voor medewerking weigerende bedrijven kon een zelfde conclusie worden getrokken.

3.8 Verandering van de opzet in 1975

Vanaf 1975 werd de ondergrens van de populatie niet meer gelegd bij 5 ha, maar bij 60 sbe. Bovendien werd een nieuw stratificatieschema geïntroduceerd, waarin achtereenvolgens naar bedrijfstype en sbe-grootteklasse en waar nodig en mogelijk ook naar oppervlakteklasse, leeftijdsklasse van de ondernemer en naar gebied werd gestratificeerd. Daarmee werd gestreefd naar een grotere representativiteit. Dit schema is de basis voor de keuze van bedrijven geweest tot en met boekjaar 1987. In deel 2 van dit onderzoekverslag wordt een uitvoerige beschrijving en evaluatie van de steekproefopzet en van de kwaliteit van de gegevens uit de steekproef in de periode 1975-1987 gegeven.

Uit de BUL '75/'76 blijkt dat de nieuwe stratificatie-procedure met name consequenties had voor de groep gemengde bedrijven met intensieve veehouderij. In het Noordelijk, Oostelijk en Centraal Zandgebied ging het gemiddeld aantal sbe per hectare cultuurgrond bij de grotere bedrijven van 14,4 in 1974 naar 22,6 in 1975; voor het Zuidelijk Zandgebied is er eveneens een forse stijging in de waarde van dit kengetal. Ook het aantal bedrijven in deze categorieën nam behoorlijk toe. Soortgelijke ontwikkelingen waren te zien bij de groep kleinere intensieve veehouderijbedrijven.

Hoofdstuk 1 van de BUL is sinds 1975 nauwelijks gewijzigd. Daarmee is goed aangegeven dat er vanaf dat jaar sprake was van continuïteit in de opzet van de steekproef en de publikatie van bedrijfsuitkomsten op grond van die steekproef.

**Deel 2: de steekproef in de boekjaren
1975 tot en met 1987**

*"Vanaf 1967 is de keuze van de bedrijven voor het boekhoud-
net volledig gericht op representativiteit voor de Neder-
landse landbouw"*

(Bedrijfsuitkomsten, 1988:9)

*"Er is geen behoefte aan de term representatieve steekproef;
deze term is niet vatbaar voor een (scherpe) definitie en
leidt veelal tot verwarring"*

(J. Muilwijk, 1981:19)

4. Doelstellingen en middelen

4.1 Inleiding

Vanuit onderzoek en beleid is er behoefte aan cijfermateriaal over de ontwikkeling van de bedrijfsuitkomsten van Nederlandse landbouwbedrijven en over de financiële positie van de ondernemers op die bedrijven. De gewenste informatie heeft zowel betrekking op de totale Nederlandse landbouw als op groepen bedrijven binnen de landbouwsector. In het vorige hoofdstuk bleek dat over groepen van bedrijven - bijvoorbeeld over een bedrijfstak of over een groep bedrijven in een bepaalde regio - al tientallen jaren informatie wordt verzameld door het LEI. De gerichtheid op de Nederlandse landbouw in zijn geheel ging vanaf het midden van de jaren zestig een rol spelen. EG-verplichtingen lagen ten grondslag aan deze koerswijziging.

Vanaf 1966 en nog sterker vanaf 1975 werd weergave van de totaliteit van de Nederlandse landbouw de hoofddoelstelling van de informatieverzameling op landbouwbedrijven. Dat hield overigens niet in dat veel afbreuk werd gedaan aan de oorspronkelijke doelstelling van gegevensverzameling over groepen bedrijven. In dit hoofdstuk wordt een beschrijving gegeven van de hoofddoelstelling van het boekhoudnet in de periode 1975-1987. De opzet van de steekproef op basis van deze doelstellingen komt aan bod in de hoofdstukken 5 en 6. De verdere concretisering van de neven-doelstellingen - informatieverzameling over groepen bedrijven binnen de populatie - en de uitwerking daarvan in de richting van de steekproefopzet, vindt plaats in hoofdstuk 7.

4.2 De doelpopulatie

Het LEI-boekhoudnet richt zich - evenals het EG-informatienet - niet op de gehele populatie van landbouwbedrijven, maar op bedrijven met een bepaalde minimale omvang. De reden daarvoor is dat het bijhouden van boekhoudingen van de kleinste bedrijven dermate veel kosten met zich meebrengt dat het - mede gezien het geringe aandeel van deze groep bedrijven in de totale landbouwproductie, de heterogeniteit van deze groep, en de geringe bereidheid tot deelname in deze groep - niet mogelijk of niet rationeel is (gezien de informatiewaarde) ook van die groep bedrijven boekhoudingen bij te houden.

Tot 1982 werd Nederland geacht zich te richten op de populatie van bedrijven met meer dan 0,75 arbeidsjaareenheden (AJE) (Poppe, 1984:71). Vanaf 1982 is de benedengrens vastgesteld op 6 Europese Grootte Eenheden (EGE) (zie voor dit begrip Poppe (1984)) Per 1 januari 1986 heeft de EG de benedengrens voor het

Nederlandse boekhoudnet vastgesteld op 16 EGE-80. Dit komt grofweg overeen met 60 sbe (Poppe, 1986).

De populatie waarop het Nederlandse boekhoudnet voor landbouwbedrijven zich sinds 1975 richt, komt wel in grote lijnen overeen met de door de EG aangegeven populatie, maar valt daar niet geheel mee samen. Als ondergrens werd voor het boekjaar 1975 gekozen voor 60 sbe. Deze grens is van jaar tot jaar verhoogd en staat voor het boekjaar 1987 op 79 sbe. Deze grens komt bij benadering overeen met de helft van de produktie-omvang per man die gemiddeld bij een doelmatige bedrijfsgrootte wordt gerealiseerd. Het boekhoudnet van tuinbouwbedrijven legt de ondergrens bij 70 sbe (van der Plas, 1985).

Een afgrenzing ten opzichte van het boekhoudnet voor de tuinbouw wordt gemaakt door bedrijven waar meer dan 50% van het totale aantal sbe betrekking heeft op tuinbouw-activiteiten, uit te sluiten van de populatie 1). Bovendien worden - om praktische redenen - bedrijven met meer dan 2.000 sbe niet in de populatie opgenomen. Om te beoordelen of een bedrijf aan de populatie-criteria voldoet, wordt uitgegaan van de meitellingsgegevens aan het begin van een boekjaar.

De doelpopulatie bevatte voor het boekjaar 1975 ruim 80.000 bedrijven. Tien jaar later was dit aantal gedaald tot minder dan 70.000 bedrijven. Ruim een derde deel van de bedrijven in de meitelling waarop landbouw de hoofdactiviteit is, valt buiten de populatie. Deze bedrijven omvatten echter minder dan 10% van de Nederlandse landbouw, indien deze in sbe's gemeten wordt. Het aandeel van deze bedrijven in de totale landbouwproduktie is waarschijnlijk nog kleiner. Voor een verdere verantwoording van de begrenzing van de steekproefbasis, zij verwezen naar Lodder (1987:12-13).

4.3 De doelvariabelen

Naast een doelpopulatie ("over wie wil je iets weten?") heeft iedere informatieverzameling te maken met doelvariabelen ("wat wil je weten?"). De gewenste informatie van het landbouwnet heeft betrekking op de bedrijfsuitkomsten van bedrijven en de financiële positie van ondernemers uit de doelpopulatie. Voor het opzetten van een steekproef ter verkrijging van deze informatie, is het noodzakelijk om deze enigszins vage doelstelling nader te concretiseren; met name om eisen ten aanzien van de te verzamelen informatie in de steekproefopzet te kunnen meenemen (zie paragraaf 4.4). Een steekproef die kwalitatief goede uitkomsten geeft voor de ene (groep van) variabele(n), geeft niet automatisch goede uitkomsten voor de andere (groep van) variabele(n).

1) Voor een precieze beschrijving van deze afgrenzing, zie bijlage 1.

Vanaf 1975 zijn de normatieve toegevoegde waarde per bedrijf en de arbeidsopbrengst van de ondernemer de doelvariabelen van het boekhoudnet (Lodder, 1987:19). De eerste variabele als representant van de groep doelvariabelen in de sfeer van de bedrijfsuitkomsten en de tweede als representant van variabelen die betrekking hebben op de financiële positie van de ondernemer.

4.4 Kwaliteit van de informatie

4.4.1 Inleiding

De uitkomsten die een steekproef oplevert, moeten aan bepaalde eisen voldoen. De LEI-steekproef is er op gericht om zowel systematische als toevallige afwijkingen tussen steekproef en populatie terug te dringen. Anders gezegd, de LEI-steekproef streeft naar een zo groot mogelijke representativiteit en betrouwbaarheid.

4.4.2 Representativiteit

De steekproefbedrijven moeten representatief zijn voor de populatie waarop ze betrekking hebben. Dat wil zeggen dat ze een zo getrouw mogelijke afspiegeling geven van de situatie in de populatie. Representativiteit is een begrip dat vaak verwarring schept. Sommige onderzoekers willen het begrip daarom helemaal niet gebruiken (Moors en Muilwijk, 1975). Theoretisch gezien is dat ook mogelijk; aselechte trekking van steekproef-elementen garandeert immers "representativiteit". In de praktijk van het werken met steekproeven is het - onder andere vanwege het optreden van non-respons - echter nuttig om iets meer te zeggen dan dat er aselechte trekking heeft plaats gevonden. Representativiteit is een begrip dat daar behulpzaam bij kan zijn. Het geeft aan in hoeverre er achteraf gezien ook sprake is geweest van aselechteid in het verschijnen van bedrijven in de steekproef-uitkomsten. Het heeft dus niet alleen betrekking op de wijze van trekken van de steekproef, maar ook op de schattingsmethode voor het berekenen van uitkomsten (zie ook Bethlehem, 1985).

Een steekproef wordt (ten aanzien van een bepaalde variabele) representatief geacht voor een populatie indien de steekproefschattingen (voor deze variabele) niet systematisch verschillen van de overeenkomstige grootheden in de populatie. Een steekproef kan representatief zijn ten aanzien van een bepaalde variabele en tegelijkertijd niet-representatief ten aanzien van een andere variabele. Bij representativiteit moet niet alleen gedacht worden aan overeenkomst in gemiddelden van variabelen, maar ook aan overeenkomst in spreiding en andere verdelingskenmerken van die variabelen.

Representativiteit kan alleen worden vastgesteld voor variabelen waarvoor de waarde voor alle bedrijven in de populatie be-

kend is. Representativiteit ten aanzien van dergelijke hulpvariabelen zegt op zichzelf niets over representativiteit ten aanzien van de doelvariabelen. In het algemeen mag echter verwacht worden dat indien er een sterke samenhang is tussen de hulp- en doelvariabelen, representativiteit ten aanzien van de hulpvariabelen nuttige aanwijzingen kan geven over de mate waarin de waarden van de doelvariabelen in de steekproef systematisch afwijken van die in de populatie. Als bijvoorbeeld bedrijven met varkens gemiddeld een hoger inkomen hebben dan bedrijven zonder varkens, en bedrijven met varkens relatief weinig voorkomen in de steekproef, dan wordt het gemiddelde inkomen in de populatie wellicht iets onderschat - ervan uitgaande dat er overigens geen andere over- en onderschatting veroorzakende factoren aanwezig zijn.

Systematische afwijkingen tussen steekproef en populatie kunnen onder andere ontstaan indien bij de opzet van de steekproef geen rekening wordt gehouden met verschillen in responsgedrag binnen de populatie. Bij de steekproef voor 1975 was er bijvoorbeeld sprake van systematische afwijkingen in de leeftijdsopbouw van de bedrijven in de steekproef.

Om vast te kunnen stellen of er voor een bepaalde variabele sprake is van systematische afwijkingen, moet er iets bekend zijn over de orde van grootte van de toevallige afwijkingen. Immers, afwijkingen tussen steekproef en populatie zullen er altijd zijn bij een steekproef. Een afwijking is pas systematisch indien deze niet verklaard kan worden uit de rol van het toeval.

4.4.3 Betrouwbaarheid

Naast de eis van representativiteit, wordt aan de steekproefuitkomsten de eis van een zo groot mogelijke betrouwbaarheid gesteld. Dat wil zeggen dat de rol van het toeval in de steekproefcijfers zo klein mogelijk moet zijn. Toevallige afwijkingen tussen steekproef en populatie zijn inherent aan het trekken van een steekproef, tenzij de steekproef de gehele populatie omvat. De kans op toevallige afwijkingen van een bepaalde grootte kan kleiner gemaakt worden door meer steekproefbedrijven te kiezen, of door de steekproefmethode of de berekeningsmethode aan te passen. In de hoofdstukken 5 en 6 wordt besproken hoe betrouwbaarheid gemeten kan worden, van welke factoren betrouwbaarheid afhankelijk is, en welke maatregelen er in de LEI-steekproefopzet zijn genomen om een maximale betrouwbaarheid van de steekproefuitkomsten te garanderen.

4.4.4 Concretisering

De waarden van de doelvariabelen dienen representatief en betrouwbaar weergegeven te worden voor de bedrijven in de doelpopulatie. Daarbij worden de totaalcijfers voor de Nederlandse landbouw in een bepaald jaar als uitgangspunt gekozen. Concreet

houdt dat in dat de betrouwbaarheid van de in paragraaf 4.3 genoemde doelvariabelen bevredigend moet zijn voor deze totaalcijfers (de "nationale boerderij") onder de voorwaarde dat de uitkomsten geen systematische afwijkingen ten opzichte van de populatie vertonen.

Bij het steven naar een redelijke betrouwbaarheid van cijfers op nationaal niveau wordt in eerste instantie niet gelet op de betrouwbaarheid van totaal- of gemiddelde cijfers voor afzonderlijke groepen bedrijven. Indien blijkt dat deze niet voldoende is, wordt de steekproef - binnen de daarvoor beschikbare capaciteit - zodanig aangepast dat ook voor de afzonderlijke groepen een acceptabele betrouwbaarheid wordt gerealiseerd (zie paragraaf 7.4). In de hoofdstukken 5 en 6 zal de aandacht echter uitsluitend gericht zijn op de steekproefopzet voor de representatieve en betrouwbare schatting van totaalcijfers van de nationale boerderij.

4.5 Middelen

De mate waarin de doelstellingen van het boekhoudnet gerealiseerd kunnen worden, wordt bepaald door de aanwezige capaciteit van het LEI om bedrijven in administratie te nemen. Deze capaciteit wordt mede bepaald door de eisen die de EG ten aanzien van de Nederlandse bijdrage aan het EG-informatienet stelt. Het LEI wordt geacht jaarlijks circa 1.500 boekhoudingen van land- en tuinbouwbedrijven aan de EG te leveren. Voor het landbouw-deel - waarover dit onderzoekverslag gaat - kunnen maximaal 1.150 bedrijven in administratie worden genomen. De realisering van de doelstellingen die in de voorgaande paragrafen zijn geformuleerd, betekent dan ook dat gegeven dit aantal bedrijven, een zodanige steekproefopzet wordt gekozen dat representatieve uitkomsten met een zo groot mogelijke betrouwbaarheid verkregen worden. Daarbij worden 1.100 bedrijven gekozen vanuit de optiek van de betrouwbaarheid van de cijfers van de nationale boerderij en ruim 50 bedrijven als aanvulling hierop voor bepaalde groepen binnen de populatie.

5. Opzet van de steekproef

5.1 Inleiding

In dit hoofdstuk wordt beschreven hoe vanuit de doelstellingen die in hoofdstuk 4 zijn geformuleerd, de LEI-steekproef voor de landbouwsector is opgezet. Of beter gezegd, hoe de LEI-steekproeven zijn opgezet. Het gaat immers niet om een eenmalige steekproef, maar om een reeks van steekproeven, waaraan er ieder jaar één wordt toegevoegd. Elke steekproef geeft een beeld van wat in het betreffende jaar de situatie in de landbouw was. Samen leveren de steekproeven de ontwikkeling van de landbouwsector in de loop van de tijd. Zoals zal blijken in dit hoofdstuk is de steekproef geen panel in de strikte zin van het woord. Er is sprake van een roterend panel, waarin ieder jaar een deel van de bedrijven wordt vervangen door nieuwe bedrijven.

De statistische achtergrond van de steekproefopzet wordt in het volgende hoofdstuk behandeld. In hoofdstuk 7 wordt ingegaan op de bijstelling van de in dit hoofdstuk beschreven opzet, die voortvloeit uit het rekening houden met nevendoelstellingen.

5.2 Stratificatie

Om alle delen van de populatie goed weer te kunnen geven en om de betrouwbaarheid van de steekproefuitkomsten te vergroten, wordt gewerkt met een gestratificeerde steekproef. Voor 1975 gebeurde dat ook al - zie hoofdstuk 3 -, maar in boekjaar 1975 is een nieuw stratificatie-systeem ingevoerd. De ervaringen tot dan toe hadden namelijk geleerd dat er verbeteringen mogelijk waren. De indeling van de populatie in strata gebeurt aan de hand van gegevens die via de jaarlijkse landbouwtelling (de meitelling) beschikbaar komen. Dat zijn gegevens die voor alle bedrijven in de populatie beschikbaar zijn. De meitelling levert onder andere informatie over de omvang van een bedrijf, de samenstelling van het bouwplan en de veestapel, de aanwezige arbeidskrachten, de leeftijd van de ondernemer, de rechtspositie en de geografische ligging van een bedrijf. In afzonderlijke jaren is nog een aantal andere gegevens beschikbaar, bijvoorbeeld over de eigendoms-/pacht-verhouding of de aanwezigheid van een ligboxenstal. Deze laatste gegevens zijn echter niet geschikt voor gebruik bij de opzet van de steekproef omdat ze niet ieder jaar beschikbaar zijn.

Het doel van stratificatie is het vormen van homogene groepen bedrijven, waarbij uit iedere groep (stratum) een aantal bedrijven gekozen wordt dat dit stratum representeert. In het ver-

volg van dit onderzoekverslag zal blijken dat een dergelijke opzet zowel systematische als toevallige fouten terugdringt.

5.3 Indeling in grootte-klassen

Op grond van meerjarige ervaring werd in 1975 besloten tot een stratificatie aan de hand van vijf criteria, die de heterogene populatie zo goed mogelijk verdelen in homogene stukken. Homogeniteit heeft daarbij vooral betrekking op de doelvariabelen. Het lag voor de hand om de populatie eerst in te delen in een aantal grootteklassen, omdat een groot deel van de verschillen in bedrijfsuitkomsten en financiële positie tussen bedrijven samenhangt met de bedrijfsgrootte. Een indeling in sbe-grootte-klassen verdeelt de populatie in groepen bedrijven die ten aanzien van de doelvariabelen van elkaar verschillen. De sbe is immers een maat voor de normatieve toegevoegde waarde en daarmee ook voor de met deze variabele samenhangende andere doelvariabelen.

Oppervlakte is een veel minder geschikt indelingscriterium om de grootte van een bedrijf vast te stellen. Indien bijvoorbeeld een akkerbouwbedrijf en een pluimveebedrijf ieder 10 hectare grond bezitten, zullen ze in veel gevallen sterk verschillen ten aanzien van de hoogte van de toegevoegde waarde en van andere doelvariabelen.

5.4 Indeling in bedrijfstypen

De indeling van de populatie aan de hand van het aantal sbe per bedrijf om daarmee homogene strata te vormen, heeft een bezwaar en dat hangt samen met het gestandaardiseerde karakter van de sbe. Deze geeft namelijk aan wat er onder *generaliseerde* productie-omstandigheden aan toegevoegde waarde voortgebracht kan worden. In de praktijk is het zo dat de productie-omstandigheden - bijvoorbeeld het weer en de prijzen - voor bedrijven binnen een bepaald bedrijfstype in afzonderlijke jaren voor de meeste bedrijven sterk overeenkomen. Daardoor zal per bedrijfstype de sbe ook voor de *gerealiseerde* toegevoegde waarde in een bepaald jaar, en daarmee ook voor andere doelvariabelen in de sfeer van de bedrijfsuitkomsten en de financiën, een goede homogeniteitsmaat zijn. Maar tussen de bedrijfstypen kunnen de productie-omstandigheden in een specifiek jaar sterk verschillen. Een goed jaar voor de akkerbouw kan samenvallen met een slecht jaar voor de melkveehouderij. Daardoor kunnen bedrijven die in dezelfde sbe-klasse vallen, aanzienlijke verschillen in bedrijfsuitkomsten en financiële positie laten zien, die louter met het bedrijfstype samenhangen. Gemiddeld over alle jaren zullen de bedrijven in die sbe-klasse vergelijkbaar zijn ten aanzien van de gerealiseerde toegevoegde waarde, maar in een individueel jaar - en daar richt de steekproef zich in eerste instantie op - kunnen er behoorlijke

verschillen optreden. Daarom worden de bedrijven bij de stratificatie ook ingedeeld naar bedrijfstype.

5.5 32 basisstrata

Om zowel de heterogeniteit ten aanzien van de verschillende produktie-omstandigheden in een individueel jaar als die ten aanzien van verschillen in grootte tussen bedrijven goed op te vangen, worden de bedrijven in de populatie eerst verdeeld over acht bedrijfstypen en vervolgens worden per bedrijfstype vier sbe-klassen onderscheiden. Naast akkerbouw-, rundvee-, varkens- en pluimveebedrijven, worden vier gemengde bedrijfstypen onderscheiden. In bijlage 1 is aangegeven hoe deze bedrijfstypen worden gedefinieerd. In hoofdstuk 6 wordt ingegaan op de statistische achtergrond van de indeling in sbe-klassen. De klasse-grenzen verschuiven ieder jaar en verschillen per bedrijfstype. In bijlage 1 zijn de klasse-grenzen voor boekjaar 1987 weergegeven.

Op grond van statistische criteria wordt bepaald hoeveel bedrijven in elk van de 32 nu ontstane basisstrata gekozen moeten worden voor de steekproef (zie paragraaf 6.5). Het steekproefpercentage varieert daarbij van basisstratum tot basisstratum. De reden daarvoor is dat de grootst mogelijke betrouwbaarheid voor de steekproefschattingen van de waarden van de doelvariabelen wordt bereikt, indien het steekproefpercentage varieert met de heterogeniteit van de strata ten aanzien van de doelvariabelen. Zo ligt het steekproefpercentage voor basisstrata met akkerbouwbedrijven in het algemeen iets hoger dan bij rundveebedrijven. Ook ligt het percentage voor de hogere sbe-klassen hoger dan voor lagere sbe-klassen, omdat de sbe-trajecten van eerstgenoemde klassen groter zijn.

Achteraf wordt door middel van een wegingsprocedure (zie hoofdstuk 9) gecorrigeerd voor de verschillen in steekproefpercentage. Daarbij krijgt ieder bedrijf ten behoeve van de aggregatie en middeling van gegevens een gewicht overeenkomstig het aantal bedrijven in de populatie dat het representeert.

5.6 Verdergaande stratificatie

5.6.1 Vorming van substrata

Dat de steekproefpercentages per basisstratum worden vastgesteld, betekent niet dat het stratificatieproces stopt bij twee stratificatiecriteria en 32 basisstrata. Er wordt geprobeerd om binnen de basisstrata een indeling te maken in vijf oppervlakteklassen, drie leeftijdsklassen en dertien regio's (zie bijlage 1). In principe zou dat kunnen leiden tot 195 (sub-)strata binnen ieder basisstratum en tot een totaal van 6.240 strata voor de hele populatie. Omdat het uit statistische overwegingen gewenst is

om minimaal twee steekproefbedrijven per stratum te kiezen, en er slechts 1.150 bedrijven in de steekproef kunnen worden opgenomen, zal het aantal strata beperkt moeten blijven tot hooguit 575.

In de praktijk wordt als volgt te werk gegaan. Uitgaande van het steekproefpercentage voor een bepaald basisstratum, wordt nagegaan of er substrata gevormd kunnen worden op een zodanige wijze dat er minimaal twee bedrijven per te vormen substratum resulteren. Daarbij wordt eerst gekeken naar de mogelijkheid om het basisstratum te verdelen in oppervlakteklassen. Vervolgens komen leeftijdsklassen en regio's aan de beurt om binnen de gevormde oppervlakte-strata te zorgen voor een verdergaande indeling in substrata. In basisstrata waar veel steekproefbedrijven gekozen moeten worden - bijvoorbeeld in de vier basisstrata die betrekking hebben op het bedrijfstype "rundveebedrijven" - is een verdergaande stratificatie mogelijk. In andere basisstrata kan slechts een tweedeling plaatsvinden. Een dergelijke tweedeling hoeft niet noodzakelijkerwijs betrekking te hebben op oppervlakte-classes, ook al is dat het eerste criterium waarnaar gekeken wordt. Een voorbeeld kan dat verduidelijken.

5.6.2 Een voorbeeld van sub-stratificatie

Een indeling van de populatie zoals die op grond van de telling in mei 1982 kon worden vastgesteld, leert dat er 730 bedrijven zijn in de laagste sbe-klasse in de groep pluimveebedrijven (Van Dijk, 1984). Van deze 730 bedrijven dienden tien bedrijven in de steekproef te worden opgenomen. Een indeling in oppervlakte-classes levert 604, 104, 21, 1 en 0 bedrijven in de achtereenvolgende classes (zie bijlage 1). Gezien het steekproefpercentage van 1,37% ($=10/730 \cdot 100\%$) kunnen de classes 2,3 en 4 noch afzonderlijk noch gezamenlijk een substratum vormen waarin minstens twee steekproefbedrijven gekozen moeten worden. Bij een indeling in leeftijds-classes vallen 138, 528 en 64 bedrijven in de opeenvolgende leeftijdsclasses. Klasse 1 en 3 zouden samen een stratum kunnen vormen, maar het ligt niet voor de hand om twee niet-op-eenvolgende classes samen te voegen omdat dan een heterogeen stratum zou ontstaan. Een goede sub-stratificatie naar leeftijdsclasses is dus niet mogelijk. Een indeling van de populatie naar regio's levert meer mogelijkheden op omdat er meer classes gevormd kunnen worden en de eis van aangrenzendheid bij combinatie van classes tot een substratum niet gesteld wordt. Het blijkt dat er drie groepen van regio's binnen het basisstratum gevormd kunnen worden waaruit per groep minstens twee bedrijven gekozen moeten worden. In de regio's 1 tot en met 8 liggen 156 bedrijven, in de regio's 9 tot en met 11 230 bedrijven, en in de regio's 12 en 13 de resterende 344 bedrijven. Er waren enkele andere combinatie-mogelijkheden, maar er is gekozen voor een combinatie van geografisch aansluitende gebieden (zie Bedrijfsuitkomsten, 1988:15).

De gehele populatie wordt op deze manier per basisstratum verder gestratificeerd. Er resulteren uiteindelijk circa 275 strata, waarvan er circa 55 betrekking hebben op akkerbouwbedrijven, 120 op rundveebedrijven, en 100 op gemengde en intensieve veehouderijbedrijven.

5.7 Bewaking van de representativiteit

5.7.1 Inleiding

De indeling van de populatie in strata en de afzonderlijke trekking van steekproefbedrijven uit elk van die strata, moet voorkomen dat boekhoudnet en populatie qua opbouw systematische verschillen vertonen. Deze systematische verschillen zouden kunnen ontstaan indien de bereidheid of de mogelijkheid tot deelname aan het boekhoudnet, samenhangt met de aard van het bedrijf. In deze paragraaf zal blijken dat voor het ondervangen van de non-respons bij de keuze van bedrijven in feite geen stratificatie vooraf binnen de basisstrata hoeft plaats te vinden. In paragraaf 5.7.4 wordt echter wel een andere reden voor stratificatie vooraf gegeven, namelijk de tussentijdse uitval van bedrijven die al in het boekhoudnet zijn opgenomen.

5.7.2 Non-respons bij de keuze

"De ervaring heeft (...) geleerd dat de non-response samenhangt met variabelen als bedrijfstype, bedrijfssomvang, leeftijd van het bedrijfshoofd en regio" (Lodder, 1987:14). Mede daarom zijn deze variabelen als stratificatievariabelen gekozen. Indien een bedrijf in een bepaald stratum niet kan of wil deelnemen, wordt uit hetzelfde stratum een ander bedrijf gekozen ter vervanging van het oorspronkelijk gekozen bedrijf. Omdat in het stratificatieproces niet alle variabelen aan bod kunnen komen, wordt bij deze vervangende keuze expliciet dat bedrijf gekozen dat het meest - in termen van de vijf stratificatie-criteria - op het niet-meewerkende bedrijf lijkt. Daarbij wordt zelfs gekeken naar de gemeente waarin een bedrijf ligt. Aangezien voor ieder stratum geldt dat alle bedrijven in elk geval hetzelfde bedrijfstype hebben en dat ze allemaal in dezelfde grootte-klasse vallen, betekent dit concreet dat een bedrijf gezocht wordt dat in dezelfde oppervlakte-klasse, leeftijds-klasse en regio valt als het oorspronkelijk gekozen bedrijf. In circa 95% van de gevallen is een dergelijk "identiek" bedrijf te vinden, en in de meeste gevallen kan er voor gezorgd worden dat dat bedrijf ook in dezelfde gemeente ligt.

Deze vervanging van door het toeval aangewezen bedrijven door bedrijven die er op lijken gebeurt weliswaar binnen de strata, maar zou ook uitgevoerd kunnen worden zonder een stratifica-

tie. Slechts indien het niet mogelijk is om voor ieder niet-deelnemend bedrijf een vervanger te kiezen, bijvoorbeeld omdat het door herhaald weigeren van vervangende bedrijven niet lukt om voor het begin van de steekproefperiode - in de praktijk voor 1 mei - de gewenste bedrijven in de steekproef te krijgen, is stratificatie noodzakelijk om geen systematische afwijkingen te krijgen. Maar dan gaat het in feite om een stratificatie die achteraf opgelegd kan worden en niet om een stratificatie die vooraf moet plaatsvinden. Paragraaf 9.2 gaat in op een dergelijke stratificatie.

5.7.3 Non-respons in de uitwerkingsfase

Een andere reden voor stratificatie die eveneens samenhangt met het streven naar representativiteit is de non-respons nadat de bedrijfskeuze heeft plaatsgevonden. Ook al is in ieder stratum het kiezen van steekproefbedrijven die mee willen doen in het boekhoudnet, volledig geslaagd, dan nog is er een deel van de bedrijven dat uiteindelijk niet met een uitgewerkte boekhouding in de steekproef terechtkomt. Deze bedrijven trekken zich na het begin van het boekjaar alsnog terug of leveren onvolladige of onbetrouwbare gegevens. Zouden dit bijvoorbeeld juist de bedrijven zijn met een relatief grote bedrijfsoppervlakte of met een oudere ondernemer, dan treedt uiteindelijk toch een systematische vertekening op in de steekproefuitkomsten. Door achteraf te stratificeren en per stratum te bekijken hoeveel bedrijven er in de populatie voorkomen en hoeveel er een uitgewerkte boekhouding in het kader van het boekhoudnet hebben geleverd, kan via het gebruik van wegingsfactoren toch een goede weergave van alle delen van de populatie worden verkregen.

Vooraf stratificeren is voor deze vorm van non-respons niet noodzakelijk tenzij door de "uitwerkings-non-respons" te kleine aantallen steekproefbedrijven resulteren voor bepaalde groepen bedrijven. Ophoging van de aantallen steekproefbedrijven voor de betreffende groepen bij de opzet van de steekproef kan dan leiden tot de gewenste aantallen uitgewerkte boekhoudingen per groep.

5.7.4 Het roterende karakter van de steekproef

De totale steekproef van 1.100 bedrijven wordt niet ieder jaar geheel opnieuw getrokken. Alleen bedrijven die al zes jaar in administratie zijn geweest en bedrijven die zelf te kennen geven niet meer mee te willen werken aan het boekhoudnet, worden vervangen. Een half jaar voor het begin van een boekjaar wordt voor ieder stratum nagegaan hoeveel bedrijven er aan het begin van het volgende boekjaar waarschijnlijk in administratie zullen blijven. Als dit aantal lager is dan het gewenste steekproefaantal dat uit de steekproefopzet volgt, dan worden er bedrijven bijgekozen in dat stratum. "Waarbij dus de aanname wordt gemaakt dat de reeds in administratie zijnde bedrijven een toevallige

steekproef vormen uit alle in het stratum aanwezige bedrijven" (Lodder, 1983:18; Lodder, 1987:25). Dat betekent dat geen rekening gehouden wordt met de oorspronkelijke trekkingskansen van de aanwezige steekproefbedrijven. Voor een onderbouwing van deze procedure, zie Lodder (1983:17-23) en hoofdstuk 8. Laatstgenoemd hoofdstuk gaat ook meer gedetailleerd in op de keuze van nieuwe steekproefbedrijven ter aanvulling van de al aanwezige bedrijven in een stratum.

In feite is de enige reden waarom - uit het oogpunt van representativiteit - ook *vooraf* een stratificatie van de populatie moet plaatsvinden, gelegen in het continue karakter van de steekproef. De aanvullingen met nieuwe bedrijven zullen niet in alle strata even hard nodig zijn. Het valt te verwachten, en het is in de praktijk ook te constateren dat met name aanvullingen nodig zijn in strata waar de bedrijfs hoofden jong zijn, of waar de tussentijdse opzeggingen voor het boekhoudnet hoger liggen dan in andere strata.

5.7.5 Non-respons voor financiële gegevens

Naast de totnutoe genoemde redenen om de steekproef - al dan niet achteraf - te stratificeren uit het oogpunt van representativiteit, is er nog een reden die te maken heeft met de tweërlei soort gegevens die de steekproef oplevert. Van alle deelnemers aan het boekhoudnet wordt verwacht dat ze gegevens over de bedrijfsuitkomsten verstrekken, terwijl op het verstrekken van gegevens die betrekking hebben op de financiële positie van de ondernemer wel wordt aangedrongen, maar zonder dat dit verplicht is voor deelname aan het boekhoudnet. In de praktijk blijkt dat de groep steekproefbedrijven die beide soorten gegevens verstrekt circa 10% kleiner is dan de groep bedrijven waarvoor de bedrijfsuitkomsten aanwezig zijn. Om ook met die kleinere groep bedrijven de gehele populatie te kunnen beschrijven, en daarbij niet de kans te lopen dat systematische afwijkingen gaan ontstaan - doordat juist bedrijven met bepaalde kenmerken de financiële gegevens weigeren - wordt de kleinere groep bedrijven opnieuw ingedeeld in de al gevormde strata, waarbij eventuele lege strata worden samengevoegd. Het zal duidelijk zijn dat met name de stratificatievariabele "leeftijd" in dit verband een belangrijke rol kan spelen vanwege de samenhang tussen leeftijd en bijvoorbeeld de opbouw van het bedrijfsvermogen, de besparingen en de investeringen.

6. Statistische achtergronden

6.1 Inleiding

In het vorige hoofdstuk zijn enkele punten met betrekking tot de opzet van de steekproef aangeroerd die om een nadere statistische uitwerking vragen. Het betreft met name zaken die verband houden met de tweede reden die werd aangevoerd om een gestratificeerde steekproef in plaats van een enkelvoudige steekproef te hanteren, namelijk het realiseren van een grotere betrouwbaarheid van de steekproefuitkomsten.

Centrale onderwerpen in dit hoofdstuk zijn de (optimale) stratificatie van de populatie en de (optimale) allocatie van de steekproefbedrijven over de strata. In eenvoudiger Nederlands: hoe moeten de stratum-grenzen bepaald worden en welke steekproefpercentages moeten per stratum gekozen worden, om een zo groot mogelijke betrouwbaarheid van de steekproefuitkomsten van de totale of van de gemiddelde nationale boerderij te realiseren?

6.2 Het meten van betrouwbaarheid

Wie met een steekproef werkt, weet bij voorbaat dat er door het steekproeftoeval verschillen kunnen optreden tussen de steekproefuitkomsten en de "werkelijke" uitkomsten in de populatie. Deze toevallige afwijkingen kunnen slechts vermeden worden indien de gehele populatie in de steekproef valt. De kans op grote toevallige fouten wordt groter naarmate het aantal steekproefbedrijven kleiner is en naarmate de populatie heterogener is.

Indien een steekproef getrokken is, kan op grond van de steekproefgrootte en de variatie in de waarde van de doelvariabele, worden aangegeven hoe groot de kans is dat de werkelijke waarde van deze variabele, meer dan een bepaalde hoeveelheid van de geschatte waarde aflight. Anders gezegd, met behulp van informatie uit de steekproef kan worden bepaald hoe betrouwbaar de steekproefuitkomsten zijn.

De informatie over de betrouwbaarheid van een schatting kan worden samengevat in een getal, de zogenaamde standaardfout. Met behulp van de standaardfout kunnen intervallen rond de geschatte waarde van een variabele worden berekend, waarbinnen de werkelijke waarde van die variabele met een bepaalde kans ligt. Indien de steekproef bijvoorbeeld oplevert dat de gemiddelde arbeidsopbrengst van de ondernemer f 30.000 is, en de bijbehorende standaardfout is f 1.500, dan betekent dit dat de kans dat de werkelijke arbeidsopbrengst in het interval (f 27.000, f 33.000) ligt, ongeveer 95% is. Dat wil zeggen dat als er 100 maal een steek-

proef uit dezelfde populatie getrokken zou worden, dit in 95 gevallen zou leiden tot een gemiddelde arbeidsopbrengst die ligt tussen f 27.000 en f 33.000.

De constructie van een betrouwbaarheidsinterval gaat als volgt. Het midden van het interval is de geschatte waarde van een variabele en de uiteinden worden berekend door T respectievelijk $-T$ maal de standaardfout bij deze geschatte waarde op te tellen. De waarde van T is afhankelijk van de mate van betrouwbaarheid die gehanteerd wordt. Bij een 95%-betrouwbaarheidsinterval hoort een T -waarde van ongeveer 2, bij 99% betrouwbaarheid is T gelijk aan 3, en bij 68% betrouwbaarheid heeft T een waarde 1. Bij een $\alpha\%$ -betrouwbaarheidsinterval hoort een T -waarde van $T(0.5*\alpha+0.5)$, waarbij $T(x)$ het (x) -e percentiel van de normale verdeling is. Indien de steekproefgrootte (n) klein is (bijvoorbeeld kleiner dan 50), dient te worden uitgegaan van een t -verdeling (Student-verdeling) met $(n-1)$ graden van vrijheid.

6.3 Berekening van standaardfouten

Voor het berekenen van de standaardfout van een schatting, bijvoorbeeld van het gemiddelde of het totaal van een bepaalde variabele, is het noodzakelijk iets te weten over de grootte van de populatie waaruit de steekproef getrokken is, over het aantal steekproefbedrijven waarop de schatting gebaseerd is, en over de variantie van de betreffende variabele in de populatie. We definiëren de volgende symbolen:

y = variabele waarvoor schattingen worden gemaakt
 N = de omvang van de populatie
 y_i = de waarde van y op het i -de bedrijf ($i=1, \dots, N$)

S^2 = de variantie van y in de populatie
 \bar{Y} = het gemiddelde van y in de populatie
 Y = het totaal van y in de populatie

De variantie van y in de populatie is gedefinieerd als:

$$S^2 = \sum_{i=1}^N \frac{(y_i - \bar{Y})^2}{(N-1)} \quad (6.3.1)$$

Definiëren we verder voor een enkelvoudige aselechte steekproef ("simple random sample"):

n = de omvang van de steekproef 1)
 y_i = de waarde van y waarvoor een schatting wordt ver-
richt, op het i -de steekproefbedrijf ($i=1, \dots, n$)

1) Zonder verlies van algemeenheid wordt aangenomen dat de populatie zodanig genummerd is van 1 t/m N dat de steekproefbedrijven de nummers 1 t/m n hebben.

- \bar{y} - het gemiddelde van y op de steekproefbedrijven
- \hat{y} - het totaal van y op de steekproefbedrijven
- s^2 - de variantie van y op de steekproefbedrijven
- $\text{Var}(\bar{y})$ - de variantie van de steekproefschatting \bar{y} in de ruimte van alle steekproeven
- $\text{var}(\bar{y})$ - de steekproefschatting van $\text{Var}(\bar{y})$
- $\text{stf}(\bar{y})$ - de steekproefschatting van de standaardfout van de schatting \bar{y} (is de wortel uit $\text{var}(\bar{y})$)

Dan geldt:

$$\bar{y} = \sum_{i=1}^n y_i / n \quad (6.3.2)$$

en

$$s^2 = \sum_{i=1}^n (y_i - \bar{y})^2 / (n-1) \quad (6.3.3)$$

Te bewijzen valt (Cochran, 1977:23-24) dat de variantie van \bar{y} gelijk is aan:

$$\text{Var}(\bar{y}) = \frac{S^2}{n} * \frac{(N-n)}{N} \quad (6.3.4)$$

en

$$\text{Var}(\hat{y}) = N * \text{Var}(\bar{y}) \quad (6.3.5)$$

Omdat S^2 betrekking heeft op de totale populatie en de waarden van y_i in de populatie juist onbekend zijn - daarom wordt er immers een schatting met behulp van een steekproef verricht - kan S^2 niet bepaald worden. S^2 kan echter wel geschat worden op basis van gegevens van de steekproef. Een statistisch zuivere schatter voor S^2 is namelijk s^2 (Cochran, 1977:25-27). De variantie van y kan dan berekend worden als:

$$\text{var}(\bar{y}) = \frac{s^2}{n} * \frac{(N-n)}{N} \quad (6.3.6)$$

en analoog voor $\text{var}(\hat{y})$.

Voor de afleiding van standaardfouten (varianties) van percentages of aantallen bedrijven en voor andere kengetallen die uit de steekproef zijn te berekenen - bijvoorbeeld quotiënten - zie bijlage 9 en paragraaf 12.7.

6.4 Standaardfouten bij een gestratificeerde steekproef

De LEI-steekproef is een gestratificeerde steekproef waarbij de populatie eerst in H strata wordt verdeeld, waarna binnen ieder stratum een enkelvoudige aselechte steekproef wordt getrokken. Standaardfouten die betrekking hebben op de schatting voor een bepaald stratum, kunnen met behulp van de formules in de vorige paragraaf worden berekend. Vaak gaat de interesse uit naar standaardfouten van schattingen die betrekking hebben op de gehele populatie of op een aantal strata gezamenlijk. De berekening van standaardfouten houdt in dergelijke gevallen in dat een gewogen sommatie van de varianties van afzonderlijke strata moet worden uitgevoerd. Om de formule voor de betreffende standaardfout te kunnen weergeven, worden de volgende symbolen gedefinieerd, waarbij de index h slaat op het stratum-nummer en de index i op het nummer van het bedrijf binnen een stratum.

H	= het aantal strata
N_h	= het totaal aantal bedrijven in de populatie (in stratum h ; $h=1, \dots, H$)
n_h	= het totaal aantal bedrijven in de steekproef (in stratum h)
y_{hi}	= de waarde van y (in stratum h op bedrijf i) ($i=1, \dots, n_h$)
f_h	= de steekproef-fractie (in stratum h)
\bar{Y}_h	= het populatie-gemiddelde van y (in stratum h)
y_h	= het steekproefgemiddelde van y (in stratum h)
S^2_h	= de variantie van y in de populatie (in stratum h)
s^2_h	= de variantie van y in de steekproef (in stratum h)
\bar{y}_{st}	= het gemiddelde van y in de hele steekproef
\bar{y}_{st}	= het totaal van y in de hele steekproef
W_h	= de wegings- of ophoogfactor voor bedrijven in stratum h

Het gemiddelde van y in de populatie (\bar{Y}) kan zuiver geschat worden als:

$$\bar{y}_{st} = \frac{1}{N} * \sum_{h=1}^H N_h \bar{y}_h \quad (6.4.1)$$

waarbij

$$\bar{y}_h = \sum_{i=1}^{n_h} y_{hi} / n_h \quad (6.4.2)$$

De schattingen \bar{y}_h uit de individuele strata krijgen hun juiste gewicht in de berekening voor alle strata gezamenlijk door

de vermenigvuldigings-factoren (N_h/N) in de formule voor \bar{y}_{st} . De schatting y_{st} kan herschreven worden als:

$$\bar{y}_{st} = \frac{1}{N} * \left[\sum_{h=1}^H \sum_{i=1}^{n_h} (N_h/n_h) * y_{hi} \right] \quad (6.4.3)$$

De factoren (N_h/n_h) worden de wegings- of ophoogfactoren (W_h) van de individuele waarnemingen y_{hi} genoemd. Alle bedrijven binnen een bepaald stratum hebben dezelfde wegingsfactor W_h .

Een stelling uit de steekproeftheorie zegt dat als steekproeven in verschillende strata onafhankelijk van elkaar getrokken worden, geldt dat

$$\text{Var}(\bar{y}_{st}) = \sum_{h=1}^H W_h^2 * \text{Var}(\bar{y}_h) \quad (6.4.4)$$

Op grond hiervan kan afgeleid worden dat de variantie van \bar{y}_{st} gelijk is aan:

$$\text{Var}(\bar{y}_{st}) = \left(\frac{1}{N^2} \right) * \sum_{h=1}^H N_h * (N_h - n_h) * \left(\frac{S_h^2}{n_h} \right) \quad (6.4.5)$$

waarbij in de praktijk weer gewerkt wordt met $\text{var}(\bar{y}_{st})$ (s_h^2 in plaats van S_h^2). De variantie van de schatting \hat{y}_{st} van het populatie-totaal Y is:

$$\text{Var}(\hat{y}_{st}) = N^2 * \text{Var}(\bar{y}_{st}) \quad (6.4.6)$$

De relatieve standaardfout (de standaardfout in procenten van de schatting) is dus voor de schatting van een totaal gelijk aan die van het bijbehorende gemiddelde.

Voor het afleiden van betrouwbaarheidsintervallen geldt in principe dezelfde procedure als bij de schatter y voor een enkelvoudige steekproef, althans indien er per stratum voldoende graden van vrijheid zijn. De te volgen werkwijze indien dat niet het geval is, is te vinden in Cochran (1977:96). Voor het berekenen van standaardfouten voor deelpopulaties van strata, zie bijlage 9. Daar worden ook de standaardfouten gegeven voor geschatte aantallen bedrijven en percentages.

Het is duidelijk dat bij de berekening van de standaardfouten niet voorbijgegaan kan worden aan de wegingsfactoren. Immers, bij de berekening van schattingen worden ze ook meegenomen. In de praktijk gebeurt het meenemen van de wegingsfactoren ook wel op

een andere manier dan via vergelijking 6.4.6. Berekening van standaardfouten vindt dan plaats met een aangepaste versie van vergelijking 6.3.6.

$$\text{var}(\bar{y}_{st}) = \frac{s_w^2}{n} * \frac{(N-n)}{N} \quad (6.4.7)$$

waarbij:

$$s_w^2 = \frac{\sum_{h=1}^H \sum_{i=1}^{n_h} W_h * (x_i - \bar{x})^2}{(n-1) * (N/n)} \quad (6.4.8)$$

In de hoofdstukken 12 en 13 zal nagegaan worden in hoeverre deze berekening een goede benadering geeft voor de berekening via vergelijking 6.4.6.

6.5 Verdeling van de steekproefbedrijven over de strata

6.5.1 Optimale allocatie

Ervan uitgaande dat de populatie is ingedeeld in strata - hoe die indeling plaats vindt komt in paragraaf 6.6 aan de orde - kan het totaal van n steekproefbedrijven op verschillende manieren over de strata worden verdeeld. Deze verdeling heeft rechtstreeks invloed op de standaardfout van de schattingen, omdat n_h in de formule voor de standaardfout voorkomt. Om de standaardfout van de schatting y_{st} of van de schatting Y_{st} zo klein mogelijk te maken, moeten de bedrijven op een zodanige wijze over de strata worden verdeeld (gealloceerd) dat

$$n_h = n * \frac{N_h * S_h}{\sum_{h=1}^H N_h * S_h} \quad (6.4.9)$$

Deze allocatie die optimale of Neyman-allocatie wordt genoemd, houdt dus in dat de steekproef-fractie f_h ($=n_h/N_h$) recht-evenredig met S_h wordt gekozen. De minimale variantie die gerealiseerd kan worden is bij deze keuze van n_h gelijk aan (Cochran, 1977:99):

$$\text{Var}_{\min}(\bar{y}_{st}) = \frac{\sum_{h=1}^H \left(\frac{N_h}{N} * S_h \right)^2}{n} - \frac{\sum_{h=1}^H N_h * S_h^2}{N^2} \quad (6.4.10)$$

Bij een indeling in strata waarbij de steekproefbedrijven binnen een stratum een homogene groep vormen en er sprake is van heterogeniteit tussen de strata, zal in vrijwel alle gevallen $\text{Var}(\bar{y}_{st})$ kleiner zijn dan $\text{Var}(\bar{y})$. Ook indien in alle strata hetzelfde steekproefpercentage wordt toegepast (evenredige allocatie), zal $\text{var}(\bar{y}_{st})$ vaak kleiner zijn dan $\text{var}(\bar{y})$ (Cochran, 1977:99-101).

Stratificatie is dus niet alleen een middel om te zorgen voor een goede representativiteit, maar ook een middel om te komen tot een grotere betrouwbaarheid van de steekproefuitkomsten.

6.5.2 Sbe per bedrijf als hulpvariabele

De allocatie van steekproefbedrijven over de strata, vindt plaats voordat een steekproef getrokken worden. Op dat moment is vaak weinig bekend over S^2_h , omdat deze variantie betrekking heeft op de variabele die waargenomen moet worden in de steekproef. In de praktijk van het steekproefonderzoek moet vaak gekozen worden uit de volgende twee mogelijkheden. In de eerste plaats kan S^2_h bepaald worden uit eerdere steekproeven, waarin dezelfde variabele werd waargenomen voor dezelfde strata. Een andere mogelijkheid is het benaderen van S^2_h door de variantie van een andere variabele waarvan de waarde wel bekend is voor alle bedrijven in de populatie en die nauw samenhangt met de onderzoeksvariabele. Een nadeel hiervan is dat de allocatie minder optimaal uit kan vallen. In veel praktijksituaties blijkt echter dat kleine afwijkingen ten opzichte van optimale allocatie, weinig invloed hebben op de standaardfout. De strata-vorming heeft in het algemeen een grotere invloed op de standaardfout dan de allocatie.

Voor de LEI-steekproef wordt de S^2_h van de doelvariabelen benaderd met behulp van de hulpvariabele (sbe in metelling)³. Deze variabele is gekozen vanwege de samenhang met een groot deel van de doelvariabelen van het boekhoudnet. In principe levert optimale allocatie voor iedere doelvariabele een andere gewenste allocatie op. Aangezien er slechts één allocatie kan resulteren, moet een tussenweg tussen alle gewenste allocaties worden bewandeld.

Een eerste stap daartoe is het centraal stellen van twee doelvariabelen die geacht worden goede representanten te zijn van het hele scala van doelvariabelen waarop het boekhoudnet zich richt. Gekozen is voor de variabelen "normatieve toegevoegde waarde" en "arbeidsopbrengst van de ondernemer" (zie paragraaf 4.3). De eerste is per definitie gelijk aan het aantal sbe op een bedrijf - op een vermenigvuldigingsconstante na - en de tweede variabele neemt ongeveer rechtevenredig toe met de wortel uit het aantal sbe per bedrijf. Daarom wordt als hulpvariabele bij de allocatie het meetkundig gemiddelde van "sbe" en " $(sbe)^{1/2}$ " gehanteerd. De spreiding S^2_h in $(sbe)^{1/2}$ wordt zo gebruikt als een maat-

staf voor de spreiding in de nog onbekende waarden van de doelvariabelen (Lodder, 1983:12; Lodder, 1987:19). In paragraaf 12.9 wordt nagegaan in hoeverre deze hulpvariabele inderdaad een geschikte hulpvariabele is bij de allocatie.

6.6 Vorming van strata

6.6.1 Optimale stratificatie

In paragraaf 5.6 is ingegaan op de indeling van de populatie in circa 275 strata. Non-respons, tussentijdse uitval van deelnemende bedrijven en structuur-wijzigingen in de populatie nopen ertoe om te stratificeren, om daardoor systematische afwijkingen tussen steekproef en populatie tegen te gaan. Daarbij wordt getracht zodanig te stratificeren dat groepen bedrijven die ten aanzien van de doelvariabelen in waarde verschillen en die - om welke reden dan ook - verschillende kansen hebben om in het boekhoudnet terecht te komen, in afzonderlijke strata terechtkomen. Op grond van ervaring kon in 1975 worden aangegeven dat de in paragraaf 5.6 weergegeven indeling in strata de kans op systematische afwijkingen behoorlijk zou verkleinen.

De indeling in 275 strata en dan met name de indeling in 32 basisstrata, heeft nog een andere achtergrond. Stratificatie kan namelijk ook leiden tot een reductie in de standaardfouten van de doelvariabelen. Deze reductie wordt niet alleen bewerkstelligd door een goede allocatie - zie paragraaf 6.5 - maar ook en vooral door het vormen van strata die in termen van de doelvariabelen zoveel mogelijk van elkaar verschillen. De stratificatie die de tussenvariantie maximaliseert levert een minimalisatie van de standaardfout. Deze stratificatie wordt een optimale stratificatie genoemd.

Een optimale stratificatie wordt bereikt door de totale populatie te ordenen naar de waarde van de doelvariabele en daarbij de stratumgrenzen zo te kiezen dat het produkt van het populatie-aantal (N_h) en de standaarddeviatie van de doelvariabele (S_h) in alle strata gelijk is.

Omdat de populatie voor de steekproef niet geordend kan worden naar de waarde van de doelvariabele, wordt ook hier gewerkt met de hulpvariabele (sbe in metelling)^{3/4}. Per bedrijfstype worden alle bedrijven geordend naar $(sbe)^{3/4}$ en de stratumgrenzen worden zodanig bepaald dat in alle sbe -strata binnen een type $N_h * S_h$ konstant is, waarbij S_h^2 de variantie van de variabele $(sbe)^{3/4}$ is. Op de vraag hoeveel strata onderscheiden dienen te worden, wordt in paragraaf 6.7 ingegaan.

6.6.2 Een voorbeeld

Een voorbeeld kan de procedure van optimale stratificatie verduidelijken. Voor het boekjaar 1984 werd bij de steekproefopzet uitgegaan van de populatie in de metelling 1982. In de betreffende metelling werden 11.677 bedrijven als akkerbouwbedrijf getypeerd. Het ging daarbij om bedrijven die groter waren dan 76 sbe en kleiner dan 2.000 sbe. Uitgaande van een indeling in vier basisstrata van deze groep bedrijven, leverde het principe van optimale stratificatie de volgende akkerbouw-basisstrata op:

- 1) 76- 153 sbe
- 2) 153- 245 sbe
- 3) 245- 403 sbe
- 4) 403-2000 sbe.

Deze basisstrata bevatten respectievelijk 4.321, 3.967, 2.630 en 759 bedrijven, en voor ieder stratum gold dat $N_h \cdot S_h$ gelijk was aan dit produkt in elk ander stratum (Van Dijk, 1984).

Voor de overige zeven bedrijfstypen werden eveneens dergelijke basisstrata gevormd, waarbij de stratumgrenzen van type tot type verschilden.

6.6.3 Optimale allocatie binnen optimale strata

Het steekproefaantal van 1.100 bedrijven wordt via optimale allocatie over de 32 gevormde strata verdeeld aan de hand van de berekende S_h . Aangezien optimale allocatie betekent dat (n_h/N_h) evenredig aan S_h gekozen moet worden, moet n_h dus evenredig met $N_h \cdot S_h$ gekozen worden. Dat wil zeggen dat binnen een bepaald bedrijfstype in elk van de vier basisstrata een gelijk aantal steekproefbedrijven gekozen moet worden; $N_h \cdot S_h$ is immers constant binnen dat bedrijfstype.

In het voorbeeld betekent dit voor boekjaar 1984 dat in ieder akkerbouw-basisstratum 58 bedrijven gekozen moesten worden 1). De aantallen van 58 per basisstratum komen overeen met steekproefpercentages van respectievelijk 1,34%, 1,46%, 2,21% en 7,64% in de vier sbe-klassen binnen de akkerbouw. Dit verschil in percentages vindt zijn oorzaak in de verschillende breedtes van de sbe-klassen, die voortvloeit uit de niet-willekeurige verdeling van de 11.677 bedrijven over het gehele sbe-traject van 79-2.000 sbe. Er is een duidelijke concentratie aan het begin van dat traject.

De hoge percentages voor de grotere bedrijven hebben in directe zin niets te maken met het feit dat een groep grotere be-

1) Voor de overige zeven bedrijfstypen resulteerden aantallen van respectievelijk 140, 25, 10, 7, 6, 13 en 16 per basisstratum.

drijven met een zelfde bedrijfsomvang - bijvoorbeeld alle bedrijven met een omvang van 650-660 sbe - heterogener van aard zou zijn dan een groep kleinere bedrijven met eveneens een gelijke omvang - bijvoorbeeld alle bedrijven met een omvang van 85-95 sbe. Zouden de meeste bedrijven geconcentreerd zijn rond 500 sbe, met slechts weinig bedrijven in de buurt van 79 sbe, dan zouden de percentages er heel anders uitzien.

6.6.4 Allocatie in de praktijk

De procedure die in de LEI-praktijk wordt toegepast ten aanzien van de optimale allocatie over de optimale strata, wijkt enigszins af van wat hiervoor is uiteengezet. In eerste instantie wordt uitgegaan van een steekproefaantal van 1.200 bedrijven die optimaal over de 32 basisstrata worden verdeeld. Vervolgens wordt in elk van de vier basisstrata van het bedrijfstype "rundveehouderij" het aantal te kiezen bedrijven met 25 verminderd, zodat een uiteindelijk steekproefaantal van 1.100 bedrijven resulteert. De gedachte hierachter is dat (sbe)² weliswaar een goede indicator is voor de spreiding in de doelvariabelen binnen een bepaald bedrijfstype, maar dat een bepaalde hoeveelheid spreiding in (sbe)² in het ene bedrijfstype een grotere hoeveelheid spreiding in de doelvariabelen inhoudt dan in het andere bedrijfstype. Bij het type "rundveehouderij" is de verhouding tussen de spreiding in de waarde van de doelvariabelen en die in de waarde van (sbe)² in het algemeen kleiner dan in de overige bedrijfstypen. De "1.200 minus 4*25"-allocatie stuurt daarom de allocatie die optimaal is voor de hulpvariable (sbe)² zodanig bij dat deze optimaal wordt voor de doelvariabelen.

6.7 Van 32 basisstrata naar 275 strata

Bij de indeling van de bedrijfstypen in sbe-klassen, is uitgegaan van een indeling in vier klassen. Theoretisch kan worden aangetoond dat een groter aantal klassen nauwelijks zinvol is uit het oogpunt van reductie van standaardfouten van doelvariabelen, met name indien de indeling in klassen gebaseerd is op de variatie van een hulpvariabele, en de correlatie tussen hulp- en doelvariabele kleiner is dan 0,90 (Cochran, 1977:132-134). Indien daartoe de mogelijkheid aanwezig is, verdient het de voorkeur "om meer indelingscriteria met een beperkt aantal klassen te hantieren, liever dan reeds opgenomen criteria in meer klassen te splitsen" (Lodder, 1987:20).

In hoofdstuk 5 kwam al aan de orde dat binnen de 32 basisstrata een verdere stratificatie wordt uitgevoerd met behulp van de variabelen oppervlakte (oppervlakte cultuurgrond gemeten maat), leeftijd (van de ondernemer) en regio(nale ligging van het bedrijf). Voornaamste reden daarvoor is het representatief houden van de steekproef. Systematische afwijkingen tussen populatie en

steekproef kunnen door een stratificatie worden tegengegaan. Voor een boekjaar begint wordt bekeken hoe de meetingspopulatie eruit ziet en hoeveel steekproefbedrijven er in de verschillende delen (strata) van de populatie aanwezig zijn. Daardoor kan de keuze van nieuwe bedrijven gaten in de aansluiting van steekproef en populatie opvullen (zie hoofdstuk 8).

In principe zou de verdergaande stratificatie ook tot een verdere reductie van de standaardfout kunnen leiden. De betekenis daarvan is echter niet groot. Vanwege de veel geringere samenhang tussen oppervlakte, leeftijd en regio enerzijds en de doelvariabelen anderzijds, ligt nauwelijks nog enige betrouwbaarheidsvergroting in het verschiet. Voor alle strata binnen een basisstratum wordt daarom hetzelfde steekproef-percentages toegepast. Er wordt niet opnieuw een optimale stratificatie en allocatie naar bijvoorbeeld oppervlakte uitgevoerd.

Over de volgorde waarin de stratificatievariabelen een rol spelen bij het vormen van strata, merkt Lodder (1983:16) op: "Wel is duidelijk dat de gekozen volgorde bij de strata-vorming (eerst hectare, dan leeftijd en tenslotte regio) invloed heeft op de uitkomsten van dit proces. De gekozen volgorde is gebaseerd op twee uitgangspunten. Ten eerste zal de variantie van de schatters door de stratificatie het meest gereduceerd worden door de inbreng van de oppervlakte; de leeftijd zal veelal minder invloed hebben en de regio vaak nog minder. Ten tweede volgt daaruit onmiddellijk dat in geval deze stratificatie niet toegepast wordt de invloed van non-response, voor zover deze gekoppeld is aan bedrijven met een andere structuur in respectievelijk hectare, leeftijd en regio, het meest vertekende beeld geeft als niet minstens op hectare en zo dat kan ook nog op leeftijd en regio gelet wordt".

De stratificatie-variabelen zijn dus variabelen die zowel met de doelvariabelen als met de non-respons samenhangen. Naar een variabele die van grote invloed is op de non-respons, bijvoorbeeld het al dan niet lid zijn van een politieke partij, hoeft niet gestratificeerd te worden - zo dat al mogelijk zou zijn - indien deze variabele totaal geen samenhang vertoont met de waarden van de doelvariabelen. Ook hoeft niet gestratificeerd te worden naar variabelen die wel enige samenhang met de doelvariabelen vertonen, maar niet zoveel dat een reductie in de standaardfouten enige zoden aan de dijk zet, indien deze variabelen niet samenhangen met de respons.

6.8 De indeling in strata nader bekeken

In het voorafgaande is gebleken dat het stratificeren gericht is op het vormen van homogene strata om zo de representativiteit te waarborgen en de betrouwbaarheid te vergroten. De aard van de stratificatie-variabelen verdient nog enige aandacht.

Er worden acht bedrijfstypen onderscheiden, waarvan vier "zuiver" kunnen worden genoemd en vier "gemengd". De zuivere typen zijn "akkerbouwbedrijf", "rundveebedrijf", "varkensbedrijf" en "pluimveebedrijf". Een bedrijf heet bijvoorbeeld varkensbedrijf indien meer dan twee derde deel van het totale aantal sbe op het bedrijf betrekking heeft op de produktierichting varkenshouderij. Deze grens wordt voor alle genoemde zuivere typen gehanteerd (Lodder, 1983:13). De gemengde typen zijn "gemengd akkerbouwbedrijf", "gemengd rundvee/akkerbouwbedrijf", "gemengd rundvee/intensieve veehouderij-bedrijf" en "gemengd intensief-bedrijf". Een bedrijf krijgt een van deze typeringen indien het geen zuiver bedrijf is, maar er wel accenten zijn op bepaalde produktierichtingen. Voor de bepaling van het type van een niet-zuiver bedrijf wordt het totale aantal sbe gesplitst in drie groepen, te weten "sbe akkerbouw en tuinbouw", "sbe weidedieren en grasland" en "sbe intensieve veehouderij" (zie bijlage 1).

Het is niet op voorhand duidelijk dat deze indeling in acht bedrijfstypen de meest ideale is met het oog op representativiteit en betrouwbaarheid. Lodder (1983:13) onderbouwt de gekozen indeling door te stellen dat "uit empirisch onderzoek is gebleken dat normaliter de totale activiteit op een landbouwbedrijf bestaat uit activiteiten in niet meer dan drie sectoren 1) en dat per sector gewoonlijk minstens 1/3 deel van de totale bedrijfsactiviteit gebundeld is". Daardoor "lijkt de homogeniteit van de reële bedrijfsuitkomsten ten opzichte van de norm het evenwichtigst te worden weergegeven door de bedrijven in te delen naar zuivere bedrijven - meer dan 2/3 deel van de totale bedrijfsactiviteiten in de betreffende sector - en gemengde bedrijven - elk der sectoren minder dan 2/3. Theoretisch leidt dit tot een indeling in vier zuivere typen en 24 gemengde typen (...) doch in de praktijk blijkt dat slechts een viertal gemengde typen reëel zijn".

In hoofdstuk 11 wordt aan de hand van gegevens uit recente boekjaren nagegaan hoe evenwichtig de gekozen typering is met betrekking tot de representativiteit van bedrijven. In hoofdstuk 12 wordt teruggekomen op de invloed van de typering op de betrouwbaarheid van de uitkomsten.

1) Er worden vier "sectoren" onderscheiden, namelijk akkerbouw, rundveehouderij, varkenshouderij en pluimveehouderij.

7. Bijstelling van de steekproef op grond van nevendoelestellingen

7.1 Inleiding

Naast de hoofddoelestelling om op nationaal niveau representatieve en betrouwbare steekproefuitkomsten te produceren, is er de doelestelling om ook voor bepaalde groepen bedrijven binnen de populatie betrouwbare gegevens te kunnen leveren. Daaraan is in de hoofdstukken 5 en 6 in eerste instantie voorbij gegaan.

Om aan de nevendoelestellingen te kunnen voldoen is het noodzakelijk om voor sommige groepen bedrijven extra steekproefbedrijven te kiezen, naast de bedrijven die gekozen worden op grond van de optimale allocatie die hoort bij de hoofddoelestelling. Voor dit doel worden jaarlijks ruim 50 bedrijven aan de steekproef van 1.100 bedrijven toegevoegd.

7.2 BUL-groepen

Ongeveer een jaar na afloop van een boekjaar publiceert het LEI de "Bedrijfsuitkomsten in de landbouw" (BUL) voor dat boekjaar. "De BUL beoogt onder andere op basis van gemiddelde cijfers van groepen bedrijven een naar bedrijfstype, gebied en bedrijfsomvang gedifferentieerd beeld te geven van de ontwikkeling in bedrijfsvoering en bedrijfsuitkomsten" (Bedrijfsuitkomsten, 1988:10). De indeling naar bedrijfstype en bedrijfsomvang is sinds 1968 nauwelijks aan verandering onderhevig geweest. In hoofdstuk 3 is deze indeling beschreven. In tabel 1.2 van de BUL wordt aangegeven welke afzonderlijke regionale groepen in de BUL onderscheiden worden en welke aantallen bedrijven volgens de meitellingsgegevens in deze groepen thuishoren. Vergelijking van deze tabel met een soortgelijke tabel van bijvoorbeeld tien jaar eerder, laat zien dat de meeste groepen maar weinig in omvang zijn veranderd. Alleen de groep grotere gemengde bedrijven met overwegend akkerbouw is in aantal gehalveerd, terwijl ook de groepen kleinere (gemengde) weidebedrijven fors in aantal zijn teruggelopen. Deze structurele ontwikkelingen hebben er achter niet toe geleid dat over genoemde groepen niet meer afzonderlijk gepubliceerd kon worden.

Naast deze indeling in twintig groepen bedrijven, wordt een groep "gespecialiseerde melkveebedrijven" en een groep "bedrijven met gespecialiseerde varkenshouderij" onderscheiden, waarbij de regionale ligging van de bedrijven buiten beschouwing wordt gelaten. Ook wordt in de BUL per bedrijfstype een indeling in grootte-classes gemaakt, waarbij afhankelijk van het beschikbare aantal bedrijven per bedrijfstype een aantal classes wordt onderscheiden.

7.3 FIP-groepen

In de "Financiële positie van de landbouw" (FIP) geeft het LEI een jaarlijks totaalbeeld van de Nederlandse landbouwbedrijven met een bepaalde minimale omvang. "Het totaalbeeld omvat de kapitaals- en vermogenspositie (...) (balans) en de inkomensvorming, de inkomensbesteding en de financiering van de bedrijven" (Aukema en Overgaw, 1988:9). Naast tabellen voor de gehele Nederlandse landbouw ("de nationale boerderij") worden ook tabellen gepresenteerd voor groepen bedrijven. Naast de BUL-groepen worden vanuit verschillende gezichtspunten groepsindelingen gemaakt. De belangrijkste indelingen zijn die naar eigendoms- en pachtbedrijven, naar bedrijfsomvang, naar duur van bedrijfsuitoefening en naar omvang van bruto-investeringen in duurzame produktiemiddelen (Aukema en Overgaw, 1988:10).

7.4 Regionale aanvullingen

7.4.1 Vijftig bedrijven

Indien alle aandacht bij de steekproefopzet gericht zou worden op de hoofddoelstelling, zou er eenvoudig een optimale allocatie kunnen plaatsvinden van 1.150 bedrijven. Dat levert de grootst mogelijke betrouwbaarheid voor totalen en gemiddelden van variabelen op nationaal niveau. De toewijzing van 50 bedrijven aan strata los van deze allocatie, houdt in dat - terwille van de nevendoeelstellingen - op het niveau van totaalcijfers van de Nederlandse landbouw genoeg genomen wordt met standaardfouten die iets hoger zijn dan bij optimale allocatie van alle 1.150 bedrijven. Deze vergroting is echter erg klein. De standaardfout van een schatting neemt immers omgekeerd evenredig toe met de wortel uit het aantal steekproefbedrijven. Een vermindering van 1.150 naar 1.100 bedrijven vergroot de standaardfout met een factor 1,02 ($\sqrt{1.150/1.100}$). Bovendien dragen de 50 bedrijven wel degelijk bij aan de reductie van de standaardfout, zij het dan niet optimaal.

Op de standaardfouten van schattingen voor groepen bedrijven waaraan deze 50 bedrijven worden toegevoegd, heeft deze extra allocatie meestal een veel grotere invloed. Een toename van het aantal steekproefbedrijven in een groep van bijvoorbeeld 25 naar 36, vermindert de standaardfout van schattingen voor die groep met bijna 17% ($1 - \sqrt{25/36}$).

7.4.2 Aanvulling voor akkerbouwgroepen

De aanvullende bedrijven die worden gekozen hebben betrekking op groepen die in paragraaf 7.2 zijn omschreven. Zonder deze aanvullingen zou een aantal groepen gegevens leveren die niet voldoende betrouwbaar worden geacht. De laatste jaren zijn met

name voor drie groepen akkerbouwbedrijven extra bedrijven in de steekproef opgenomen. Het betrof de groepen "grotere bedrijven in het Noordelijk Kleigebied", "grotere bedrijven in de Veenkoloniën en het Noordelijk Zandgebied" en "kleinere bedrijven in de Veenkoloniën en het Noordelijk Zandgebied". De oorzaak hiervan moet met name gezocht worden in de kleine populatie-aantallen in deze groepen (in het boekjaar 1986 respectievelijk 1.119, 1.240 en 1.009) en niet alleen in een grotere heterogeniteit in deze groepen in termen van de doelvariabelen. In de Veenkoloniën en het Noordelijk Zandgebied is juist sprake van vrij homogene groepen.

De keuze voor het aanvullen van het aantal steekproefbedrijven in een aantal strata, is gebaseerd op een afweging die wordt gemaakt tussen het belang van meer betrouwbaarheid op nationaal niveau en het belang van betrouwbare cijfers over afzonderlijke groepen. Dat de aanvullende keuze zich met name op akkerbouwbedrijven richt, en niet op bijvoorbeeld mestvarkensbedrijven, is de vertaling van de doelstelling om voor vier afzonderlijke gebieden in Nederland cijfers te publiceren over de groep grote akkerbouwbedrijven en in het geheel geen afzonderlijke cijfers te presenteren over mestvarkensbedrijven.

Bij de opzet van de steekproef voor boekjaar 1987 werd bijvoorbeeld van tevoren vastgesteld dat er voor de vier regionale groepen grotere akkerbouwbedrijven respectievelijk 55, 40, 45 en 55 bedrijven per gebied gekozen moesten worden per 1 mei 1987, om na afloop van het boekjaar voldoende bedrijven per gebied over te houden om redelijk betrouwbare cijfers per gebied te kunnen publiceren. Deze aantallen zijn gebaseerd op ervaring. Impliciet zitten achter deze aantallen beschouwingen ten aanzien van een redelijke betrouwbaarheid, dat wil zeggen ten aanzien van de toegestane standaardfout van de schattingen.

7.4.3 Afstemming bedrijfskeuze op BUL-groepen

Een probleem bij de keuze van aanvullende bedrijven, zijn de verschillen in definities bij het vormen van enerzijds strata en anderzijds BUL- en FIP-groepen (zie ook paragraaf 13.2 en 13.3). In de BUL wordt een akkerbouwbedrijf gedefinieerd als een bedrijf waar meer dan 80% van het aantal sbe betrekking heeft op akkerbouw-activiteiten. Bij de strata-indeling wordt het percentage van 66 2/3% gehanteerd, terwijl bovendien iets anders wordt verstaan onder een akkerbouw-activiteit. Ook de definitie van strata ten aanzien van de regio's kan problemen geven. In een stratum zitten vaak bedrijven uit verschillende regio's, doordat "regio" pas in laatste instantie een rol speelt bij de vorming van strata. Indien ten behoeve van bijvoorbeeld de BUL-groep "grotere akkerbouwbedrijven in regio 1" een aanvullend bedrijf wordt getrokken in een akkerbouwstratum met bedrijven uit de regio's 1 tot en met 3 dan kunnen in principe - vanwege de aseletheid van het trekken binnen een stratum - bedrijven gekozen worden die minder dan 80% akkerbouw hebben en/of - wat vaker voorkomt - in regio 2

of 3 liggen. Deze bedrijven zouden dan uiteindelijk toch niet in de betreffende BUL-groep terechtkomen, en de aanvulling zou niet het gewenste effect hebben. Daarom worden de extra bedrijven voor een BUL-groep voornamelijk gekozen in die strata waar vrijwel uitsluitend bedrijven uit de betreffende regio voorkomen. Kleinere akkerbouwbedrijven worden gekozen uit de strata die gevormd zijn binnen het basisstratum dat betrekking heeft op de laagste sbe-klasse en grotere akkerbouwbedrijven uit de drie hoogste sbe-klassen. Daarbij is het een gelukkige omstandigheid dat de sbe-grens tussen de eerste en de tweede sbe-klasse niet veel afwijkt van de sbe-grens tussen grotere en kleinere bedrijven in de BUL.

Overigens zal er ook bij gelijke indelingscriteria voor de stratificatie en de BUL nooit een complete aansluiting van beide gerealiseerd kunnen worden, omdat bij de stratificatie de bedrijven ingedeeld worden op basis van de metelling en in de BUL op basis van de situatie op het bedrijf zoals die gemiddeld gedurende het boekjaar is.

De keuze van aanvullende bedrijven gebeurt zodanig dat daardoor de gewichten van de bedrijven in de strata binnen een bepaald basisstratum - en dan met name die strata die betrekking hebben op dezelfde regio's - meer naar elkaar toegebracht worden. Strata met bedrijven die lage gewichten hebben vanwege een (te) groot aantal bedrijven die in dat stratum al in administratie zijn - op de oorzaak daarvan wordt in het volgende hoofdstuk ingegaan - komen dus niet zo snel in aanmerking voor een aanvullende keuze.

Het werkelijk aantal aanvullende bedrijven dat de laatste zeven boekjaren tot en met 1987 is bijgekozen, bedraagt gemiddeld 61, dat wil zeggen iets meer dan 5% van alle steekproefbedrijven.

8. De keuze van de bedrijven

8.1 Inleiding

In hoofdstuk 5 werd aangegeven dat na de indeling van de populatie in 275 strata, en de vaststelling van de gewenste aantallen per stratum, allereerst alle bedrijven waarvan verwacht mag worden dat ze aan het begin van het komende boekjaar nog in administratie zullen zijn, ingepast worden in het nieuwe stratificatie-systeem. Dat betreft alle bedrijven die minder dan 6 jaar in administratie zijn en die niet tussentijds hun administratie stoppen, omdat bijvoorbeeld het bedrijf opgeheven wordt of omdat verdere medewerking geweigerd wordt.

In dit hoofdstuk wordt daar verder op ingegaan en wordt besproken hoe nieuwe bedrijven in de steekproef terecht komen.

8.2 Blijvende bedrijven

Voor het boekjaar 1987 blijven de bedrijven die gekozen zijn vanaf boekjaar 1982 in administratie. Van deze bedrijven is bekend in welke strata ze voorkwamen bij de keuze voor boekjaar 1986; deze strata zijn gebaseerd op metelling 1984. Een formeel juiste inpassing van de blijvende steekproefbedrijven in de strata van de "nieuwe" metelling 1985 - die gebruikt wordt bij de opzet voor boekjaar 1987 - zou inhouden dat ieder nieuw stratum zou moeten worden verdeeld in substrata die aangeven in welk mei-1984-stratum de betreffende populatie- en steekproefbedrijven vielen. Een dergelijke post-stratificatie - niet te verwarren met de post-stratificatie die uiteindelijk bij het berekenen van steekproefuitkomsten wordt uitgevoerd op basis van metelling 1987 - is echter in de praktijk onuitvoerbaar. Veel van de 275*275 substrata zullen vrijwel leeg blijven, omdat er relatief gezien weinig verschuivingen tussen strata plaatsvinden van jaar tot jaar. Alleen de substrata waarin bedrijven vallen die zowel in 1984 als in 1985 in hetzelfde stratum vallen, zullen in de praktijk voldoende grote aantallen bedrijven bevatten om als afzonderlijk stratum in aanmerking te komen. De overige substrata zullen moeten worden samengevoegd met substrata waarin bedrijven voorkomen die veel lijken op de bedrijven in de betreffende substrata. Dat houdt in de praktijk in dat alle bedrijven die in 1985 in een bepaald stratum voorkomen, ongeacht het stratum waarin ze voorkomen in 1984, geacht worden een stratum te vormen. De indeling van steekproefbedrijven in strata op basis van metelling 1985 wordt dus gebruikt zonder verder te letten op het stratum-verleden van de bedrijven en hun oorspronkelijke trekkingskansen (Lodder, 1983:17-18; Lodder, 1987:25).

8.3 Nieuwe bedrijven

Nadat voor ieder van de 275 strata is vastgesteld hoeveel steekproefbedrijven uit voorafgaande jaren daarin voorkomen, wordt bepaald in hoeverre deze stratum-vulling afwijkt van de gewenste allocatie (paragraaf 6.5 en 6.6). De allocatie van regionale aanvullingen (zie paragraaf 7.2) vindt in een later stadium plaats.

Er kunnen te weinig steekproefbedrijven in een stratum voorkomen, maar ook te veel. Een aantal oorzaken daarvan zijn:

- 1) de structuur van de populatie verandert en daardoor veranderen de populatie-aantallen en de gewenste steekproefaantallen;
- 2) blijvende bedrijven in het boekhoudnet komen soms in andere strata terecht door verandering in specifieke bedrijfssituaties, die niet noodzakelijk parallel lopen aan de veranderingen in de populatie-structuur;
- 3) hoewel de basisprincipes voor strata-vorming ieder jaar hetzelfde zijn, is de concrete uitwerking naar specifieke stratum-definities afhankelijk van de populatiestructuur. In het ene jaar kan bijvoorbeeld in een bepaald basisstratum wel een indeling in vijf oppervlakteklassen plaatsvinden en in het andere jaar zijn maar vier klassen mogelijk;
- 4) regionale aanvullingen ten behoeve van specifieke BUL-groepen in het verleden kunnen ertoe leiden dat in een aantal strata meer steekproefbedrijven voorkomen dan volgens de optimale allocatie gewenst is;
- 5) bedrijven die zes boekjaren geleden zijn gekozen, worden uitgesloten van verdere deelname, evenals bedrijven die verdere medewerking weigeren na aanvankelijke deelname (Lodder, 1983:18; Lodder, 1987:25-26).

Bedrijven die al in administratie zijn, maar volgens metelling 1985 buiten de populatie vallen, kunnen in sommige gevallen aangehouden worden, bijvoorbeeld indien verwacht mag worden dat het bedrijf de volgende boekjaren wel binnen de populatie zal vallen.

Bij een teveel aan bedrijven in een stratum wordt geprobeerd dit aantal te verminderen, tenzij het een stratum betreft waarin regionale aanvullingen noodzakelijk zijn. Daarbij worden met name die bedrijven uit het boekhoudnet gehaald die:

- 1) veel inzet van mankracht voor de administratieve verwerking van gegevens vereisen;
- 2) sterk van structuur zijn veranderd en niet meer lijken op het oorspronkelijk gekozen bedrijf, bijvoorbeeld door opvolging, fusie, en dergelijke (Lodder, 1983:18; Lodder, 1987:26).

Indien het aantal aanwezige bedrijven in een stratum niet toereikend is, worden met behulp van toevalsgetallen zoveel be-

drijven uit dit stratum getrokken als nodig is. Samen met de al aanwezige bedrijven worden ze geacht een toevallige steekproef uit dit stratum te zijn. Door de verschillende trekkingkansen van de blijvende bedrijven in het verleden, wordt er weliswaar enige afbreuk gedaan aan volledige aseleetheid, maar deze bedrijven behoren op het moment van trekking wel thuis in het betreffende stratum, en ze kunnen daarom de bedrijven in het stratum vrijwel even goed representeren als willekeurig gekozen bedrijven (zie paragraaf 9.2).

In bijlage 3 wordt een overzicht gegeven van de opbouw van het boekhoudnet in de boekjaren 1982-1987 naar jaar van eerste deelname van de bedrijven.

8.4 Vervangende keuze bij non-respons

Een bedrijf dat door middel van een aselechte steekproeftrekking binnen een stratum, wordt aangewezen voor deelname aan het boekhoudnet, kan bij nadere beschouwing ongeschikt zijn voor deelname of weigeren om deel te nemen.

Een bedrijf is in de volgende gevallen ongeschikt voor of niet in staat tot deelname:

- het bedrijf behoort volgens de meitelling (1985) wel tot de populatie, maar op het moment van keuze (voorjaar 1987) is het bedrijf opgeheven of is het zeer waarschijnlijk dat het bedrijf per 1 mei (1987) niet tot de populatie zal behoren (Lodder, 1987:26);
- het bedrijf is sterk van structuur veranderd en zal daardoor uiteindelijk in een stratum terechtkomen waarin geen bedrijven bijgekozen hoeven te worden;
- het bedrijf kan niet de volle periode van zes jaar in administratie blijven, bijvoorbeeld omdat het binnenkort opgeheven zal worden;
- er is sprake van ziekte, ongevallen of invaliditeit bij het bedrijfshoofd of bij de gezinsleden;
- het bedrijf is een onderzoeksbedrijf of een bedrijf in eigen beheer van de Rijksdienst voor de IJsselmeerpolders;
- het bedrijf kan de administrative gegevens over de landbouw niet zuiver verstrekken in verband met bijvoorbeeld verwevenheid met andere bedrijven. Het landbouwbedrijf is dan vaak een niet-zelfstandige tak van de onderneming, waarin ook loonwerk, veehandel, veevoederproductie- of handel of eierhandel voorkomt. Ook pluimvee-broederijen-slachterijen zijn ongeschikt;
- het bedrijf heeft land (al dan niet zaaiklaar verhuurd) in gebruik bij zaadteelt en zaadhandel, conservenindustrie en/of aardappelhandel;
- het bedrijf kent geen normale bedrijfsvoering door bijvoorbeeld sanering, ruilverkaveling of door ligging in een be-

- stemmingsplan; bedrijven in een ruilverkaveling zijn administratief vaak moeilijk te verwerken in verband met veelvuldige kavelruil, bemestingskosten en bewerkingskosten;
- het bedrijf is ongeschikt in verband met eerdere ervaringen (slecht insturen, het achterhouden van gegevens, enz.);
 - andere redenen zoals verhuizing, overdracht, reorganisatie.

Een bedrijf kan tenslotte wel geschikt zijn voor deelname en bovendien in staat tot deelname, maar toch weigeren om deel te nemen aan het boekhoudnet. Dit ondanks de veelbelovende tekst in de wervingsbrief voor het LEI-boekhoudnet: "Bij het kiezen van bedrijven is ook uw bedrijf in de keuze gevallen. Ik wend mij daarom tot u met het verzoek of u er mee accoord gaat dat het LEI voor u een boekhouding gaat verzorgen. Voor u is daar weinig werk aan verbonden; u krijgt daarvoor van ons een zeer uitgebreid overzicht van de economische gang van zaken op uw bedrijf, waarvan u veel nut kunt hebben. Deze boekhouding brengt voor u geen kosten mee" (van Dijk, 1986).

Redenen om niet deel te nemen zijn onder andere:

- geen tijd;
- geen belangstelling;
- onenigheid met landbouw-instanties over een geweigerde subsidie;
- teleurstelling over het gevoerde landbouwbeleid.

De laatste twee redenen komen maar weinig voor. In de praktijk blijkt het vaak moeilijk te achterhalen wat de werkelijke reden tot weigering is.

Kan of wil een bedrijf niet deelnemen, dan wordt in plaats van het oorspronkelijk gekozen bedrijf een ander bedrijf in hetzelfde stratum gekozen. Dit vervangende bedrijf wordt niet select gekozen. Als vervangend bedrijf wordt dat bedrijf aangewezen dat in termen van de vijf stratificatie-variabelen - en dan ook in dezelfde volgorde van belangrijkheid als bij de vorming van strata - het meest lijkt op het niet-deelnemende bedrijf (zie paragraaf 5.7.2). Daarmee worden systematische afwijkingen van de steekproefbedrijven ten opzichte van de populatiebedrijven tegengegaan. Een bedrijf van het type "akkerbouwbedrijf" dat voorkomt in sbe-klasse 3, oppervlakte-klasse 5, leeftijdsklasse 2 en regio 5, wordt dus vervangen door een bedrijf dat voor al deze criteria dezelfde waarden heeft, zonder overigens te letten op de precieze sbe-grootte, oppervlakte of leeftijd.

Weigert een vervangend bedrijf om mee te doen dan wordt op dezelfde wijze een nieuwe vervanger aangewezen. Is het aan het begin van het boekjaar nog niet gelukt om een vervanger te vinden, dan wordt de keuze gestaakt.

9. De berekening van steekproefuitkomsten

9.1 Inleiding

Op het moment dat de gegevens van het boekhoudnet 1987 beschikbaar zijn (voorjaar 1989), is er ook ten aanzien van de structuur van de populatie gedurende dit boekjaar meer informatie aanwezig uit de metellingen van 1987 en 1988. Terwijl de keuze van de bedrijven gebaseerd was op de metelling 1985, gebeurt het schatten van totaalcijfers en gemiddelden voor de gehele landbouw en voor groepen bedrijven daarbinnen, op basis van metelling 1987.

In dit hoofdstuk wordt beschreven hoe voor de individuele steekproefbedrijven een wegingsfactor wordt berekend die aangeeft welk gewicht het bedrijf - gezien de samenstelling van de achterliggende populatie - moet krijgen in de berekening van steekproefuitkomsten.

9.2 Stratificatie achteraf

De afgrenzing en indeling van de steekproefpopulatie waarop het boekhoudnet zich bij de keuze van bedrijven richt, wordt bepaald op grond van de meest recente metelling waarover gegevens beschikbaar zijn. Dat betekent in de praktijk dat op het moment van de keuze met enigszins verouderde gegevens moet worden gewerkt. Voor de steekproef die cijfers moet leveren voor het boekjaar 1987 - dat loopt van 1 mei 1987 tot en met 30 april 1988 - wordt bij de steekproefopzet (eind 1986) uitgegaan van de gegevens van de metelling 1985. Deze populatie wordt in strata ingedeeld en er worden steekproefbedrijven gekozen. Begin 1989, als de uitgewerkte boekhoudingen beschikbaar zijn, wordt voor dit zelfde boekjaar 1987 een nieuwe stratificatie uitgevoerd op basis van metelling 1987. De bedrijven met uitgewerkte boekhoudingen worden in de ontstane strata ingedeeld, evenals alle populatiebedrijven. Daarna worden de wegingsfactoren per steekproefbedrijf vastgesteld via deling van het aantal populatie-bedrijven en het aantal boekhoudnet-bedrijven. Deze procedure zorgt ervoor dat uitgegaan wordt van een zo recent mogelijke structuur van de bedrijven, wat de representativiteit ten goede komt. Bovendien wordt daardoor gecorrigeerd voor de non-respons die is opgetreden tussen de keuze van de bedrijven en de berekening van steekproefuitkomsten.

De vorming van strata op basis van metelling 1987 gebeurt door optimale stratificatie. Deze wordt tegelijkertijd gebruikt voor de opzet van de steekproef voor boekjaar 1989.

De gang van zaken bij het stratificeren van de populatie in de meitelling 1987 heeft wel enige praktische en theoretische haken en ogen. De strata op basis van meitelling 1987 en meitelling 1985 zijn weliswaar op grond van dezelfde criteria gevormd, maar ze hoeven niet per definitie identiek te zijn. De aantallen populatie- en steekproefbedrijven zullen per stratum verschillend zijn in beide tellingen. Bedrijven kunnen opgeheven worden, bedrijven kunnen van karakter veranderen en daardoor in een ander stratum terechtkomen, en er kunnen bedrijven tot de populatie toetreden die daar voorheen niet toe behoorden. Lodder (1983) gaat op deze problematiek in. Hij komt tot de conclusie dat theoretisch gezien ieder stratum in de populatie op basis van de meitelling aan het begin van het boekjaar - in ons voorbeeld meitelling 1987 - moet worden verdeeld in substrata met die populatiebestanddelen die in mei 1985 in aparte strata vielen (Lodder, 1983:19-20). De achtergrond daarvan wordt gevormd door het feit dat bedrijven die in mei 1987 in een bepaald stratum vallen, niet allemaal dezelfde trekkingkans hebben gehad in het verleden, vanwege de verschillen in steekproefpercentage tussen de basisstrata.

Om gelijksoortige redenen als bij de bedrijfskeuze wordt een dergelijke post-stratificatie niet uitgevoerd. Vanwege het leeg zijn van veel van de substrata - zeker voor wat betreft het aantal steekproefbedrijven - wordt ook hier ieder stratum op basis van meitelling 1987 als een eenheid beschouwd, waarin alle daarin voorkomende steekproefbedrijven geacht worden een zelfde trekkingkans te hebben gehad in het verleden. Voor zover bedrijven tussen 1985 en 1987 van stratum zijn veranderd, maar niet van basisstratum, is het uitgangspunt van gelijke trekkingkansen bij benadering correct. Bedrijven die ook van basisstratum veranderen, komen meestal in een basisstratum terecht, waar de trekkingkansen weinig afwijken van het oorspronkelijke stratum; overgangen van bijvoorbeeld sbe-klasse 1 naar sbe-klasse 4 komen weinig voor.

Ook zullen bedrijven die nieuw zijn in een stratum in het algemeen geen grote verschillen vertonen met blijvende populatie- en steekproefbedrijven in dat stratum, waardoor een verschil in trekkingkansen nauwelijks gevolgen heeft voor de steekproefuitkomsten.

9.3 Meitelling per begin boekjaar als basis

De enige basis voor een (post-)stratificatie van de populatie van landbouwbedrijven, is de meitelling. Het LEI kiest de telling aan het begin van het boekjaar als stratificatie-basis. Daarmee wordt bereikt dat de gegevens die de steekproef levert, betrekking hebben op de landbouw-structuur zoals die aan het begin van het boekjaar was. Dat gekozen wordt voor de meitelling

aan het begin van het boekjaar, en niet voor die aan het eind van het boekjaar, heeft te maken met de samenhang tussen stratificatie- en doelvariabelen. "Voor de niet-grond-gebonden produktie zal stratificatie per begin of per eind boekjaar niet veel uitmaken maar voor de grond-gebonden produktie wel omdat de resultaten van de akkerbouwsector uiteraard bepaald worden door de areaalen die voor dit oogstjaar ingezet zijn en niet door de areaalen die ingezet worden voor het volgende boekjaar" (Lodder, 1983:20).

9.4 Verschil tussen doelpopulatie en keuze-populatie

De doelpopulatie voor het boekhoudnet 1987 wordt gevormd door alle bedrijven die volgens de metelling per 1 mei 1987 aan de gestelde populatie-criteria voldoen (zie hoofdstuk 4). De steekproefbedrijven voor dit boekjaar zijn echter getrokken op basis van metelling 1985. Bedrijven die tussen 1 mei 1985 en 1 mei 1987 gesticht zijn, of op andere wijze tot de populatie zijn gaan behoren, hebben dus geen vertegenwoordigers uit eigen gelederen in het boekhoudnet 1987, maar horen wel bij de populatie die beschreven wordt. Ze worden daar - via de wegingsfactoren - vertegenwoordigd door bedrijven die al langer dan twee jaar tot de populatie behoren. Omdat veel nieuwe bedrijven een bedrijfshoofd hebben dat jonger is dan 35 jaar, en alle bedrijven met een bedrijfshoofd jonger dan 35 jaar in de stratificatie zoveel mogelijk in afzonderlijke strata terechtkomen (paragraaf 5.6 en bijlage 1), zullen deze nieuwe bedrijven worden vertegenwoordigd door de bedrijven die er het meest op lijken. Temeer omdat naar vijf metellingscriteria wordt gestratificeerd.

Bedrijven die langer dan twee jaar maar korter dan acht jaar in de populatie zitten, zijn overigens ook ondervertegenwoordigd in het boekhoudnet. Immers, elk jaar wordt slechts een deel van de bedrijven in het boekhoudnet vervangen door andere bedrijven. "Daardoor kan gesteld worden dat bij aggregatie op basis van blijvende bedrijven aggregatiefouten voor zullen komen, met name voor kengetallen die sterk samenhangen met deze, toch regelmatig voorkomende bedrijfssituaties" (Lodder, 1983:20).

En andere groep bedrijven die wel impliciet in de berekeningen voor het boekjaar 1987 wordt opgenomen, maar niet in het boekhoudnet 1987 voorkomen, zijn de bedrijven die in de loop van het boekjaar opgeheven worden. Ook hier treedt enige vertekening op doordat elk bedrijf in metelling 1987 - dus ook elk opgeheven bedrijf - vertegenwoordigd wordt door het gemiddelde van de steekproefbedrijven in het stratum waarin dat metellingsbedrijf valt. Immers, een blijvend bedrijf zal in het algemeen gemiddeld gedurende het boekjaar een grotere bedrijfsoppervlakte, meer

koeien, een beter bedrijfsresultaat, enz., hebben dan een verdwijnend bedrijf 1).

Deze vertekening komt nog eens boven op de vertekening die optreedt doordat ondernemers die naar verwachting niet de volle periode van zes jaar mee kunnen doen in de steekproef, niet voor deelname in aanmerking komen (zie hoofdstuk 8). De beide vertekeningen worden enigszins gecorrigeerd door het onderscheiden van een aparte leeftijdsklasse voor ondernemers boven de 60 jaar. Veel bedrijven die opgeheven worden, hebben een ondernemer die ouder is dan 60 jaar. Daardoor worden deze bedrijven, voor zover leeftijd een rol speelt bij de stratificatie, gerepresenteerd door bedrijven die er het meest op lijken.

Figuur 9.1 geeft aan wat de verschillen zijn tussen de populatie die wordt gehanteerd bij de keuze van de steekproefbedrijven en de populatie bij het berekenen van de steekproefuitkomsten (meitelling 1987).

Bij de berekening van steekproefuitkomsten wordt uitgegaan van alle bedrijven die volgens de meintelling van 1987 (laatste kolom van figuur 9.1) voldoen aan de definitie van de steekproefpopulatie. Ook de nieuw toetgetreden populatiebedrijven doen mee in de weging, evenals de bedrijven die binnenkort zullen worden opgeheven en dientengevolge ongeschikt zijn voor opname in de steekproef. Op de consequenties hiervan zal in hoofdstuk 11 worden teruggekomen.

Een laatste punt dat hier genoemd moet worden is een overschatting van totaalcijfers en gemiddelden door het niet geheel aansluiten van de elementen in de steekproef op de elementen in de populatie. De identificatie van bedrijven in de populatie loopt via de gemeente-registratie-nummers van ondernemers in de meintelling. In sommige gevallen komt een bedrijf onder meerdere nummers voor in de meintelling. Dergelijke bedrijven hebben een grotere kans om in de steekproef te vallen. Ze zijn vaak groter dan de overige bedrijven in het stratum waarin ze voorkomen. Dat wordt veroorzaakt door het feit dat een bedrijf met meerdere ondernemers, in de meintelling soms evenredig verdeeld wordt over meerdere ondernemers (registratie-nummers). Een bedrijf met bijvoorbeeld 300 sbe en twee ondernemers, komt dan in de meintelling voor als twee bedrijven met elk 150 sbe. Indien één van beide bedrijven van 150 sbe in de steekproef voorkomt, wordt het gehele bedrijf van 300 sbe in de steekproefuitkomsten meegenomen.

Indien dergelijke bedrijven in de toekomst meer gaan voorkomen zal daar in de weging rekening mee gehouden moeten worden.

1) Om dit te voorkomen zou gewerkt moeten worden met een soort gemiddelde van het aantal bedrijven in de meintellingen aan het begin en aan het eind van het boekjaar.

Groep bedrijven	Bedrijven komen wel (+) of niet (-) voor in de betreffende populatie		
	Populatie meitelling 1985	Keuzepopulatie boekhoudnet 1987	Populatie meitelling 1987
Bedrijven die zowel in 1985 als in 1987 niet aan de populatie-definitie voldoen	-	-	-
Bedrijven die alleen in 1985 aan de pop. definitie voldoen	+	-	-
Bedrijven die alleen in 1987 aan de pop. def. voldoen	-	-	+
Bedrijven waarvan de ondernemer te kennen geeft dat het bedrijf binnen 6 jaar uit de pop. zal verdwijnen	+	-	?
Overige bedrijven	+	+	+

Figuur 9.1 Het voorkomen van specifieke groepen bedrijven in de meitellingspopulaties van 1985 en 1987 en in de keuzepopulatie voor boekjaar 1987

9.5 Wegingsfactoren

Per stratum wordt bepaald hoeveel bedrijven er voorkomen in de populatie en in de steekproef. Bij de steekproef gaat het uiteraard alleen om die bedrijven waarvoor - op het moment dat de BUL vervaardigd wordt - een volledige boekhouding beschikbaar is en niet om alle bedrijven die oorspronkelijk in administratie genomen zijn.

Door populatie- en steekproefaantallen op elkaar te delen, worden wegingsfactoren bepaald. Deze factoren geven aan wat het gewicht van een bedrijf is in de berekeningen. Ze zijn nodig om de populatie-structuur goed weer te geven. Een wegingsfactor geeft aan hoeveel bedrijven in de populatie een bepaald steekproefbedrijf vertegenwoordigt.

Bij de berekening van bedrijfsuitkomsten (BUL) worden andere wegingsfactoren gebruikt dan bij berekeningen ten aanzien van de

financiële positie van de ondernemer (FIP). Dit hangt samen met verschillen in aantallen steekproefbedrijven per stratum waarop de berekeningen betrekking hebben. Een klein deel van de bedrijven waarvoor wel bedrijfsuitkomsten beschikbaar zijn, levert geen gegevens over de financiële positie van de ondernemer. Voor de FIP zijn dus per stratum minder bedrijven beschikbaar dan voor de BUL en daarom zijn de wegingsfactoren per bedrijf voor FIP-berekeningen groter dan of gelijk aan die voor BUL-berekeningen.

9.6 Samenvoeging van strata

Voor het berekenen van standaardfouten van steekproefschattingen, is het wenselijk om over minimaal twee waarnemingen per stratum te beschikken 1). Hoewel bij de steekproefopzet per stratum minimaal twee bedrijven worden gekozen, komt het toch voor dat er in de uiteindelijke strata minder dan twee bedrijven een uitgewerkte boekhouding leveren.

Dergelijke strata worden gecombineerd met andere strata die er in termen van de stratificatie-variabelen veel op lijken. Voor strata waarin één steekproefbedrijf voorkomt, kan eventueel eerst een wegingsfactor worden berekend, voordat samenvoeging met een ander stratum plaatsvindt.

Bij de berekeningen voor de BUL en de FIP worden strata met één steekproefbedrijf niet samengevoegd met andere strata. Bij het gebruik van formule (6.4.7) in plaats van (6.4.5) voor het berekenen van standaardfouten is dat immers niet noodzakelijk.

9.7 Post-stratificatie

Naast de in paragraaf 9.2 beschreven vorm van stratificatie nadat de steekproef is getrokken, is het te overwegen om voor specifieke doeleinden alternatieve post-stratificaties uit te voeren.

Indien er een sterk verband is tussen een doelvariabele en een meetingsvariabele die geen rol in de "normale" stratificatie speelt, kan de betrouwbaarheid van schattingen voor die doelvariabele vergroot worden door in de stratificatie ook rekening te houden met de betreffende meetingsvariabele. Indien bijvoorbeeld een gemiddelde waarde van een variabele geschat moet worden waarbij bekend is dat er een sterk verband is tussen de aanwezigheid van een ligboxenstal op een bedrijf en de waarde van de doelvariabele, kan de betrouwbaarheid van die schatting vergroot worden door bij de stratificatie de bedrijven ook in te de-

1) Er zijn ook methoden om bij één waarneming per stratum standaardfouten te berekenen (Cochran, 1977:138-140).

len naar het al dan niet aanwezig zijn van een ligboxenstal. De nieuwe strata zullen dan immers homogener zijn in termen van de doelvariabele. Voorwaarde is echter dat voor alle bedrijven in de populatie bekend moet zijn of er al dan niet een ligboxenstal aanwezig was. In 1984 was dit gegeven in de metelling opgenomen, zodat voor dat jaar een dergelijke post-stratificatie mogelijk is.

In hoofdstuk 14 blijkt dat met quotiënt- en regressieschat- ters vaak hetzelfde effect bereikt kan worden als met post-strati- ficatie.

Een andere toepassing van post-stratificatie richt zich niet zozeer op de te schatten variabele als wel op de groep bedrijven waarvoor schattingen moeten worden bepaald. Voor deelpopulaties die dwars door strata heen lopen ("cross classes") kan de be- trouwbaarheid van schattingen vergroot worden indien de aantallen van de deelpopulaties per stratum exact bekend zijn. De vergro- ting van de betrouwbaarheid zit dan meestal niet in het homogener worden van de strata, maar in het opheffen van de schattingsfout in de wegingsfactoren (zie ook hoofdstuk 13). Een voorbeeld van een schatting voor "cross classes" is de schatting van het aantal meewerkende kinderen op bedrijven in de provincie Friesland. Doordat de regio's die in de oorspronkelijke stratificatie worden gehanteerd niet overeenkomen met de provinciegrenzen, kan het aantal bedrijven in Friesland niet exact worden vastgesteld uit de steekproef. Door in de stratificatie de provinciale ligging van bedrijven mee te nemen, kan de betrouwbaarheid van de schat- ting vergroot worden.

Bij de "normale" stratificatie achteraf (paragraaf 9.2) wordt niet expliciet rekening gehouden met de oorspronkelijke stratificatie vooraf, omdat beide stratificaties in grote lijnen overeenkomen (paragraaf 8.2). Bedrijven in een bepaald post-strat- um zullen allemaal ongeveer dezelfde trekkingskansen hebben. Bij een post-stratificatie waarbij andere criteria worden gehanteerd dan bij de oorspronkelijke stratificatie, moet wel expliciet re- kering gehouden worden met de strata die vooraf zijn gevormd en de verschillen in trekkingskansen tussen de strata. Dat betekent dat per oorspronkelijk stratum een post-stratificatie moet worden uitgevoerd. Omdat voor strata binnen een basisstratum geldt dat de trekkingspercentages niet ver uiteenlopen, kan in de praktijk worden volstaan met een post-stratificatie per basisstratum.

Alleen indien de trekkingskansen van een bedrijf volledig los- staat van de te schatten variabele, omdat er geen samenhang is tussen deze variabele en de oorspronkelijke stratificatie-variabe- len, kan de post-stratificatie plaatsvinden zonder rekening te houden met de oorspronkelijke strata (Lodder, 1987:31). Er zijn echter maar weinig metellingsvariabelen die in een post-strati- ficatie een rol zouden kunnen spelen - vanwege hun verband met de doelvariabele - en die losstaan van de oorspronkelijke vijf strati- ficatie-variabelen.

In het kader van het EG-informatienet wordt door de EG ook een post-stratificatie uitgevoerd naar EG-bedrijfstype en EG-grootteklasse. Daarbij wordt geen rekening gehouden met de oorspronkelijke stratificatie en de bijbehorende trekkingskansen. Dit leidt in de praktijk tot vertekende steekproefuitkomsten. Immers, omdat in een EG-stratum bedrijven terechtkomen met verschillende historische trekkingskansen, en grotere bedrijven een grotere trekkingskans hebben dan kleinere bedrijven, zullen de schattingen per stratum voor veel grootheden te hoog uitkomen. Gemiddelden over strata zullen eveneens overschat worden. Ook de verschillen in trekkingskansen tussen de verschillende akkerbouwgebieden als gevolg van de aanvullende keuzes ten behoeve van BUL-groepen, zullen in de EG-cijfers leiden tot vertekening.

10. Non-respons

10.1 Inleiding

In de hoofdstukken 5 tot en met 9 is beschreven welke werkwijze gevolgd wordt om de doelstellingen van de steekproef (hoofdstuk 4) te realiseren. In dit en de volgende hoofdstukken van deel 2 wordt nagegaan in hoeverre de doelstellingen in de boekjaren vanaf 1975 ook daadwerkelijk zijn gerealiseerd. De aandacht wordt daarbij gericht op de representativiteit en de betrouwbaarheid van de steekproefuitkomsten.

In dit hoofdstuk staat de analyse van de non-respons centraal. Het vormt een inleiding op de bespreking van de representativiteit in hoofdstuk 11. De representativiteit van steekproefuitkomsten kan namelijk beïnvloed worden door mogelijke verbanden tussen deze uitkomsten en factoren die samenhangen met de non-respons. In hoofdstuk 11 blijkt dat non-respons niet hoeft te leiden tot een aantasting van de representativiteit. De steekproefopzet (hoofdstukken 5-8) en de wijze van berekening van steekproefuitkomsten (hoofdstuk 9) dragen daar in de praktijk zorg voor.

In paragraaf 2 wordt ingegaan op het meten van (non-)respons voor een steekproef waarbij gewerkt wordt met vervangende bedrijven bij non-respons. In de paragrafen 3 tot en met 5 wordt informatie gegeven over drie vormen van non-respons die een rol spelen bij de LEI-steekproef van landbouwbedrijven. In paragraaf 3 wordt inzicht gegeven in de samenhang die er is tussen bedrijfs- en ondernemerskenmerken en de respons bij de werving van bedrijven. Ook worden cijfers gegeven over de mate waarin de werving ook daadwerkelijk leidt tot opname van het gewenste aantal bedrijven in de steekproef. Paragraaf 4 bespreekt de uitval van bedrijven tijdens het boekjaar. De partiële non-respons voor financieringsuitkomsten komt in paragraaf 5 aan de orde.

10.2 Respons

Zoals bijna iedere steekproef, heeft ook de LEI-steekproef te kampen met non-respons. Sommige bedrijven zijn niet in staat om mee te doen, andere weigeren deelname, en weer andere leveren onvoldoende gegevens voor het samenstellen van een volledige administratie. In de praktijk van de LEI-steekproef blijkt dat bedrijven die deel willen nemen aan de steekproef gemiddeld iets andere kenmerken hebben dan alle bedrijven die aangezocht worden voor deelname. Dat hoeft op zichzelf de representativiteit van de steekproefuitkomsten niet te ondermijnen. Door bij de keuze van

bedrijven rekening te houden met de bedrijfskenmerken die samenhangen met de non-respons en door er ook achteraf rekening mee te houden bij de berekening van steekproefuitkomsten, kan voorkomen worden dat er systematische verschillen ontstaan tussen steekproef en populatie.

"Respons" kan op verschillende manieren worden gedefinieerd. Het aantal deelnemende bedrijven in procenten van het aantal aangezochte bedrijven (bij de bedrijfskeuze) en het aantal bedrijven met een uitgewerkte boekhouding in procenten van alle bedrijven in de steekproef (in de verwerkingsfase), lijken voor de hand liggende maten voor het meten van de "responsgeneigdheid" van een groep bedrijven. Toch zeggen deze percentages niet alles over het respons-gedrag in de populatie. Bij de bedrijfskeuze hebben dergelijke responspercentages immers betrekking op alle bedrijven die zijn aangezocht. Wordt alleen uitgegaan van de reacties van de "eerste keus"-bedrijven - gedefinieerd als alle bedrijven die in eerste instantie zijn aangezocht; dus met uitsluiting van alle bedrijven die dienen als vervanger voor een niet-deelnemend bedrijf -, dan zullen in het algemeen hogere respons-percentages resulteren. Immers een vervangend bedrijf heeft ongeveer dezelfde kenmerken als het te vervangen bedrijf (zie paragraaf 8.4), en zal qua respons-beïnvloedende factoren dus ook lijken op dit bedrijf. Daardoor zal de groep van vervangende bedrijven een hogere non-respons te zien geven dan de groep van oorspronkelijk gekozen bedrijven. Wordt er toch uitgegaan van alle aangezochte bedrijven dan wordt een reëler beeld van de totale respons in een bepaald jaar verkregen door de respons-percentages per stratum te wegen met het aantal steekproefbedrijven dat daarin in dat jaar bijgekozen moet worden.

Verder hangt de respons bij de bedrijfskeuze in een bepaald jaar nogal af van de delen van de populatie waar de bedrijven gekozen moeten worden. Bepaalde delen van de populatie - bijvoorbeeld de bedrijven met jonge ondernemers - zullen relatief sterk vertegenwoordigd zijn in de "eerste keus". De keuze van nieuwe steekproefbedrijven betreft immers een aanvulling van ongeveer 25% op de bedrijven die resteren uit de steekproef van het voorgaande jaar. Daarnaast is de tussentijdse uitval van bedrijven niet helemaal onafhankelijk van de bedrijfskenmerken (paragraaf 10.4). Bovendien zorgen de verschillen in steekproefpercentage tussen delen van de populatie ervoor dat sommige populatie-delen sterker bij de keuze van bedrijven betrokken zijn dan andere delen en dus ook zwaarder meewegen in de berekening van respons-percentages. Dat laatste geldt ook voor de non-respons in de verwerkingsfase.

De genoemde factoren leiden ertoe dat de responspercentages op basis van "eerste keus"-bedrijven steekproef-specifiek zijn en niet gelden voor de populatie in het algemeen.

Een potentieel responspercentage voor de hele populatie zou berekend kunnen worden door de responspercentages per stratum te

wegen met de populatie-aantallen per stratum. Een dergelijke berekeningswijze omzeilt het probleem van de verschillen in steekproefpercentage tussen groepen bedrijven die in de hiervoor genoemde alternatieven doorwerken in de resultaten.

Afhankelijk van het doel kan de ene berekeningswijze de voorkeur verdienen boven de andere. Het doel kan zijn om aan te geven hoeveel moeite het kost om het vastgestelde aantal bedrijven in de steekproef te kunnen opnemen, maar ook om de responsgevoeligheid van de "eerste keus"-bedrijven of van de bedrijven in de populatie te bepalen. In de volgende paragrafen wordt steeds aangegeven welk quotiënt gebruikt wordt als respons-percentage. In de meeste gevallen gaat het om respons-percentages die betrekking hebben op de "eerste keus"-bedrijven.

10.3 Non-respons bij de bedrijfskeuze

10.3.1 Gangbare berekening van non-respons

Voor ieder boekjaar berekent het LEI welke responspercentages de bedrijfskeuze oplevert (van Dijk, 1975-1987). Deze percentages geven aan welk deel van alle aangezochte bedrijven toegezegd heeft om deel te nemen aan de steekproef. De percentages worden berekend voor de acht bedrijfstypen die bij de stratificatie onderscheiden worden. Tabel 10.1 geeft een overzicht van dit beschikbare materiaal voor de boekjaren 1981 tot en met 1987.

Tabel 10.1 Aantal bedrijven dat deelname aan het boekhoudnet toezegt in procenten van het aantal aangezochte bedrijven naar bedrijfstype en boekjaar

Bedrijfstype	Boekjaar							
	81	82	83	84	85	86	87	81-87
Akkerbouw	50	58	30	40	27	42	38	37
Rundvee	50	42	47	57	50	50	46	49
Varkens	34	28	31	32	27	22	22	28
Pluimvee	30	22	31	9	24	38	16	24
Gemengd akkerbouw	30	16	57	33	55	22	38	32
Gemengd rundvee/akkerbouw	36	41	40	40	56	40	35	41
Gemengd rundvee/int. veeh.	65	50	45	43	33	44	38	44
Gemengd intensieve veeh.	35	31	22	19	41	21	36	27
Regionale aanvulling (akk.)	47	32	42	34	41	35	34	37
Alle bedrijven	42	36	36	39	38	39	33	38

Bron: Van Dijk, 1981-1987.

Deze tabel laat duidelijk zien dat de hoogste respons (40-50%) wordt gehaald in de bedrijfstypen waarin rundveehouderij de hoofdactiviteit is. Bedrijven met intensieve veehouderij zijn beduidend minder geneigd of in staat om deel te nemen aan de steekproef (20-30% respons). De bedrijfstypen waar akkerbouw de boventoon voert, zitten qua responsgedrag in een middenpositie (ongeveer 35% respons). In de boekjaren 1981-1987 zijn totaal 3.530 bedrijven aangezocht. Daarvan zegden 1.344 (38%) medewerking aan de steekproef toe.

In paragraaf 2 is aangegeven dat bij deze wijze van berekening naar verwachting lagere responspercentages zullen resulteren dan bij berekening op basis van "eerste keus"-bedrijven. Deze laatste berekening is relevanter indien de aandacht uitgaat naar de "respons-geneigdheid" in de steekproef en niet zozeer naar de moeite die het kost om een bepaald aantal steekproefbedrijven te werven. De verschillen in responspercentage tussen de bedrijfstypen zullen echter in grote lijnen de zelfde zijn bij beide berekeningsmethoden. Alleen het niveau waarop de respons zich bevindt zal verschillen. Voor de boekjaren 1985, 1986 en 1987 bijvoorbeeld bedraagt de respons van de "eerste keus"-bedrijven respectievelijk 46%, 45% en 38% en van alle aangezochte bedrijven respectievelijk 38%, 39% en 33%. In de volgende subparagrafen wordt steeds uitgegaan van de "eerste keus"-bedrijven.

Een andere kanttekening die bij tabel 10.1 geplaatst moet worden is dat de indeling in bedrijfstypen is gebaseerd op de kenmerken van bedrijven in de metelling waarop de keuze is gebaseerd. Dat is de metelling die twee jaar voor het moment van keuze heeft plaatsgevonden. Daardoor bevat de populatie waaruit gekozen wordt ook bedrijven die inmiddels zijn opgeheven of om andere redenen niet meer tot de populatie behoren. Deze bedrijven nemen uiteraard niet deel aan de steekproef. Ze zijn echter wel opgenomen in de berekening van de responspercentages; deze worden daardoor licht negatief beïnvloed. In de volgende subparagrafen wordt zoveel mogelijk uitgegaan van de bedrijfskenmerken in de metelling in het jaar van keuze. Een dergelijke indeling heeft bovendien het voordeel dat verschillen in respons tussen verschillende groepen bedrijven nog iets duidelijker uit de verf komen.

10.3.2 De keuze in de boekjaren 1985-1987

Voor de boekjaren 1985, 1986 en 1987 is een nadere analyse van het verband tussen responspercentages en bedrijfskenmerken uitgevoerd. De responspercentages zijn daarbij berekend als de verhoudingen tussen alle aangezochte en alle deelnemende "eerste keus"-bedrijven.

In deze drie boekjaren moesten totaal 853 bedrijven worden bijgekozen voor deelname aan de steekproef. In eerste instantie werden daarvoor 853 bedrijven benaderd. In een aantal gevallen

was het aangewezen bedrijf geschikt en bereid om deel te gaan nemen aan de steekproef. In andere gevallen moesten vervangende bedrijven worden aangewezen. Totaal werden er in de genoemde boekjaren 2.129 bedrijven aangezocht voor deelname aan het boekhoudnet. Uiteindelijk leverde dat 775 bedrijven op die instemden met het verstrekken van gegevens aan het boekhoudnet. Er werden dus 72 (853-775) bedrijven minder in het boekhoudnet opgenomen dan de bedoeling was. Dat houdt in dat de bedrijfskeuze voor 91% is geslaagd. Om toch aan voldoende steekproefbedrijven in sommige strata te komen, zijn in de boekjaren 1986 en 1987 respectievelijk 6 en 7 bedrijven die al 7 jaar in administratie waren, aangehouden in het boekhoudnet.

De cijfers in de volgende subparagrafen hebben betrekking op de totale non-respons bij de bedrijfskeuze. In subparagraaf 10.3.9 wordt een splitsing naar de reden van non-respons gegeven.

10.3.3 Non-respons en bedrijfstype

Tabel 10.1 liet zien dat er sterke verschillen in respons zijn tussen de bedrijfstypen. Ook bij berekening op basis van "eerste keus"-bedrijven en op basis van de meitelling in het jaar van de keuze worden deze verschillen zichtbaar. Tabel 10.1 beperkt zich tot de acht bedrijfstypen die worden onderscheiden bij de stratificatie. Een meer gedetailleerde indeling in bedrijfstypen in tabel 10.2 brengt de samenhang tussen bedrijfstype en respons nog scherper naar voren. Deze indeling sluit aan bij de NEG-typering die sinds 1986 in Nederland gebruikt wordt (zie deel 3 van dit verslag).

In tabel 10.2 zijn alleen de bedrijven opgenomen die in de meitelling in het jaar van keuze waren terug te vinden. Bovendien zijn de bedrijven die op basis van deze meitelling meer dan 50% tuinbouw bezaten, uitgesloten van de berekeningen.

Binnen elk van de hoofdgroepen blijken de responspercentages nogal te variëren tussen subgroepen. Bij de akkerbouwbedrijven hebben de "overige akkerbouwbedrijven" een relatief lage respons. De met deze groep bedrijven verwante "gewassencombinaties" en "gewassen/veeteeltcombinaties" scoren eveneens laag. Bij de graasdierbedrijven blijkt de mate van specialisatie in de melkveehouderij samen te hangen met de respons. Kalvermestrijbedrijven hebben een responspercentage dat vergelijkbaar is met dat van de andere groepen intensieve veehouderijbedrijven. Mestvarkensbedrijven hebben binnen de groep hokdierbedrijven de hoogste respons.

Tabel 10.2 geeft dezelfde hoofdlijnen weer als tabel 10.1. Het blijft echter een eendimensionale tabel. Verschillen in respons hangen samen met meerdere factoren tegelijk. Daarom kunnen uit tabel 10.2 nog maar voorlopige conclusies getrokken worden. De volgende subparagrafen schenken achtereenvolgens aandacht aan de factoren bedrijfsomvang, bedrijfsoppervlakte, leeftijd en opleiding van de ondernemer en regio.

Tabel 10.2 Responspercentage en aantal aangezochte bedrijven
naar NEG-bedrijfstype (boekjaren 1985-1987)

Bedrijfstype	Aantal bedrijven	Respons- percentage
Akkerbouwbedrijven:		
Gespecialiseerde hakvruchtenbedrijven	69	42
Graan/hakvruchtenbedrijven	17	41
Overige hakvruchtenbedrijven	59	40
Overige akkerbouwbedrijven	27	33
Graasdierbedrijven:		
Sterk gespecialiseerde melkveebedrijven	278	56
Gespecialiseerde en ov. melkveebedrijven	48	40
Kalvermesterijbedrijven	7	29
Overige rundveebedrijven	12	33
Andere graasdierbedrijven	12	42
Hokdierbedrijven:		
Fokvarkensbedrijven	45	24
Mestvarkensbedrijven	25	44
Andere varkensbedrijven	40	33
Legkippenbedrijven	34	29
Slachtpluinveebedrijven	6	33
Andere hokdierbedrijven	7	28
Combinatiebedrijven:		
Gewassencombinaties	10	20
Veeteeltcombinaties	67	42
Gewassen/veeteeltcombinaties	38	31

10.3.4 Non-respons en bedrijfsgrootte

De non-respons blijkt niet alleen te verschillen tussen de bedrijfstypen, maar ook tussen verschillende sbe-klassen per bedrijfstype. Tabel 10.3 geeft voor alle bedrijfstypen samen de respons-percentages per sbe-klasse. Weliswaar is er een lichte samenhang tussen de grootte van een bedrijf en het bedrijfstype, maar dat verhindert niet dat de cijfers in deze tabel duidelijk aangeven wat er aan de hand is.

De lagere respons in de laagste en de hoogste sbe-klassen vallen op. Voor de laagste sbe-klasse vindt dit voor een deel zijn oorzaak in een relatief groot aantal niet-geschikte bedrijven (zie paragraaf 10.3.9). De bedrijven die kleiner zijn dan

79 sbe in de metelling in het jaar van keuze, horen inmiddels niet meer tot de populatie. In gevallen waarin een bedrijf slechts incidenteel kleiner dan 79 sbe is, wordt toch overgegaan tot opname in de steekproef. Bij de kleinere bedrijven zal in het algemeen gelden dat er een grotere kans is dat het bedrijf geen zes jaar kan deelnemen aan de steekproef of dat het bedrijf op wat langere termijn zal worden opgeheven waardoor de interesse om deel te nemen minder groot is.

Tabel 10.3 Respons-percentage en aantal aangezochte bedrijven naar sbe-aantal (boekjaren 1985-1987)

Aantal sbe	Aantal Bedrijven	Respons-percentage
< 79	45	9
79 - 100	58	33
100 - 125	72	39
125 - 150	76	46
150 - 200	117	45
200 - 250	114	51
250 - 300	71	54
300 - 350	60	55
350 - 500	102	45
500 - 750	65	37
750 - 1.000	17	29
>=1.000	8	25

Per bedrijfstype is een zelfde patroon in de respons waar te nemen als voor alle bedrijven gezamenlijk. Hoe groter het bedrijf des te groter de kans op respons, met uitzondering van de zeer grote bedrijven.

10.3.5 Non-respons en leeftijd van de ondernemer

Een indeling van de bedrijven naar de leeftijd van de ondernemer, brengt aan het licht dat de leeftijdsklasse sterk samenhangt met de non-respons. Tabel 10.4 geeft een overzicht voor de boekjaren 1985-1987.

De hogere respons bij jongere ondernemers en de lage respons bij rechtspersonen zijn de meest opvallende cijfers in deze tabel. Voor de ondernemers die jonger zijn dan 60 jaar neemt het respons-percentage af met de leeftijd van de ondernemer.

De vrij goede respons bij ondernemers boven de 60 jaar, zou te maken kunnen hebben met de aanwezigheid van opvolgers op een deel van deze bedrijven. Indien de opvolger op het moment van de bedrijfskeuze te kennen geeft de LEI-administratie te willen

voortzetten indien het bedrijf binnen zes jaar wordt overgedragen, is het bedrijf in principe geschikt voor deelname.

In de metelling van 1984 is bij de ondernemers die ouder waren dan 50 jaar, gevraagd naar de aanwezigheid van een opvolger. Voor de 834 bedrijven die betrokken waren bij de keuze voor

Tabel 10.4 Responspercentage en aantal aangezochte bedrijven naar leeftijd van de ondernemer (boekjaren 1985-1987)

Leeftijd van de ondernemer	Aantal bedrijven	Responspercentage
< 30 jaar	27	59
30 - 35 jaar	80	50
35 - 40 jaar	64	47
40 - 45 jaar	111	44
45 - 50 jaar	123	46
50 - 55 jaar	126	40
55 - 60 jaar	121	33
60 - 65 jaar	79	48
65 - 70 jaar	30	43
70 - 75 jaar	10	50
>= 75 jaar	4	25
NV's, BV's	30	17

de boekjaren 1985-1987 en waarvoor de gegevens van de metelling beschikbaar waren, hadden er 358 een ondernemer die ouder was dan 50 jaar. Bij de ondernemers boven de 65 jaar lag de respons op de 20 bedrijven met een opvolger op 30% en bij de 16 bedrijven zonder opvolger op 56%. De aanwezigheid van een opvolger hangt in deze leeftijdscategorie dus niet aantoonbaar positief samen met de respons; er is eerder sprake van een negatief verband. Waarschijnlijk hangt dit samen met het feit dat bedrijven in een periode waarin een bedrijf overgedragen wordt aan een opvolger, in het algemeen terughoudend zullen zijn in het verstrekken van financiële gegevens aan derden. De aanwezigheid van opvolgers op bedrijven met een oudere ondernemer kan gezien het voorgaande moeilijk als reden worden aangevoerd waarom de oudere ondernemers een hoge respons hebben.

Bij de ondernemers tussen de 50 en 65 jaar blijkt de aanwezigheid van een opvolger wel positief samen te hangen met de respons. Een verklaring daarvoor zou kunnen zijn dat de bedrijfs-overdracht op de betreffende bedrijven pas op langere termijn gerealiseerd zal worden, terwijl de aanwezigheid van een opvolger in veel gevallen betekent dat het bedrijf niet binnen zes jaar zal worden opgeheven. Maar ook binnen de groep bedrijven met een opvolger is de respons van ondernemers tussen 50 en 65 jaar niet

hoog. De leeftijd van de ondernemer is dus een factor die - ongeacht de aanwezigheid van een opvolger - van invloed is op de respons.

Verder onderzoek naar het verband tussen de leeftijd van de ondernemer, de aanwezigheid van een opvolger en de respons, is zeker gewenst. In een sterk veranderende landbouw zal dat verband ook in de loop van de tijd aan verandering onderhevig kunnen zijn. In het verleden was er bijvoorbeeld bij ondernemers boven de 60 jaar in het algemeen een lagere respons dan in de boekjaren waarmee hier gerekend is.

De lage respons bij NV's en BV's heeft voornamelijk te maken met de aard van deze bedrijven. Een groot deel van deze bedrijven is niet geschikt voor deelname aan de steekproef (zie paragraaf 10.3.9), onder andere vanwege nevenactiviteiten op deze bedrijven.

10.3.6 Non-respons en oppervlakte

Een indeling van alle bedrijven naar oppervlakte-klasse laat zien dat de respons in vijf oppervlakte-klassen achtereenvolgens 27% (minder dan 5 hectare), 36% (5-10 hectare), 40% (10-20 hectare), 51% (20-50 hectare) en 46% (meer dan 50 hectare) is. Daaruit kunnen echter nauwelijks nieuwe conclusies getrokken worden, omdat de oppervlakte van een bedrijf sterk samenhangt met het bedrijfstype en het aantal sbe. Wordt per basisstratum een indeling in vijf of meer oppervlakte-klassen gemaakt, dan zijn de verschillen in respons tussen de oppervlakte-klassen slechts in enkele basisstrata waar te nemen. Binnen de akkerbouwbasisstrata bijvoorbeeld, waar de oppervlakte fungeert als intensiteits-indicator, is de oppervlakte een factor die samenhangt met de respons. De zeer intensieve akkerbouwbedrijven blijken moeilijk in de steekproef op te nemen te zijn. Van de 33 akkerbouwbedrijven in sbe-klasse 1 die kleiner waren dan 5 hectare en die gedurende 1985-1987 benaderd werden (inclusief de vervangende bedrijven), waren slechts twee bedrijven in staat of bereid deel te nemen 1). Het oppervlakte-criterium in de stratificatie zorgt er binnen dit basisstratum voor dat er ondanks de hoge non-respons toch voldoende zeer intensieve akkerbouwbedrijven in de steekproef worden opgenomen.

Een ander voorbeeld waaruit blijkt dat er een samenhang is tussen oppervlakte en respons, heeft betrekking op het bedrijfstype "gemengd intensief". De oppervlakte geeft hier aan in hoeverre de minimaal 33% intensieve veehouderij op deze bedrijven gecombineerd wordt met grondgebonden activiteiten. In feite gaat

1) De lage respons op de intensieve akkerbouwbedrijven is in paragraaf 10.3.3 ook al naar voren gekomen. Deze groep bedrijven valt namelijk voor het grootste deel in de groep "overige akkerbouwbedrijven".

het ook hier om een verdergaande indeling in bedrijfstypen. De hoeveelheid grond blijkt samen te hangen met de non-respons.

10.3.7 Non-respons en regio

De regio waarin een bedrijf ligt, hangt ook enigszins samen met het responsgedrag. Het Noordelijk Veenweidegebied, het Centraal Kleigebeid, de Friese Wouden en het Centraal Zandgebied laten een meer dan gemiddelde respons zien (meer dan 50%). Het Zuidwestelijk Kleigebied, de Veenkoloniën, het Noordelijk Zandgebied, het Zuidelijk Zandgebied en Zuid-Limburg hebben een respons-percentage dat lager dan 35% ligt. Voor Zuid-Limburg is de respons zelfs lager dan 15%.

Deze verschillen in respons zeggen op zichzelf niet veel over de mate waarin de regio samenhangt met het respons-gedrag. Bedrijven in een bepaalde regio behoren veelal in hoofdzaak tot een bepaald type. De regio's met een lagere respons zijn gebieden waar relatief veel akkerbouw- en intensieve veehouderijbedrijven voorkomen. In theorie is het ook mogelijk dat akkerbouwbedrijven op zichzelf geen lage respons hebben, maar dat ze een lage respons hebben omdat ze in bepaalde regio's liggen waar - om wat voor reden dan ook - de respons voor alle bedrijfstypen laag is.

Om een zuiver beeld te krijgen met betrekking tot de respons-bepalende factoren, is het noodzakelijk om per bedrijfstype-grootteklasse (basisstratum) de verschillen in respons tussen de regio's te bekijken.

In de akkerbouw-basisstrata is er een relatief hoge respons (74%) in het Centraal Kleigebeid. Laag is de respons in de basisstrata in de Veenkoloniën, het Noordelijk Zandgebied, het Zuidelijk Zandgebied en Zuid-Limburg.

In de rundveehouderij-sector worden hoge respons-percentages geregistreerd in het Noordelijk Veenweidegebied, het Centraal Kleigebeid en het Centraal Zandgebied. De respons-percentages zijn daar respectievelijk 71%, 65% en 62%.

Sommige intensieve veehouderij-basisstrata vertonen een lage respons voor bedrijven in gebieden waar deze bedrijven relatief veel voorkomen. Voor de varkensbedrijven zijn de verschillen niet groot. Bij pluimveebedrijven is het verschil aanzienlijk; in het Zuidelijk Zandgebied is de respons 19% en in de overige gebieden 33%. Het verschil in bedrijfs-grootte-structuur is natuurlijk van invloed op de verschillen in gemiddelde respons, maar die verklaren in deze gevallen nauwelijks iets. De gemengd intensieve veehouderijbedrijven hebben een respons van 32% in het Zuidelijk Zandgebied en van 50% in de overige gebieden. Hier zal naast de regio-invloed waarschijnlijk ook het feit meespelen dat de bedrijven in het Zuidelijk Zandgebied een groter aandeel intensieve veehouderij bezitten dan de bedrijven in de overige gebieden. De meer grondgebonden bedrijven hebben in het algemeen een hogere respons.

10.3.8 Non-respons en opleiding van de bedrijfsleider

De schoolopleiding van degene die de dagelijkse leiding op het bedrijf heeft, wordt niet ieder jaar in de metelling vastgelegd. In 1985 is er echter wel naar gevraagd. Voor die bedrijven die zijn aangezocht voor de bedrijfskeuze in de boekjaren 1985 tot en met 1987 en waarvan de metellingsgegevens voor 1985 bekend zijn, is nagegaan of er een samenhang is waar te nemen tussen de respons en de vorm van voortgezet onderwijs die de bedrijfsleider heeft gevolgd. Tabel 10.5 geeft de resultaten.

De conclusie kan getrokken worden dat er een vrij sterke samenhang is tussen de schoolopleiding van de bedrijfsleider en de respons. De respons neemt toe met het niveau van de schoolopleiding. Van de groep bedrijfsleiders met voortgezet middelbaar landbouwonderwijs responderen er tweemaal zoveel als van de groep zonder voortgezet onderwijs. In de akkerbouwsector is de samenhang het meest duidelijk aanwezig. Daar is de respons bij bedrijfsleiders zonder voortgezet landbouwonderwijs 20% en deze neemt toe van 30% bij lager tot 53% bij middelbaar en tot 100% bij hoger landbouwonderwijs.

Het lage responspercentage bij de bedrijfsleiders met een voortgezette opleiding buiten de landbouw wordt gedeeltelijk "veroorzaakt" door het feit dat deze bedrijfsleiders relatief veel voorkomen op bedrijven met intensieve veehouderij.

Tabel 10.5 Responspercentage en aantal aangezochte bedrijven naar opleiding van de bedrijfsleider (boekjaren 1985-1987)

Vorm van voortgezet onderwijs die gevolgd is door de bedrijfsleider	Aantal bedrijven	Responspercentage
Geen voortgezet onderwijs	150	27
Alleen landbouwonderwijs:		
- lager landbouwonderwijs	341	41
- middelbaar landbouwonderwijs	200	58
- hoger landbouwonderwijs	6	50
Combinaties van landbouwonderwijs en overig onderwijs:		
- lager landbouwonderwijs	9	56
- middelbaar landbouwonderwijs	60	50
- hoger landbouwonderwijs	10	20
Alleen niet-landbouwonderwijs:		
- lager onderwijs	44	30
- middelbaar onderwijs	16	31
- hoger onderwijs	7	43

De geconstateerde samenhangen tussen schoolopleiding en respons is voor vrijwel alle leeftijdsklassen aanwezig. Ook per basisstratum is het patroon terug te vinden. Een mogelijke verklaring van de waargenomen verschillen kan liggen in de bekendheid met de LEI-steekproef van bedrijfsleiders met een middelbare of hogere landbouw-opleiding. In het onderwijs-programma van de betreffende opleidingen komt het LEI-boekhoudnet in elk geval ter sprake.

10.3.9 Geschiktheid en bereidheid tot deelname

De hiervoor aangegeven lijnen met betrekking tot de respons van verschillende soorten bedrijven slaan op de totale non-respons. Een uitsplitsing naar non-respons wegens ongeschiktheid en non-respons wegens weigering kan nuttig zijn voor een taxatie van de gevolgen van de non-respons voor de steekproefuitkomsten.

Een bedrijf kan aangemerkt worden als niet in staat tot deelname of ongeschikt voor deelname vanwege ziekte, gezinsomstandigheden, te verwachten opheffing, bijzonder karakter van een bedrijf, en dergelijke (zie paragraaf 8.4). Aan de tweede reden van non-respons - geen bereidheid tot deelname - kunnen verschillende oorzaken ten grondslag liggen. Tijdgebrek, geringe belangstelling, kritiek op het overheidsbeleid en dergelijke oorzaken vallen daar onder. In tabel 10.6 zijn de respons-percentages uit tabel 10.1 - die zijn berekend op basis van alle aangezochte bedrijven - uitgesplitst naar de reden van non-respons. Van alle aangezochte bedrijven wordt eerst het percentage geschikte bedrijven bepaald. Vervolgens wordt het percentage bedrijven berekend dat bereid is mee te doen, bij gebleken geschiktheid.

Van de aangezochte bedrijven blijkt 17% niet in staat tot deelname of ongeschikt voor deelname te zijn. De bereidheid tot deelname is op 45% van de geschikte bedrijven aanwezig. In de

Tabel 10.6-A Bedrijven die geschikt zijn voor deelname aan het boekhoudnet in procenten van het aantal aangezochte bedrijven

Bedrijfstype	Boekjaar							
	81	82	83	84	85	86	87	81-87
Akkerbouw	74	79	76	84	76	76	76	77
Rundvee	83	80	86	92	91	89	86	87
Gemengd akkerbouw + Gemengd rundvee/akkerbouw +	90	79	89	85	68	81	86	83
Gemengd rundvee/int. veeh. (Gemengd) intensieve veeh.	81	83	79	89	87	81	83	83
Alle bedrijven	82	81	81	88	85	83	83	83

Tabel 10.6-B Bedrijven die bereid zijn tot deelname aan het boekhoudnet in procenten van het aantal geschikte bedrijven

Bedrijfstype	Boekjaar							
	81	82	83	84	85	86	87	81-87
Akkerbouw	66	57	47	44	39	52	47	50
Rundvee	60	53	55	62	53	58	54	56
Gemengd akkerbouw + Gemengd rundvee/akkerbouw +	49	48	53	48	50	38	43	47
Gemengd rundvee/int. veeh. (Gemengd) intensieve veeh.	41	32	34	27	36	28	28	32
Alle bedrijven	51	45	45	44	45	46	41	45

Bron: Van Dijk, 1981-1987.

rundveehouderij is zelfs ruim de helft van de bedrijven bereid tot deelname. De geschiktheid voor deelname is in die sector ook het grootst.

De bereidheid tot deelname in de akkerbouwsector blijkt hoger te zijn dan de totale respons-cijfers uit de vorige subparagrafen doen vermoeden. Een kwart van de bedrijven valt al af vanwege ongeschiktheid. De verwevenheid van bedrijven met andere bedrijfstakken is daar een van de oorzaken van.

Voor de bedrijfskeuze van de boekjaren 1986 en 1987 is een onderverdeling naar reden van non-respons gemaakt, waarbij uitge-

Tabel 10.7 Geschiktheids- en bereidheidsrespons en de bijbehorende aantallen bedrijven naar bedrijfstype (boekjaren 1986-1987)

Bedrijfstype	Respons- percentage (geschikt)	Aantal aangezochte bedrijven	Respons- percentage (bereid)	Aantal geschikte bedrijven
Akkerbouw	77	126	52	97
Rundvee	89	223	58	199
Varkens	85	86	34	73
Pluimvee	66	29	37	19
Gemengd akkerbouw	77	13	20	10
Gem. rundvee/akk.	67	15	30	20
Gem. rundvee/int.	86	28	54	24
Gem. intensief	86	35	43	30
Alle bedrijven	83	555	50	462

gaan is van de "eerste keus"-bedrijven en waarbij ook naar andere factoren dan bedrijfstype is gekeken.

Van de 555 aangezochte bedrijven waren 462 bedrijven (83%) geschikt voor deelname. Tabel 10.7 geeft een onderverdeling naar bedrijfstype.

Evenals bij de berekening op basis van alle aangezochte bedrijven (tabel 10.6) is er sprake van een relatief groot aantal ongeschikte bedrijven in de sectoren waar akkerbouw de hoofd- of neven-activiteit is. Wordt een verdeling gemaakt naar NRG-typen, dan zijn het vooral de "overige akkerbouwbedrijven" en de "gewassencombinaties" waar relatief veel bedrijven ongeschikt zijn.

Bij de pluimveebedrijven is een op de drie bedrijven niet in staat tot of ongeschikt voor deelname aan de steekproef. Vooral het relatief grote aantal NV's en BV's in dit bedrijfstype draagt daaraan bij.

Van de 462 geschikte bedrijven is 50% bereid tot deelname. Dit percentage ligt hoger op de akkerbouw- en de rundveebedrijven. De mengvormen van deze bedrijfstypen scoren daarentegen laag. Ook de bedrijven in de intensieve veehouderij-sector hebben een geringe bereidheid tot deelname.

De leeftijd van de ondernemer is een tweede factor die duidelijk verband houdt met de geschiktheid tot deelname.

Tabel 10.8 Geschiktheids- en bereidheidsrespons en de bijbehorende aantallen bedrijven naar leeftijdsklasse van het bedrijfshoofd (boekjaren 1986-1987)

Leeftijd van de ondernemer	Responspercentage (geschikt)	Aantal aangezochte bedrijven	Responspercentage (bereid)	Aantal geschikte bedrijven
< 30 jaar	100	18	61	18
30 - 35 jaar	84	56	49	47
35 - 40 jaar	85	39	58	33
40 - 45 jaar	79	78	55	62
45 - 50 jaar	90	88	44	79
50 - 55 jaar	87	93	47	81
55 - 60 jaar	88	77	35	68
60 - 65 jaar	79	52	63	41
65 - 70 jaar	73	22	63	16
70 - 75 jaar	78	9	71	7
>= 75 jaar	100	3	33	3
NV's, BV's	35	20	26	7
Totaal	83	555	43	462

Tabel 10.8 laat zien dat van de bedrijven met een ondernemer die ouder is dan 60 jaar een kwart afvalt vanwege ongeschiktheid. De bereidheid tot deelname blijkt eveneens samen te hangen met de leeftijd, maar hier zijn het met name de ondernemers in de leeftijdsklasse 55-60 jaar die het laten afweten. De bereidheid neemt af van ongeveer 60% bij jongere ondernemers tot ongeveer 35% bij oudere ondernemers. De ondernemers boven de 60 jaar hebben een opvallend hoge bereidheid tot deelname aan de steekproef. Waarschijnlijk hangt dit gedeeltelijk samen met de lagere geschiktheids-respons. Een duidelijke scheiding tussen de twee redenen om niet mee doen is in de praktijk niet in alle gevallen te maken. De hogere respons hangt in elk geval nauwelijks samen met de aanwezigheid van een opvolger op een bedrijf (zie paragraaf 10.3.5).

Van de bedrijven die kleiner zijn dan 79 sbe - volgens de metelling in het jaar van keuze - blijkt toch nog 60% geschikt te zijn voor deelname. Hetzelfde percentage geldt voor bedrijven die groter zijn dan 750 sbe. De verschillen in geschiktheids-respons tussen de overige sbe-klassen zijn vrij klein. Wordt gekeken naar de bereidheid tot deelname dan komt de lijn uit tabel 10.3 weer sterk naar voren: de bereidheid neemt toe met de bedrijfsgrootte.

Regionale verschillen zijn er in de akkerbouw en in de rundveehouderij. In de Veenkoloniën komen bij deze bedrijfstypen in de boekjaren 1986 en 1987 relatief veel ongeschikte bedrijven voor. In het Centraal Kleigebeid en het Centraal Zandgebied is slechts 5 à 10% van de bedrijven ongeschikt voor deelname aan de steekproef.

De bereidheid om mee te doen is hoog bij de akkerbouwers in het Centraal Kleigebeid (84%) en laag in de Veenkoloniën en het Noordelijk en Zuidelijk Zandgebied (20 à 30%).

In de rundveehouderij is de bereidheid tot deelname het grootst in het Noordelijk Veenweidegebied (87%). In de intensieve veehouderij is de bereidheid tot deelname het laagst in het Zuidelijk Zandgebied.

In het algemeen kan gesteld worden dat groepen bedrijven met een hoog percentage geschikte bedrijven, ook een relatief grote bereidheid tot deelname aan de steekproef vertonen.

10.3.10 Realisatie keuze per 1 mei

Voor sommige oorspronkelijk aangewezen bedrijven worden meer dan tien vervangers benaderd zonder dat dit tot deelname van een bedrijf leidt. In die gevallen waar op 1 mei van het betreffende boekjaar nog geen bedrijf gevonden is dat mee wil doen, wordt de keuze gestaakt. Het percentage van de te kiezen bedrijven dat uiteindelijk ook per 1 mei in administratie genomen kan worden, verschilt niet veel van jaar tot jaar. Meestal kan circa 90% van

de steekproefaanvulling (250 à 300 bedrijven) worden gerealiseerd. De volgende tabel geeft een overzicht per bedrijfstype.

Tabel 10.9 Aantal bedrijven dat per 1 mei deelneemt aan het boekhoudnet in procenten van het aantal bij te kiezen bedrijven naar bedrijfstype en keuze-jaar

Bedrijfstype	Boekjaar							
	81	82	83	84	85	86	87	81-87
Akkerbouw	100	97	86	89	88	89	91	91
Rundvee	96	89	97	95	95	99	98	96
Varkens	91	86	87	92	85	78	76	85
Fluimvee	94	79	100	40	89	75	81	80
Gemengd akkerbouw	86	50	100	75	86	67	113	82
Gemengd rundvee/akkerbouw	80	90	100	100	100	100	90	94
Gemengd rundvee/int. veeh.	100	100	93	94	87	92	100	95
Gemengd intensieve veeh.	84	83	80	65	93	69	95	81
Regionale aanvulling (akk.)	100	85	97	83	93	75	96	90
Alle bedrijven	93	88	92	88	91	90	92	91

Bron: Van Dijk, 1981-1987.

De cijfers in tabel 10.9 sluiten aan bij de gegevens uit paragraaf 10.1. Immers, in bedrijfstypen waar de respons laag is - in de intensieve veehouderij - moeten relatief veel bedrijven benaderd worden voordat het gewenste aantal bedrijven is bereikt dat een toezegging heeft gedaan om deel te gaan nemen aan de steekproef. De kans dat op 1 mei nog steeds naar een vervangend bedrijf voor een bepaald "eerste keus"-bedrijf gezocht wordt, is in die bedrijfstypen dus het grootst.

10.3.11 Non-respons aan het begin van het boekjaar

De resultaten die in de voorgaande subparagrafen gepresenteerd zijn, gaan uit van de respons op het moment dat een bedrijf benaderd wordt. In sommige gevallen blijken bedrijven die in het voorjaar hebben toegezegd mee te werken aan de steekproef, alsnog af te vallen als het per 1 mei op daadwerkelijke medewerking aankomt.

Voor boekjaar 1986 bijvoorbeeld hadden 262 bedrijven een toezegging van deelname gedaan. Bij inventarisatie op 1 mei 1986 bleken er daarvan al 29 afgevallen te zijn. Ook deze verlate non-respondenten bleken zich in meetellingskenmerken te onderscheiden van alle overige respondenten. Bij een gemiddelde afval van 11% per 1 mei van het aantal respondenten, kwamen de bedrijven in de

vier sbe-klassen achtereenvolgens op 18%, 6%, 5% en 12%. De drie leeftijdsklassen die bij de stratificatie worden gehanteerd hadden een afvalpercentage van respectievelijk 9%, 8%, en 21%, terwijl de NV's en BV's op 12% kwamen. De kleinste en de grootste bedrijven, en de bedrijven met de oudste ondernemers dreigen hierdoor nog verder ondervertegenwoordigd te raken in de steekproef.

10.3.12 Conclusies

Bij de werving van "eerste keus"-bedrijven ter verversing en aanvulling van de steekproef, is de respons minder dan 50%. Aan de ene kant is ruim 15% van de bedrijven niet geschikt voor of niet in staat tot deelname. Aan de andere kant is van de resterende bedrijven ongeveer de helft niet bereid tot deelname aan de steekproef.

Een aantal bedrijfskenmerken hangt sterk samen met de geneigdheid tot responderen. Intensieve veehouderij-bedrijven hebben een lage respons, rundvee-bedrijven scoren relatief hoog en de akkerbouwbedrijven zitten daar tussen in. Grotere bedrijven doen in het algemeen sneller mee met de steekproef dan kleinere bedrijven.

De leeftijd van de ondernemer hangt ook vrij sterk samen met de non-respons. Hoe jonger de ondernemer, des te groter de responsgeneigdheid. Rechtspersonen hebben een zeer hoge non-respons. De opleiding van de ondernemer speelt een vrij grote rol bij de respons. Middelbaar landbouwonderwijs staat garant voor een grote kans op deelname aan de steekproef. Bedrijfsleiders zonder voortgezet onderwijs hebben een laag respons-percentage.

Regionale verschillen in respons hangen gedeeltelijk samen met verschillen in respons tussen de bedrijfstypen, maar de regio heeft ook een zelfstandige samenhang met de respons, los van de samenhangen van de respons met de andere genoemde factoren. Dat laatste geldt voor alle factoren in meerdere of mindere mate. De samenhang tussen respons en opleiding bijvoorbeeld, gaat ook op binnen een bedrijfstype of grootte- of leeftijds-klasse.

10.4 Non-respons bij de uitwerking van boekhoudingen (BUL)

10.4.1 Inleiding

Ook indien de non-respons bij de bedrijfskeuze niet of nauwelijks van invloed zou zijn op de uiteindelijke bedrijfsuitkomsten, of indien voor die invloed gecorrigeerd wordt, en als bovendien de bedrijfskeuze volledig geslaagd zou zijn voor 1 mei, is het nog mogelijk dat er systematische fouten in de uitkomsten gaan optreden. De oorzaak daarvan is dat gedurende het boekjaar bedrijven uit de steekproef verdwijnen, bijvoorbeeld door opheffing of door nalatigheid bij het insturen van gegevens. Sommige

bedrijven blijven wel in de steekproef maar leveren hun gegevens te laat in.

10.4.2 Samenhang met stratificatie-variabelen

Voor de boekjaren 1982-1986 is nagegaan in hoeverre er in het niet - of te laat - leveren van een uitgewerkte boekhouding aan het eind van het boekjaar, een zekere systematiek aanwezig is. Van de circa 5.700 steekproefbedrijven die per 1 mei van deze boekjaren voldeden aan de populatie-eisen, kon 9 maanden na afloop van het boekjaar - dat wil zeggen op het moment dat de BUL geproduceerd werd - voor ruim 9% geen uitgewerkte boekhouding worden geleverd. Daarbij is het opmerkelijk dat dit voor bijna 14% van de rundveebedrijven het geval was, voor bijna 8% van de pluimveebedrijven, en voor ongeveer 5% van de overige bedrijven.

Tussen de vier sbe-classes die worden onderscheiden bij de stratificatie, zijn ook verschillen te signaleren. Non-responspercentages van respectievelijk 5%, 8%, 13% en 11% in de uitwerkingsfase wijzen erop dat het moeilijker is om voor grote bedrijven een uitgewerkte boekhouding te realiseren dan voor kleine bedrijven. De hoogste non-responspercentages gelden voor rundveebedrijven in de sbe-classes 3 en 4 (respectievelijk 20% en 16%) en voor pluimveebedrijven in sbe-klasse 4 (17%). De laagste sbe-classes in de intensieve-veehouderijsector laten een beeld zien dat iets afwijkt ten opzichte van de overige sectoren; hier worden van relatief veel bedrijven (circa 10%) de boekhouding niet uitgewerkt.

Tussen de leeftijdsclasses zijn geen grote verschillen in respons te constateren. Van de ondernemers jonger dan 35 jaar levert 11% geen uitgewerkte boekhouding, voor de boeren ouder dan 60 jaar is dat 9%, en voor de tussengroep 8%. De categorie NV's en BV's heeft echter een duidelijk afwijkende non-respons van 24% - op een totaal van 78 van dergelijke bedrijven in de 5 betreffende boekjaren. In vrijwel alle 32 basisstrata zijn deze tendensen terug te vinden.

Verschillen tussen oppervlakteclasses volgen in grote lijnen de verschillen tussen sbe-classes. Voor de classes 0-5 ha, 5-10 ha, 10-20 ha, 20-50 ha, en meer dan 50 ha, bedragen de non-responspercentages respectievelijk 7%, 3%, 4%, 13% en 8%. Binnen sommige basisstrata zijn de verschillen tussen de oppervlakteclasses ook aanwezig. Binnen de akkerbouwsector hebben de intensieve bedrijven (meer dan 12 sbe/ha) een hogere non-respons. Dat geldt ook voor rundveebedrijven die extensief zijn (minder dan 6 sbe/ha). De zeer intensieve rundveebedrijven (meer dan 21 sbe/ha) hebben een lage non-respons.

In de intensieve veehouderij zijn het vooral de bedrijven met een kleinere oppervlakte die een hoger non-responspercentage hebben binnen een bepaalde sbe-klasse.

Regionale verschillen zijn ook vrij opvallend. Het Noordelijk Veenweidegebied (19% non-respons), het Westelijk weidegebied (18%), en Zuid-Limburg (23%) leveren relatief gezien weinig uitgewerkte boekhoudingen. Met name de rundveebedrijven - die in deze regio's een groot deel van het totaal aantal bedrijven omvatten - dragen in deze regio's aan deze resultaten bij 1). In de pluimveehouderij is het opvallend dat 99% van de bedrijven in het Zuidelijk Zandgebied een uitgewerkte boekhouding levert, tegen 84% in de overige regio's. Bij gemengde akkerbouwbedrijven doet zich een soortgelijk iets voor, maar dan voor andere regio's. In het Zuidelijk Zandgebied en in Zuid-Limburg is de respons 85% en in de overige regio's 99%.

10.4.3 De noodzaak tot stratificatie

Gegeven het cijfermateriaal kan geconcludeerd worden dat het met het oog op het voorkomen van systematische fouten, gewenst is om achteraf voor een aantal factoren die met de non-respons in de uitwerkingsfase en met de steekproefuitkomsten samenhangen, te corrigeren. Vanwege de verschillen in steekproefpercentages tussen de 32 basisstrata bij de bedrijfskeuze, is al een poststratificatie in 32 basisstrata noodzakelijk (zie hoofdstuk 9). Voor de sub-strata die binnen de 32 basisstrata zijn gevormd bij de keuze, zijn ongeveer dezelfde steekproefpercentages toegepast, zodat vanuit het gezichtspunt van de *historische trekkingskansen* een poststratificatie binnen de 32 post-basisstrata niet strikt noodzakelijk is 2). Maar omdat de *uitwerkingskansen* binnen een basisstratum wel enigszins verschillen tussen oppervlakte-, leeftijds-, en regionale klassen, is een verdergaande poststratificatie met behulp van deze klassen gewenst.

10.5 Non-respons bij de uitwerking van boekhoudingen (FIP)

10.5.1 Inleiding

In de vorige paragraaf is naar voren gekomen dat een deel van de bedrijven die medewerking toezeggen aan het boekhoudnet, geen bruikbare bedrijfseconomische boekhouding levert. Daarnaast zijn er bedrijven die wel een bedrijfseconomische boekhouding le-

-
- 1) Ook binnen de groep van rundveebedrijven hebben deze regio's een hogere non-respons. Het bedrijfstype-effect en het regionale effect versterken elkaar dus ten aanzien van de non-respons. Van de in administratie zijnde rundveebedrijven in Zuid-Limburg heeft minder dan twee derde deel een uitgewerkte boekhouding geleverd.
 - 2) Behalve voor de akkerbouwstrata waar regionale aanvullingen (paragraaf 7.4) zorgen voor verschillende trekkingskansen.

veren, maar geen financieringsboekhouding (willen) verstrekken. In deze paragraaf zal aandacht besteed worden aan de factoren die met deze vorm van non-respons samenhangen.

Van de ruim 1.000 bedrijven die gebruikt worden voor het samenstellen van de BUL, leveren er gemiddeld ruim 900 ook een financieringsboekhouding. Dit laatste aantal kan in principe gebruikt worden voor het samenstellen van de FIP. In de praktijk wordt echter een aantal van deze 900 bedrijven uitgesloten van de berekeningen voor de FIP, omdat ze niet een volledige boekhouding kunnen leveren. Hierop wordt in paragraaf 10.5.3 ingegaan.

10.5.2 Bereidheid tot het leveren van een financieringsboekhouding

Van de 3.033 bedrijven die in de boekjaren 1984-1986 een uitgewerkte bedrijfseconomische boekhouding aan het LEI verstrekten, leverden er 2.736 (91%) ook een betrouwbare financieringsboekhouding. Op akkerbouwbedrijven lag dit uitwerkingspercentage op 85%; op akkerbouwbedrijven met minder dan 4 sbe/hectare op 74% en op bedrijven met meer dan 9 sbe/hectare op 92%. Bij de rundveebedrijven werd van de bedrijven met minder dan 6 sbe/hectare 97% uitgewerkt, 95% van de bedrijven met meer dan 15 sbe/hectare, en van de overige bedrijven 92%. Van de overige bedrijfstypen scoorden vooral de gemengde bedrijven met rundveehouderij en intensieve veehouderij hoog (95%).

Voor de vier sbe-klassen gold dat respectievelijk 94%, 92%, 91% en 84% van de bedrijven met een BUL-wegingsfactor, ook een financieringsboekhouding leverde. Per bedrijfstype levert de hoogste sbe-klasse ook in vrijwel alle gevallen de laagste FIP-uitwerkingsrespons.

Wordt naar leeftijdsklasse gekeken, dan is de respons iets minder dan 90% voor ondernemers van 50-60 jaar, en rond 80% voor ondernemers ouder dan 60 jaar. Wellicht speelt bedrijfsoverdracht en de financiële afwikkeling daarvan een rol bij de lagere respons van oudere ondernemers. Voor NV's en BV's werd in de jaren 1984-1986 een FIP-respons van 67% genoteerd.

Regionale verschillen zijn duidelijk te constateren bij de FIP-respons. In de akkerbouw leveren relatief weinig bedrijven een FIP-boekhouding in het Noordelijk Zeekleigebied (74%), het Zuidelijk Zandgebied (72%) en Zuid-Limburg (33%). De bedrijven in het Centraal Kleigebied (92%) en in het Noordelijk Zandgebied (94%) hebben minder moeite met het leveren van een financieringsboekhouding. In de rundveehouderij heeft het Centraal Kleigebied eveneens een hoge respons (97%), evenals de Veenkoloniën (100%), het Noordelijk Veenweidegebied (95%) en het Zuidelijk Zandgebied (97%). In het Noordelijk Zeekleigebied (85%), het Rivierkleigebied (85%) en in Zuid-Limburg (59%) kost het meer moeite om financieringsboekhoudingen geleverd te krijgen.

10.5.3 Uitsluiting van bedrijven met een onvolledige boekhouding

Van de 2.736 bedrijven die gedurende de boekjaren 1984-1986 een uitgewerkte financieringsboekhouding leverden en die een FIP-wegingsfactor kregen, zijn er 199 niet in de resultaten van de FIP opgenomen, omdat ze geen *volledige* financieringsboekhouding konden leveren. Het ging hier vooral om NV's, BV's en maatschappijen (Aukema en Overgauw, 1988:9). Een nadere analyse leert dat deze uitsluiting nogal specifieke bedrijven betreft.

Van het totale aantal bedrijven werd bijna 93% opgenomen in de FIP. Voor de akkerbouwbedrijven was dit 88%, voor de gemengde rundvee/akkerbouwbedrijven 89%, en voor de gemengde rundvee/intensieve veehouderij-bedrijven 99%.

Van de bedrijven die in de metelling opgeven dat ze rechtspersoonlijkheid bezaten (20 bedrijven, waarvan 10 varkensbedrijven en 6 akkerbouwbedrijven), werd toch nog 30% in de FIP opgenomen. Bij de overige bedrijven werden relatief veel bedrijven met een jong bedrijfshoofd opgenomen. In de leeftijdscategorie "jonger dan 35 jaar" werd 97% van de bedrijven in de FIP opgenomen.

Een indeling van de bedrijven naar sbe-klasse geeft te zien dat met name veel grotere bedrijven van de FIP-berekeningen worden uitgesloten. De 4 sbe-klassen die bij de bedrijfskeuze worden gebruikt, laten FIP-opname-percentages zien van respectievelijk 100%, 97%, 94% en 78%. Het merendeel van de NV's en BV's is in de laatste grootte-klasse te vinden. Van de 20 akkerbouwbedrijven die groter zijn dan 650 sbe worden slechts 3 bedrijven in de FIP opgenomen. Bij de rundveebedrijven in deze klasse ligt dit percentage rond de 50%.

Regionale verschillen zijn vooral aanwezig bij akkerbouw- en rundvee-bedrijven. Bij de akkerbouwbedrijven valt de lage opname in het Noordelijk Zeekleigebied (71%) op; ook al wordt rekening gehouden met het grote aantal grote bedrijven in dit gebied, dan nog is het cijfer duidelijk lager dan in andere gebieden. In dit gebied was de FIP-respons ook al laag, zodat ten opzichte van het aantal bedrijven in de BUL een flinke reductie optreedt. In het Noordelijk Zandgebied worden alle 102 bedrijven die daarvoor in aanmerking komen opgenomen in de FIP.

In de rundveehouderij-sector zijn de verschillen tussen regio's vrij opmerkelijk. Weliswaar speelt ook daar dat in regio's waar de bedrijven relatief gezien groot zijn, een groter deel van de bedrijven met een FIP-wegingsfactor niet in de FIP wordt opgenomen. Maar los van dit grootte-aspect is er ook een regionaal aspect in de verschillen aanwezig. In het Noordelijk Zeekleigebied doet 84% van de bedrijven mee in de FIP, in het Noordelijk Veenweidegebied 88%. In sommige andere gebieden is dit percentage 99 of 100. Dat betreft het Centraal Kleigebied, het Zuidwestelijk Zeekleigebied, het Centraal Zandgebied en het Oostelijk Zandgebied.

11. Representativiteit

11.1 Inleiding

In dit hoofdstuk wordt nagegaan of de steekproefuitkomsten al dan niet systematisch afwijken van de werkelijke uitkomsten in de populatie. Representatieve steekproefuitkomsten zijn immers niet gegarandeerd door een aselechte steekproeftrekking 1). Door systematische non-respons bij de werving van bedrijven of bij de uitwerking van boekhoudingen - hoofdstuk 10 gaf aan dat die systematiek in sterke mate aanwezig is - kunnen de steekproefuitkomsten systematisch gaan afwijken van de uitkomsten in de populatie. Dergelijke afwijkingen kunnen echter voorkomen worden indien bij de steekproefopzet en bij het berekenen van steekproefuitkomsten op een juiste wijze rekening wordt gehouden met de non-respons. In de hoofdstukken 5 tot en met 9 van dit onderzoekverslag is aangegeven hoe dat gebeurt. In dit hoofdstuk blijkt dat de gekozen steekproefmethode - een gestratificeerde steekproef - in combinatie met de gekozen methode voor het berekenen van steekproefuitkomsten - stratifikatie achteraf -, in de periode vanaf 1975 hebben geleid tot steekproefuitkomsten die in veel opzichten representatief zijn.

In dit hoofdstuk wordt op twee manieren nagegaan of de steekproefuitkomsten al dan niet systematisch afwijken van de werkelijke uitkomsten op de landbouwbedrijven in de populatie.

In de eerste plaats wordt nagegaan in hoeverre de meetlingskenmerken van de steekproefbedrijven overeenkomen met de meetlingskenmerken van alle bedrijven die aan de populatie-eisen voldoen. Onderzocht wordt of de bedrijven in beide groepen qua grootte overeenkomen, of de gemiddelde leeftijd van het bedrijfshoofd ongeveer gelijk is, of het aandeel van de verschillende bedrijfs-activiteiten niet te sterk afwijkt, enz.

In de tweede plaats worden de steekproefuitkomsten geconfronteerd met gegevens die uit andere bronnen beschikbaar zijn. Daarbij doet zich het probleem voor dat externe gegevens vaak betrekking hebben op andere - meestal grotere - populaties dan de steekproefgegevens. Door een schatting te maken voor het in de steekproefpopulatie ontbrekende deel van de grotere populatie waarvoor externe gegevens beschikbaar zijn, kan een vergelijking worden gemaakt. Kilogram-opbrengsten per hectare gewas, melkgift per koe, en dergelijke, zijn grootheden waarvoor extern vergelijkingsmateriaal voorhanden is. Paragraaf 11.4 gaat in op deze benadering.

1) Voor een omschrijving van het begrip "representativiteit" zie paragraaf 4.4.2.

In paragraaf 11.5 wordt nagegaan wat de gevolgen voor de representativiteit zouden zijn van een eenvoudiger systeem van post-stratificatie. Paragraaf 11.6 kijkt concluderend terug op dit hoofdstuk.

11.2 Vergelijking steekproef met metelling (BUL)

11.2.1 Inleiding

Voor de totale steekproefpopulatie van circa 70.000 bedrijven kan worden nagegaan wat de totalen en gemiddelden zijn van grootheden die beschikbaar komen via de metelling. Met behulp van de ruim 1.000 steekproefbedrijven kan - via de wegingsfactoren die aan de bedrijven zijn toegekend - eveneens bepaald worden wat de totalen en gemiddelden van *metellingsgegevens* voor de gehele populatie zijn. Voor de jaren 1982 t/m 1986 zijn de betreffende cijfers naast elkaar gelegd. Voor enkele gegevens is ook gekeken naar de periode 1975-1981.

Op deze manier kan worden bepaald of er sprake is van "sample-population consistency" (Brackstone and Rao, 1979), dat wil zeggen of de steekproef goede schattingen geeft voor variabelen waarvoor de waarde van alle bedrijven bekend is.

In de eerste jaren na 1975 werden de wegingsfactoren nog niet berekend op basis van de exakte metellingsgegevens, maar uit een grote steekproef uit de metelling. Dat vergrootte de kans op afwijkingen tussen steekproef en populatie. Daarom worden deze jaren hier maar zijdelings in beschouwing genomen. Voor de FIP zijn pas sinds 1982 afzonderlijke wegingsfactoren beschikbaar. Met het oog op de vergelijkbaarheid tussen BUL- en FIP-cijfers (paragraaf 11.3) wordt in deze paragraaf dan ook vooral gekeken naar de boekjaren 1982-1986.

11.2.2 Kleine afwijkingen

Voor variabelen die een belangrijke rol spelen in het stratificatie- en bedrijfskeuzeproces, zijn er - zoals verwacht mag worden - slechts kleine afwijkingen tussen steekproef- en populatie-berekeningen. Het gemiddeld aantal sbe is in de steekproef ruim 1% groter dan in de populatie. De gemiddelde oppervlakte cultuurgrond wijkt meestal minder dan 1% af, evenals het aantal bedrijven met cultuurgrond. De gemiddelde leeftijd van het bedrijfshoofd volgens de cijfers van de steekproefbedrijven, ligt minder dan een half jaar onder de gemiddelde leeftijd die resulteert uit de cijfers van alle metellingsbedrijven.

Ook voor de meeste variabelen die betrekking hebben op het aantal arbeidskrachten, het bouwplan en de samenstelling van de veestapel, is er in grote lijnen overeenstemming tussen metellingscijfers van steekproefbedrijven en van de overige bedrijven

in de populatie. Voor een aantal specifieke variabelen is er echter een opmerkelijke verschil te constateren. Tabel 11.1 geeft daarover nadere informatie. In deze tabel wordt voor een groot deel van de meetingsvariabelen de gemiddelde waarde voor alle bedrijven gegeven en deze wordt vergeleken met de gemiddelde waarde op steekproefbedrijven. Daarbij wordt aangegeven in hoeverre een afwijking tussen beide waarden veroorzaakt wordt door een afwijkend aantal bedrijven met een waarde groter dan 0 voor de betreffende variabele, en in hoeverre door een afwijkende gemiddelde waarde op bedrijven waar de betreffende variabele een positieve waarde heeft.

11.2.3 Arbeidskrachten

Een eerste groep van variabelen in de meting heeft betrekking op het aantal arbeidskrachten. Het aantal mannelijke arbeidskrachten wordt vrij goed benaderd door de steekproef, waarbij echter het aantal vreemde arbeidskrachten op steekproefbedrijven circa 25% lager ligt dan op alle meetingsbedrijven. Het aantal vrouwelijke arbeidskrachten wordt over de gehele linie te laag ingeschat door de steekproef. Het aantal meewerkende echtgenotes is in de steekproef ongeveer 10% lager dan in de gehele meting. Vrouwelijke bedrijfshoofden zijn er bijna twee maal zoveel dan de steekproef wil doen geloven, terwijl ook meewerkende dochters op steekproefbedrijven minder voorkomen dan gemiddeld in de populatie. Voor meewerkende zoons geldt overigens hetzelfde. Vaste vreemde vrouwelijke arbeidskrachten zijn sterk ondervertegenwoordigd in het boekhoudnet. Waarschijnlijk is het uitsluiten van bedrijven met veel nevenactiviteiten daar de oorzaak van.

De aanwezigheid van vreemde, vrouwelijke of andere gezinsarbeidskrachten zijn dus factoren die blijkbaar samenhangen met de non-respons. Misschien niet zozeer direkt als wel via bijvoorbeeld de opvolgings- en gezinssituatie, of de aard van de bedrijven.

11.2.4 Oppervlakten

Grootheden die betrekking hebben op de bedrijfsoppervlakte en de verdeling daarvan over verschillende bestemmingen, vertonen geen grote verschillen tussen populatie en steekproef met betrekking tot de meetingsgegevens. Opvallend is echter wel dat de gemiddelde hoeveelheid braakland op steekproefbedrijven 30 à 40% hoger ligt dan op de rest van de bedrijven, terwijl de oppervlakte "tuin voor eigen gebruik" ruim 20% hoger is. Het betreft hier slechts 2 respectievelijk 1 are gemiddeld per bedrijf over de populatie, zodat hier geen problemen uit voortvloeien. Het kan een aanwijzing zijn dat steekproefbedrijven hun meetingsformulieren nauwkeuriger invullen dan de overige bedrijven. Dat zou ook kunnen blijken uit de oppervlakte overige gronden die systema-

tisch 25% wordt onderschat. In de periode 1975-1981 was deze onderschatting nog groter. Daarbij speelt waarschijnlijk ook een rol dat bedrijven met veel "overige gronden" vaker ongeschikt zullen zijn voor deelname aan het boekhoudnet dan bedrijven met minder "overige gronden". Het gaat hier gemiddeld om een oppervlakte van bijna één hectare per bedrijf in de populatie. Ook de oppervlaktes grienden, riet en biezen, waarnaar gevraagd is in de metellingen van 1975 en 1979, en de hoeveelheid bos worden sterk onderschat, maar ook hierbij moet aangetekend worden dat het om kleine oppervlaktes gaat en dat verwevenheid met neventakken en de daaruit voortvloeiende ongeschiktheid tot deelname (paragraaf 8.4) de oorzaak van deze vertekening kan zijn.

11.2.5 Intensieve veehouderij

Metallingsgrootheden die inzicht geven in de representatie van de intensieve veehouderij in de steekproef, zijn de aantallen dieren en de hokcapaciteiten voor verschillende diersoorten.

Het aantal bedrijven met hokcapaciteit voor mestkalveren wordt met circa 22% onderschat door de steekproef, terwijl het aantal bedrijven met mestkalveren met 16% onderschat wordt. Het aantal mestkalveren per bedrijf op de betreffende bedrijven wordt ook onderschat.

Het totale aantal mestvarkens wordt vrij goed weergegeven door de steekproef, maar dit is het resultaat van een overschatting van het aantal bedrijven met mestvarkens met 8% en een onderschatting van het gemiddeld aantal aanwezige mestvarkens op die bedrijven met 10%.

De hokcapaciteit voor slachtkuikens wordt meestal te laag ingeschat door de steekproef. Zowel het aantal bedrijven met hokcapaciteit voor slachtkuikens als de gemiddelde hokcapaciteit dragen daaraan bij. De hokcapaciteit voor moederdieren van slachtrassen daarentegen, wordt met 50% overschat. Het aantal moederdieren wordt ook overschat, en dan vooral voor moederdieren jonger dan 6 maanden. De overschatting wordt veroorzaakt door een te groot aantal bedrijven met moederdieren in de steekproef.

Voor leghennen ouder dan 5 maanden zit de steekproef vrij goed met schattingen voor de hokcapaciteit. Het totale aantal leghennen wordt in de meeste jaren te hoog ingeschat, meestal door een te hoge schatting van het aantal bedrijven waar deze dieren voorkomen. Het aantal leghennen dat ouder is dan 20 maanden wordt onderschat.

Het aantal kalkoenen wordt voortdurend overschat met een factor 2 à 3, vooral doordat het aantal bedrijven met kalkoenen wordt overschat.

Voor de periode 1975-1981 zijn dezelfde tendensen te zien in de intensieve veehouderij.

Concluderend kan gesteld worden dat er vrij grote verschillen tussen de diersoorten voorkomen voor wat betreft hun weergave

Tabel 11.1 *Vergelijking meetellingsgegevens van steekproef- en meetellingsbedrijven voor de gemiddelde nationale boerderij (gemiddelde van boekjaren 1982-1986)*

Variabele	Gemiddelde waarde in populatie	Index voor boekhoudnetbedrijven		
		gemiddelde waarde	bedrijven met waarde > 0	gem. waarde
Sbe totaal	209,9	101	100	101
Bss totaal	58,8	102	100	102
Oppervlakte (aren):				
- cultuurgrond vorig jr.	2206	102	102	100
- cultuurgrond	2294	101	101	100
- bouwland	854	101	102	99
- blijv. grasland	1342	100	101	100
- tijd. grasland	45	87	101	86
- tuinbouw open gr.	50	106	109	97
- tuinbouw glas	0	1	4	13
- braakland	2	129	109	119
- eigendom	1) 1330	98	102	97
- pacht	1) 881	103	103	100
- erfpacht	1) 71	100	98	102
- pacht van ouders	3) 122	138	142	97
- andere expl. vormen	1) 15	94	110	86
- totaal kadastraal	2360	101	101	100
- niet in gebruik	1) 1	110	106	104
- bos	1) 23	28	106	27
- overige gronden	93	74	101	73
- tuin eigen gebruik	1) 1	122	117	103
- totaal bedrijf	2467	99	100	99
Hokcapaciteiten:				
- mestkalveren	8	71	78	92
- mestvarkens	87	99	109	90
- slachtkuikens	2) 548	92	97	95
- moederdieren	2) 74	152	157	97
- leghennen > 5 mnd.	2) 480	104	100	104
Aantal melkkoeien	33	103	101	102
Sbe				
- rundvee fokkerij	104,1	103	100	102
- rundvee mesterij	2,3	81	96	85

Tabel 11.1 (1e vervolg)

Variabele	Gemiddelde waarde in populatie	Index voor boekhoudnetbe- drijven		
		gemid- delde waarde	bedrijven met waarde > 0	
			aantal	gem. waarde
- mestkalveren	2,4	69	84	83
- schapen	0,9	81	105	77
- geiten	6) 0,1	9	86	10
- paarden	1,4	52	83	63
- grasland	15,3	100	101	99
- mestvarkens	10,8	98	108	90
- fokvarkens	18,4	105	106	100
- slachtkuikens	1,4	89	86	103
- leghennen (incl. moederd.)	6,0	113	112	101
- moederdieren < 6 mnd.	0,3	142	192	74
- moederdieren > 6 mnd.	1,1	123	134	92
- leghennen < 5 mnd.	0,7	126	139	91
- leghennen 5 -20 mnd.	3,4	109	99	111
- leghennen >= 20 mnd.	0,6	84	79	107
- eenden	0,1	0	0	0
- kalkoenen	0,1	275	308	89
- peulvruchten	0,6	114	121	94
- landbouwzaden	0,7	116	112	103
- handelsgewassen	0,4	98	111	88
- aardappelen totaal	15,2	106	106	100
- pootaardappelen	5,2	110	111	99
- fabrieksaardappelen	3,3	110	118	93
- cons. aardappelen	6,7	101	102	99
- granen	6,0	102	105	97
- suikerbieten	11,7	106	105	100
- overige akkerbouw	8,6	96	101	94
- voedergewassen	6,5	90	99	91
- tuinakkergewassen	2,4	87	106	82
- tuinbouwgewassen	3,1	86	113	77
Arbeidskrachten:				
- mannelijk	1,38	98	101	97
- bedrijfshoofd	0,83	102	102	100
- vreemde arb. kr.	0,12	75	94	80
- meewerkende echtgen.	0,00	20	20	100
- vrouwelijk	0,44	89	90	99
- bedrijfshoofd	0,01	62	62	100
- vreemde arb. kr.	0,01	38	47	81
- meewerkende echtgen.	0,41	91	91	100

Tabel 11.1 (2e vervolg)

Variabele	Gemiddelde waarde in populatie	Index voor boekhoudnetbe- drijven		bedrijven met waarde > 0	
		gemid- delde waarde		aantal	gem. waarde
Aantal kavels	3) 4,01	103		101	102
Opp. huiskavel (aren)	4) 914	105		103	102
% bedr. met ligboxenstal	4) 31	105		105	100
Beregenbare opp.(1983)	3) 590	116		112	103
Te beregenen opp.(1985)	5) 413	105		101	104
- opp. gras	5) 323	108		103	104
- opp. bouwland	5) 77	98		96	102
- opp. tuinbouw	5) 13	77		93	83
Leeftijd bedrijfshoofd	48,55	99		100	99
Geboortjaar opvolger (indien aanwezig)	(1959,1) 4)	106		106	100
% bedrijven met					
- rechtspersoonlijkheid	1	41		41	100
- hoofdberoep landbouwer	97	98		98	100
- machtiging landbouwschap	99	101		100	100

- 1) Alleen gevraagd in metelling 1983 en 1985.
- 2) Niet gevraagd in metelling 1985.
- 3) Alleen gevraagd in metelling 1983.
- 4) Alleen gevraagd in metelling 1984.
- 5) Alleen gevraagd in metelling 1985.
- 6) Alleen gevraagd in metelling 1984-1986.

in de steekproef. Dit hangt vrijwel zeker samen met de respons bij de bedrijfskeuze. Bedrijven die al veel bedrijfseconomische en technische informatie over hun bedrijf ontvangen van hun mengvoederfabrikant of via de Technisch-Economische Administratie (TEA) van de Consulentschappen zijn minder snel geneigd om deelname aan de LEI-steekproef toe te zeggen dan bedrijven die over minder bedrijfsinformatie beschikken. Bedrijven met kalkoenen of met moederdieren van slachtrassen zijn in het algemeen minder met informatie toegerust dan bedrijven met slachtkuikens. Hetzelfde geldt voor kleinere bedrijven met intensieve veehouderij. Verder speelt een rol dat in sommige intensieve veehouderijtakken de dieren op contract-basis worden gehouden, wat de behoefte aan informatie bij de veehouder niet vergroot.

Een factor die ook van invloed is op de weergave van bepaalde categorieën bedrijven in de intensieve veehouderij is de leeg-

stand van de hokken op 1 mei op een deel van de bedrijven. Deze bedrijven vallen door dit verschijnsel buiten de populatie en de betreffende steekproefbedrijven worden daarom niet meegenomen in de berekening van steekproefuitkomsten. Voor de slachtkuikenbedrijven bijvoorbeeld resulteert dit in een minder sterke vertegenwoordiging in de steekproefbedrijven in de pluimveebasisstrata. Zouden de slachtkuikenbedrijven in een afzonderlijk stratum zijn opgenomen, dan zou er geen sprake zijn van ondervertegenwoordiging door in en uit de populatie "springende" bedrijven.

11.2.6 Overige dieren

Systematisch ondervertegenwoordigd in de steekproef zijn het rundvee voor de mesterij, de schapen, de geiten en de paarden. De steekproef telt circa 20% minder mestvee, 20% minder schapen, 90% minder geiten en 50% minder paarden dan de mettellingspopulatie bij integrale telling. Vooral het gemiddeld aantal dieren per bedrijf waar ze voorkomen, wordt te laag ingeschat. Bij paarden is dat bijvoorbeeld te verklaren uit het feit dat manege-bedrijven niet in de steekproef worden opgenomen (zie paragraaf 8.4). Bij de geiten ontbreken de grotere bedrijven die recent met de geitenhouderij van start zijn gegaan en daarom (nog) niet in de steekproef zijn opgenomen.

Een meer algemene oorzaak van de ondervertegenwoordiging van de genoemde diersoorten ligt in de bedrijfstypering bij de bedrijfskeuze. Daarbij wordt niet expliciet onderscheid gemaakt tussen verschillende soorten rundveehouderijbedrijven. In paragraaf 10.3.3 bleek dat de gespecialiseerde melkvee-bedrijven een hoger respons-percentages hebben dan de andere rundveehouderijbedrijven. Dat verklaart ook waarom het aantal melkkoeien met ongeveer 3% overschat wordt door de steekproefbedrijven.

11.2.7 Gewassen

Bij de meeste gewassen geeft de steekproef een overschatting van het totale areaal met enkele procenten. Het graanareaal vertoont nauwelijks afwijkingen groter dan 2%, waarbij wel opgemerkt moet worden dat een teveel aan bedrijven met graan in de steekproef gecompenseerd wordt door een gemiddeld iets kleinere oppervlakte graan op deze bedrijven. Het totale areaal aardappelen wordt met circa 6% overschat, evenveel als bij suikerbieten het geval is. Dat heeft voornamelijk te maken met een overschatting van het aantal bedrijven waarop deze gewassen voorkomen. Opvallend is dat de gewassen die op wat minder grote schaal in Nederland voorkomen, sterk vertegenwoordigd zijn in de steekproef. De totale oppervlakte peulvruchten wordt met 14% overschat, de oppervlakte landbouwzaden met 16%, de oppervlakte poot- en fabriksaardappelen beide met 10%. Daar staat tegenover dat de oppervlakte voedergewassen met 10% wordt onderschat, en de oppervlakte tuinakkergewassen met 13%. Het areaal tuinbouwgewassen

(exclusief tuinakkergewassen) dat binnen de populatie voorkomt is ook groter dan de steekproef doet vermoeden, hoewel het aantal bedrijven met tuinbouwgewassen overschat wordt.

11.2.8 Overige meetellingsgegevens

Naast de gebruikelijke gegevens over arbeidskrachten, gewassen, diersoorten en grondgebruik, zijn er in de meeste jaren enige aanvullende gegevens met betrekking tot specifieke onderwerpen, in de metelling opgenomen.

In 1984 werd de aanwezigheid van een ligboxenstal vastgelegd. Volgens de steekproef zijn er 22.826 bedrijven met een of meerdere ligboxenstallen, terwijl telling van de gehele populatie uitwijst dat er 21.807 zijn. De steekproef overschat het aantal bedrijven dus met 5%. Wordt rekening gehouden met de overschatting van het aantal bedrijven met melkkoeien dan is de overschatting van het aantal bedrijven met een ligboxenstal vrij klein.

In 1983 en in 1985 is gevraagd naar de beregenbare oppervlakte. De steekproef schat het aantal bedrijven met beregeningsmogelijkheden in 1983 op 23.815 tegen een populatieaantal van 21.251. Deze overschatting van 12% levert samen met een lichte overschatting van de gemiddeld beregenbare oppervlakte per bedrijf, een totale beregenbare oppervlakte die 16% hoger is dan in de populatie. Volgens de metelling in 1985 werd er 288.246 hectare berekend, terwijl berekening van deze grootte op basis van het boekhoudnet een oppervlakte van 302.513 hectare oplevert, waarvan 225.435 hectare op grasland (overschatting 8%), 53.729 hectare op bouwland (-2%), en 9.082 hectare op tuinbouwland (-23%).

Een ander gegeven over de bedrijfsvoering is de verkavelingssituatie. In 1983 is in de metelling een gemiddeld aantal kavels van 4,01 geregistreerd in de populatie. De steekproef komt uit op een gemiddelde van 4,14, wat neerkomt op een lichte overschatting. De oppervlakte van de huiskavel - gemeten op 1 mei 1984 - verschilt eveneens niet veel tussen populatie- en steekproefbedrijven. De huiskavel was gemiddeld respectievelijk 9,1 en 9,6 hectare. In 1978 werden gemiddelden van 8,4 en 8,7 waargenomen.

Het percentage hoofdberoepsbedrijven is in de steekproef iets groter dan in de populatie. Bedrijven met rechtspersoonlijkheid zijn sterk ondervertegenwoordigd in de steekproef. In de periode 1975-1981 was er slechts een vertegenwoordiging van 20%. In de jaren daarna was dat 40%. Het gaat hier overigens om een kleine groep bedrijven binnen de populatie.

Het aantal opvolgers wordt door het boekhoudnet met 6% overschat. In de periode 1975-1981 was er sprake van een lichte onderschatting op dit punt.

Een laatste gegeven waarnaar min of meer regelmatig gevraagd wordt in de metelling is de exploitatievorm van de grond. Zowel in de periode 1975-1981 als in de periode 1982-1986 is de gepachte oppervlakte oververtegenwoordigd in de steekproef. Daar staat een ondervertegenwoordiging van de oppervlakte in eigendom tegenover.

11.3 Vergelijking steekproef met metelling (FIP)

11.3.1 Inleiding

Om de representativiteit van de steekproefbedrijven die voor het samenstellen van de FIP worden gebruikt, te controleren, staan in principe twee wegen open. In de eerste plaats kunnen alle bedrijven die een FIP-wegingsfactor krijgen, vergeleken worden met de bedrijven in de populatie. In de tweede plaats kunnen alle bedrijven die meegenomen worden in de berekeningen - dat is dus exclusief de rechtspersonen en dergelijke - vergeleken worden met de bedrijven in de populatie.

Bij de eerste vergelijking zal de som van de wegingsfactoren gelijk zijn aan het aantal bedrijven in de populatie en kan bekeken worden in hoeverre bedrijven die bereid zijn een financieeringsboekhouding te leveren, representatief zijn voor alle bedrijven. Bij de tweede vergelijking zal de som van de wegingsfactoren kleiner zijn en is vergelijking minder goed mogelijk, omdat bewust een aantal bedrijven buiten beschouwing wordt gelaten. De tweede vergelijking geeft wel een goed beeld van de groep bedrijven waar het eigenlijk over gaat in de FIP. Tabel 11.2 draagt het materiaal aan voor beide vergelijkingen.

11.3.2 Vergelijking voor de totale populatie

Evenals bij de vergelijking van BUL-bedrijven met populatiebedrijven, zijn er in de jaren 1982-1986 slechts kleine afwijkingen ten aanzien van de waarden van de stratificatie-variabelen tussen FIP-steekproef en populatie. Het aantal bedrijven met rechtspersoonlijkheid is nog sterker ondervertegenwoordigd dan bij de BUL-bedrijven. De FIP-bedrijven leveren een totaal van circa 200 van dergelijke bedrijven, terwijl er circa 800 in de populatie aanwezig zijn. Het areaal handelsgewassen en granen wordt met 11% respectievelijk 5% onderschat, terwijl dit bij de BUL-bedrijven niet het geval was in de betreffende jaren. Het aantal vreemde arbeidskrachten wordt nog minder sterk gerepresenteerd dan in de BUL. Hetzelfde geldt voor de oppervlakte grond die niet in gebruik is, de oppervlakte bos en de oppervlakte overige gronden. Tabel 11.2 geeft een overzicht voor een groot aantal variabelen (kolom FIP+).

11.3.3 Uitsluiting van bedrijven met een onvolledige boekhouding

Worden de bedrijven zonder volledige financieringsboekhouding (de NV's en BV's) buiten de vergelijking gehouden, dan ontstaat een iets ander beeld. Ongeveer 4% van de bedrijven (wegingsfactoren) valt dan buiten de vergelijking. Het blijkt dat dan 93% van het totale aantal sbe in de populatie door de FIP-bedrijven wordt gerepresenteerd. Naast een kleiner aantal bedrijven is er dus ook een kleinere gemiddelde grootte van de FIP-bedrijven. De FIP-bedrijven zijn gemiddeld 3% kleiner dan het gemiddelde populatiebedrijf. In oppervlakte gemeten zijn ze 8% kleiner.

Het aantal rechtspersonen op FIP-bedrijven is gemiddeld 11% van dit aantal in de populatie.

De arealen granen en handelsgewassen bedragen respectievelijk 84% en 68% van de arealen in de populatie. Overigens laten vrijwel alle akkerbouwgewassen een iets sterkere daling dan 7% (100-93%) zien. Dat betekent dat op bedrijven zonder volledige financieringsboekhouding relatief veel van deze gewassen voorkomen. Dat spoort met het gegeven uit paragraaf 10.5 dat akkerbouwbedrijven vrij vaak niet een dergelijke boekhouding kunnen leveren. In tabel 11.2 wordt meer informatie gegeven (kolom FIP-).

Tabel 11.2 Vergelijking meetellingsgegevens van steekproef- en meetellingsbedrijven. Totalen voor de nationale FIP-boerderij in procenten van de totalen van de totale populatie (gemiddelde van boekjaren 1982-1986)

Variabele	Gemiddelde waarde in populatie	Totaal-index voor boekhoudnetbedrijven		
		gemiddelde waarde (tussen haakjes het aantal bedrijven met een waarde > 0 en de gemiddelde waarde op die bedr.)		
		FIP +	FIP -	a)
Sbe totaal	209,9	101 (100*101)	93 (96* 97)	
Bss totaal	58,8	101 (100*101)	94 (96* 98)	
Oppervlakte (aren):				
- cultuurgrond vorig jr.	2206	101 (102* 99)	92 (98* 94)	
- cultuurgrond	2294	100 (101* 99)	92 (97* 95)	
- bouwland	854	100 (102* 98)	90 (97* 92)	
- blijv. grasland	1342	100 (100*100)	93 (96* 97)	
- tijd. grasland	45	86 (103* 84)	75 (95* 79)	
- tuinbouw open gr.	50	101 (110* 92)	93 (105* 88)	
- tuinbouw glas	0	0 (0* 0)	0 (0* 0)	
- braakland	2	122 (106*115)	103 (94*109)	

Tabel 11.2 (1e vervolg)

Variabele	Gemiddelde waarde in populatie	Totaal-index voor boekhoudnetbedrijven		
		FIP +	FIP - a)	
- eigendom	b)	1330	97 (102* 95)	89 (98* 91)
- pacht	b)	881	103 (103*100)	96 (99* 97)
- erfpacht	b)	71	99 (100* 99)	84 (93* 91)
- pacht van ouders	c)	122	130 (137* 95)	123 (132* 93)
- andere expl.vormen	b)	15	81 (107* 75)	79 (106* 75)
- kadastraal		2360	100 (101* 99)	92 (97* 95)
- niet in gebruik	b)	1	86 (92* 94)	86 (92* 94)
- bos	b)	23	25 (104* 24)	24 (99* 24)
- overige gronden		93	67 (101* 67)	63 (97* 65)
- tuin eigen gebruik	b)	1	122 (118*103)	119 (115*103)
- totaal bedrijf		2467	98 (100* 98)	90 (96* 94)
Aantal melkkoeien		33	102 (101*101)	95 (98* 97)
Sbe:				
- rundvee fokkerij		104,1	102 (100*102)	95 (97* 98)
- rundvee mesterij		2,3	84 (98* 86)	79 (93* 85)
- mestkalveren		2,4	71 (86* 82)	68 (84* 81)
- schapen		0,9	82 (106* 78)	78 (101* 77)
- geiten	g)	0,1	8 (80* 10)	8 (80* 10)
- paarden		1,4	51 (84* 61)	49 (81* 61)
- grasland		15,3	100 (100* 99)	92 (96* 96)
- mestvarkens		10,8	99 (109* 91)	95 (106* 90)
- fokvarkens		18,4	105 (105*100)	101 (103* 98)
- slachtkuikens		1,4	90 (83*109)	88 (82*108)
- leghennen				
(incl. moederdieren)		6,0	110 (110*100)	104 (107* 97)
- moederdieren < 6 mnd.		0,3	141 (170* 83)	141 (170* 83)
- moederdieren > 6 mnd.		1,1	124 (132* 94)	123 (131* 94)
- leghennen < 5 mnd.		0,7	137 (144* 95)	134 (137* 98)
- leghennen 5 -20 mnd.		3,4	102 (96*106)	94 (93*101)
- leghennen >= 20 mnd.		0,6	83 (83*100)	72 (78* 93)
- eenden		0,1	0 (0* 0)	0 (0* 0)
- kalkoenen		0,1	277 (310* 89)	277 (310* 89)
- peulvruchten		0,6	115 (122* 94)	106 (115* 92)
- landbouwzaden		0,7	120 (111*108)	112 (102*110)
- handelsgewassen		0,4	89 (107* 83)	68 (94* 72)
- aardappelen totaal		15,2	107 (106*101)	95 (99* 95)
- pootaardappelen		5,2	115 (116*100)	96 (103* 93)
- fabrieksaardappelen		3,3	112 (118* 95)	103 (114* 91)
- cons. aardappelen		6,7	98 (102* 97)	90 (96* 94)

Tabel 11.2 (2e vervolg)

Variabele	Gemiddelde waarde in populatie	Totaal-index voor boekhoud-netbedrijven		
		FIP +	FIP -	a)
- granen	6,0	95 (104* 92)	84 (98* 85)	
- suikerbieten	11,7	104 (105* 99)	93 (98* 94)	
- ov. akkerbouw	8,6	96 (102* 94)	89 (98* 91)	
- voedergewassen	6,5	91 (100* 92)	85 (96* 89)	
- tuinakkergewassen	2,4	84 (107* 78)	78 (101* 77)	
- tuinbouwgewassen	3,1	83 (115* 72)	78 (110* 71)	
Hokcapaciteiten:				
- mestkalveren	8	73 (80* 92)	70 (77* 91)	
- mestvarkens	87	99 (110* 90)	95 (107* 89)	
- slachtkuikens	c) 439	91 (94* 97)	88 (92* 96)	
- moederdieren	c) 59	155 (150*104)	140 (143* 98)	
- leghennen > 5 mnd.	c) 385	98 (98*100)	88 (95* 92)	
Arbeidskrachten:				
- mannelijk	1,38	98 (101* 97)	92 (97* 94)	
-bedrijfshoofd	0,83	103 (103*100)	103 (103*100)	
-vreemde arb. kr.	0,12	67 (89* 76)	58 (80* 73)	
-meewerkende echtgen.	0,00	20 (20*100)	20 (20*100)	
- vrouwelijk	0,44	91 (91* 99)	90 (91* 99)	
-bedrijfshoofd	0,01	56 (55*100)	53 (53*100)	
-vreemde arb. kr.	0,01	33 (44* 75)	33 (44* 75)	
-meewerkende echtgen.	0,41	93 (93*100)	92 (92* 99)	
Aantal kavels	d) 4,01	101 (101*101)	96 (96* 99)	
Opp. huiskavel (aren)	e) 914	104 (103*101)	95 (98* 96)	
% bedr. met ligb. stal	e) 31	106 (106*100)	99 (99*100)	
Beregenbare opp.(1983)	d) 590	114 (112*102)	106 (107* 99)	
Te beregenen opp.(1985)	f) 413	103 (99*103)	98 (95*103)	
- opp. gras	f) 323	104 (102*102)	99 (98*101)	
- opp. bouwland	f) 77	106 (93*114)	101 (89*113)	
- opp. tuinbouw	f) 13	63 (90* 69)	56 (85* 66)	
Leeftijd bedrijfshoofd	48,55	99 (100* 99)	95 (96* 99)	
Geb. jr. opvolger (indien aanwezig)	(1959,1) e)	105 (105*100)	98 (98*100)	
% bedr. met				
- hoofdberoep landbouwer	97	98 (98*100)	94 (94*100)	
- rechtspersoonlijkheid	1	25 (25*100)	11 (11*100)	
- machtiging landbouwschap	99	101 (101*100)	97 (97*100)	

Noten bij tabel 11.2:

- a) FIP + : op basis van alle bedrijven met een uitgewerkte financieringsboekhouding.
FIP - : op basis van alle FIP-bedrijven (excl. NV's, BV's, en dergelijke).
- b) Alleen gevraagd in metelling 1983 en 1985.
- c) Niet gevraagd in metelling 1985.
- d) Alleen gevraagd in metelling 1983.
- e) Alleen gevraagd in metelling 1984.
- f) Alleen gevraagd in metelling 1985.
- g) Alleen gevraagd in metelling 1984-1986.

11.4 Vergelijking steekproef met externe bronnen

11.4.1 Inleiding

Een belangrijk punt bij de evaluatie van de representativiteit van het boekhoudnet, is de vraag in hoeverre de steekproef niet alleen qua metellingskenmerken overeenstemt met de populatie, maar ook voor wat betreft de bedrijfsuitkomsten. De wijze waarop beslist en gewerkt wordt is uiteraard niet gelijk op alle bedrijven die qua metellingskenmerken sterk overeenkomen. Het management van de ondernemer, de kwaliteit van de grond, de inrichting en de efficiëncy van het produktieproces, en dergelijke, kunnen een belangrijke invloed hebben op de bedrijfsuitkomsten.

Voorkomen moet worden dat het LEI-boekhoudnet te veel bedrijven met een minder dan gemiddeld resultaat zou bevatten, of juist te veel bedrijven die zeer gunstige resultaten boeken. Een aantal "kwaliteits"-kenmerken van het gemiddelde boekhoudnetbedrijf zijn via vergelijking met externe bronnen te controleren. Voor de weidebedrijven zijn bijvoorbeeld de gemiddelde melkproduktie per koe, het aantal ligboxenstallen en het aantal voercomputers bekend. In de akkerbouw zijn er de gemiddelde opbrengsten per hectare, die vergeleken kunnen worden met CBS-oogstramingen. In de intensieve veehouderij kunnen voederconversies, aantal biggen per zeug, aantal eieren per leggen, en dergelijke, geconfronteerd worden met externe gegevens.

De mate waarin de *spreiding* van uitkomsten van het boekhoudnet representatief is voor de spreiding in de populatie is moeilijker te controleren. Op dit aspect van de representativiteit wordt in deze paragraaf niet ingegaan.

11.4.2 Rundveehouderij

Volgens Landbouwcijfers 1987 bedroeg de gemiddelde melkproduktie per koe in 1985 5.350 kg met een vetpercentage van 4,17%. In 1986 zijn deze cijfers 5.637 kg melk en 4,26% vet. Het boekhoudnet - dat betrekking heeft op alle bedrijven die groter zijn dan 79 sbe en kleiner dan 2.000 sbe en waarop 96% van alle melk-

koeien in Nederland voorkomen - komt voor boekjaar 1985/86 op 5.688 kg melk met een vetpercentage van 4,22%.

In onderstaande tabel zijn de gegevens uit de twee bronnen voor een aantal jaren naast elkaar gezet. De LEI-cijfers komen iets hoger uit dan de CBS-cijfers.

Op zichzelf zegt de vergelijking van cijfers over de produktie per koe niet alles, omdat de gemiddelde cijfers het resultaat zijn van een deling van de totale melkproduktie door een totaal aantal melkkoeien. Voor de bepaling van het aantal melkkoeien volgt het LEI een andere berekeningswijze dan het CBS. Een vergelijking van de totale melk- en melkvet-produktie kan daarom het inzicht vergroten.

Tabel 11.3 Melkgift per koe en vetpercentage volgens boekhoudnet en volgens Landbouwcijfers naar (boek-)jaar

Boekjaar/ Kalenderjaar	Kg melk per koe		Vetpercentage	
	LEI	CBS	LEI	CBS
1981	5.346	5.080	4.12%	4.00%
1982	5.496	5.300	4.09%	4.09%
1983	5.455	5.305	4.15%	4.13%
1984	5.415	5.210	4.17%	4.13%
1985	5.688	5.350	4.22%	4.17%
1986	6.008	5.637	4.31%	4.26%
1987 *)		5.815		4.31%

*) Voorlopig.

Bron: Landbouwcijfers, tabel 52-a (diverse jaren); Bedrijfsuitkomsten, tabel 2.2 (diverse jaren).

Tabel 11.4 Totale melk- en melkvet-produktie volgens boekhoudnet en volgens Landbouwcijfers naar (boek-)jaar

Kalender- jaar of boekjaar	Totale melkproduktie (mln. tonnen)			Totale melkvetproduktie (mln. tonnen)		
	CBS	LEI	LEI/CBS	CBS	LEI	LEI/CBS
1980	11.851	12.081	102%	474	487	103%
1981	12.185	12.465	103%	495	513	104%
1982	12.817	13.317	104%	524	545	104%
1983	13.207	13.473	102%	545	559	102%
1984	12.732	12.996	102%	526	542	103%
1985	12.528	13.236	106%	522	558	107%
1986	12.710	13.132	103%	540	566	105%

Bron: Landbouwcijfers, tabel 52-a (diverse jaren); Bedrijfsuitkomsten, tabel 2.1 (diverse jaren).

Tabel 11.4 laat zien dat de hogere melkproduktie per koe in het boekhoudnet niet alleen veroorzaakt wordt door een kleiner aantal koeien in het boekhoudnet dan in de CBS-cijfers. Immers, ook de totale melk(vet-)produktie in Nederland wordt door het boekhoudnet enigszins overschat. Daarbij moet aangetekend worden dat het boekhoudnet betrekking heeft op een groep bedrijven waar ruim 96% van het aantal melkkoeien in Nederland wordt gehouden, terwijl de CBS-cijfers betrekking hebben op 100% van het aantal melkkoeien. Ervan uitgaande dat de koeien op de bedrijven buiten de populatie gemiddeld per koe een iets lagere produktie hebben dan de koeien op bedrijven in de populatie, komt de produktie buiten de populatie op ruim 3% van de totale produktie. De overschatting van het boekhoudnet op het punt van de totale melkvetproduktie is dus nog ongeveer 3% groter dan de 4% die de laatste kolom in tabel 11.4 aangeeft.

Van deze totale overschatting van ongeveer 7% is ongeveer 3% toe te schrijven aan een overschatting van het aantal melkkoeien (tabel 11.1). De produktie *per koe* is dus volgens de LEI-steekproef 4% hoger dan volgens de CBS-cijfers. Daarbij moeten wel de volgende kanttekeningen geplaatst worden. In de eerste plaats is er een bepaalde mate van overschatting van de totale melk(vet-)produktie te verwachten door de manier van berekening van de wegingsfactoren (zie paragraaf 9.4 en 9.5). Vooral in jaren waarin veel bedrijven worden opgeheven speelt dat een rol. In de tweede plaats loopt het LEI-boekjaar enige maanden achter bij het kalenderjaar, waardoor een exacte vergelijking niet helemaal mogelijk is. Deze twee factoren leiden tot de conclusie dat de gemiddelde melkvetproduktie per koe waarschijnlijk minder dan 4% - ongeveer 3% - overschat wordt. Het is mogelijk dat deze overschatting samenhangt met een overschatting van het aandeel zwartbont vee in de steekproef (zie verderop in deze paragraaf).

Wat de gevolgen zijn van een overschatting van de melkproduktie per koe van 3% voor de bedrijfsuitkomsten, is niet eenvoudig aan te geven. Het is namelijk moeilijk na te gaan in hoeverre de steekproefbedrijven ook aan de kostenkant overschattingen te zien geven, bijvoorbeeld in de vorm van hogere krachtvoerkosten dan het gemiddelde bedrijf in de populatie.

Een gegeven dat in het boekhoudnet wordt vastgelegd voor de bedrijven met melkkoeien, is de deelname aan de melkcontrole. Daaruit kan worden berekend hoeveel bedrijven in de gehele populatie deelnemen aan de melkcontrole. In Landbouwcijfers worden ook deelname-cijfers gepresenteerd, gebaseerd op gegevens van de Centrale Melk Controle (CMD). In tabel 11.5 worden de cijfers uit de twee bronnen naast elkaar gezet.

Hoewel de populatie van het LEI-boekhoudnet kleiner is dan die waarop de gegevens van de CMD betrekking hebben, komt het LEI tot een fors hogere schatting van het aantal bedrijven met deelname aan de CMD-controle. Bedrijven die deelnemen aan de melkcontrole zijn blijkbaar meer geneigd om deel te nemen aan het LEI-

Tabel 11.5 Aantal bedrijven met melkcontrole volgens boekhoudnet en volgens Centrale Melkcontrole Dienst (CMD) naar boekjaar

Boekjaar	Melkcontrole volgens:	
	boekhoudnet	CMD *)
1980	42.001	34.025
1981	42.534	33.622
1982	41.785	33.853
1983	40.966	34.447
1984	40.826	34.835
1985	38.585	34.455
1986	37.886	34.273

*) Per 1 juli van het betreffende boekjaar.

Bron: LEI-databank; Landbouwcijfers, tabel 43-d; Jaarverslagen NRS.

boekhoudnet dan bedrijven die geen melkcontrole laten uitvoeren. Het verschil tussen de twee bronnen lijkt de laatste jaren iets minder groot te worden.

Het al dan niet bezitten van een ligboxenstal is een belangrijk gegeven in de rundveehouderij. Voor de bedrijven in het boekhoudnet is dit gegeven bekend. Confrontatie met externe cijfers (tabel 11.6) leidt tot de conclusie dat het LEI-boekhoudnet het aantal bedrijven met een ligboxenstal overschat - of dat de andere bron dit aantal onderschat. Een aanwijzing voor dit laatste is te vinden in paragraaf 11.2.8. Daar bleek dat het aantal bedrijven dat in de metelling van 1984 opgaf een ligboxenstal te bezitten, door de steekproef slechts met 5% werd overschat.

Tabel 11.6 Aantal bedrijven met een ligboxenstal volgens boekhoudnet en volgens Landbouwcijfers naar (boek-)jaar

(Boek-)jaar	Bedrijven met een ligboxenstal volgens boekhoudnet	Aantal ligboxenstallen volgens Landbouwcijfers
1980	20.849	17.933
1981	22.054	19.448
1982	22.765	20.127
1983	24.033	20.799
1984	23.795	21.750
1985	24.538	22.195
1986	24.777	22.496

Bron: Landbouwcijfers, tabel 32-d.

De bedrijven in het boekhoudnet geven een lichte overschatting van het aantal (bedrijven met) roodbonte koeien. Volgens Landbouwcijfers (1986, tabel 43-d) bestond de Nederlandse melkveestapel in 1981 voor 35% uit roodbonten en voor 62,5% uit zwartbonten. Indien alle koeien op LEI-bedrijven die geregistreerd staan als bedrijven met (overwegend) zwartbont vee, gerekend worden als zwartbont vee, en hetzelfde principe wordt toegepast voor roodbont vee, dan komt het LEI-boekhoudnet in 1981 - en ook voor de jaren daarna - op ongeveer 28% roodbonte melkkoeien en 70% zwartbonte koeien. Omdat er meer roodbont vee zal voorkomen op bedrijven met (overwegend) zwartbont vee dan omgekeerd, zullen deze cijfers iets ten nadele van het roodbonte vee vertekend zijn, maar zeker niet zoveel dat daarmee het verschil tussen 28% en 35% verklaard kan worden. Regionale verschillen in non-respons die maar gedeeltelijk door de stratificatie gecorrigeerd worden, liggen hier wellicht aan ten grondslag; roodbont vee komt namelijk in specifieke regio's veel voor en in andere regio's nauwelijks.

11.4.3 Akkerbouw

Voor een aantal gewassen maakt het CBS in het najaar een oogstraming. Door deze te vergelijken met de definitieve LEI-boekhoudnetcijfers - die anderhalf jaar later gepubliceerd worden - kan een indruk verkregen worden van de mate waarin steekproefbedrijven voor wat betreft de kilogram-opbrengsten per hectare overeenkomen met alle bedrijven in de populatie. Bij de vergelijking moet wel bedacht worden dat de LEI-steekproef betrekking heeft op de landbouwbedrijven die groter zijn dan een bepaalde minimale sbe-omvang, terwijl het CBS zich richt op de gehele oppervlakte bouwland. Daarnaast spelen soms definitiever-schillen een rol in de vergelijking. Per gewas zullen deze verschillen worden aangegeven.

Tabel 11.7 geeft de vergelijking voor wintertarwe en zomergerst. De CBS-cijfers hebben betrekking op het handelsgewicht bij 17% vocht. De LEI-cijfers hebben betrekking op de werkelijk geconstateerde gewichten en zijn bovendien gebaseerd op gegevens van akkerbouwbedrijven.

Voor wintertarwe is er een opmerkelijk goede aansluiting tussen gegevens van de oogstraming en het boekhoudnet. De iets hogere opbrengsten in de LEI-steekproef worden voor een deel veroorzaakt door het buiten beschouwing laten van wintertarwe op niet-akkerbouwbedrijven. Voor zomergerst zijn de verschillen wat groter, al lijkt er vanaf 1983 een vrij nauwe aansluiting tussen beide bronnen te zijn.

Bij haver is er eveneens een goede aansluiting tussen steekproef- en oogstramings-gegevens. Gedurende de boekjaren 1975-1979 kwam het LEI uit op 4.750 kg per hectare (Bedrijfsuitkomsten,

1981:56-57), terwijl het CBS een gemiddelde van 4.800 kg per hectare noteerde (Landbouwcijfers, 1980).

Tabel 11.8 geeft een vergelijking voor de gewassen fabrieksaardappelen en suikerbieten. Een handicap in de vergelijking van de twee bronnen is dat de CBS-oogstramingen voor fabrieksaardappelen inclusief pootgoed zijn, en de LEI-cijfers exclusief pootgoed. Verder berekent het LEI de hectare-productie op basis van

Tabel 11.7 Kilogram-opbrengsten van wintertarwe en zomergerst (CBS-oogstraming en LEI-boekhoudnet) *)

(Boek-) jaar	Wintertarwe		Zomergerst	
	CBS-oogst raming	LEI-boekhoudnet	CBS-oogst raming	LEI-boekhoudnet
1976	5.700	5.900	4.200	4.800
1977	5.400	5.600	4.300	4.200
1978	6.800	7.000	4.900	4.900
1979	6.100	6.300	4.500	4.600
1980	6.400	6.600	4.600	5.100
1981	6.900	7.100	4.500	4.900
1982	7.600	7.800	5.700	6.100
1983	7.100	7.200	4.500	4.600
1984	7.900	8.100	5.300	5.400
1985	6.700	6.600	5.000	5.200
Gemiddeld	6.700	6.800	4.800	5.000

*) De LEI-cijfers hebben alleen betrekking op akkerbouwbedrijven.

Bron: Landbouwcijfers, tabel 51-c, diverse jaren; Poppe (1987) tabel 2.4.

een onderwatergewicht van 400 gram. Aan de vergelijking kan dus zonder aanvullende informatie nauwelijks een conclusie worden ontleend. Daarom is gebruik gemaakt van door het LEI berekende "aangepaste oogstramingen" die de CBS-cijfers vergelijkbaar maken met de LEI-cijfers. Deze aanpassingen worden ieder jaar gemaakt ten behoeve van de LEI-prognoses (paragraaf 15.5).

Het LEI komt voor de gemiddelde hectare-opbrengst van fabrieksaardappelen in de periode vanaf 1977 iets hoger uit dan het CBS. De verschillen zijn echter klein.

Bij suikerbieten zijn de verschillen ook vrij klein, zeker voor wat betreft de laatste boekjaren. De LEI-cijfers liggen gemiddeld 3% hoger dan de CBS-cijfers.

Bij consumptie-aardappelen doet zich het probleem voor dat deze in de CBS-ramingen worden samengevoegd met de pootaardappe-

len. Om een goede vergelijking te kunnen maken, zouden ook de LEI-cijfers berekend moeten worden over deze totale categorie. Daar komt bij dat de oppervlaktes voor de beide categorieën verschillen in de twee bronnen, waardoor een totaalcijfer ook beïnvloedt wordt door de weging van de verschillende categorieën. Door het grotere aandeel pootaardappelen in het totaal in de LEI-cijfers ligt een lagere hectare-opbrengst dan in de CBS-cijfers

Tabel 11.8 Kilogram-opbrengsten in tonnen per ha van fabrieksaardappelen en suikerbieten (CBS-oogstraming en LEI-boekhoudnet)

Boekjaar	Fabrieksaardappelen		Suikerbieten	
	CBS-oogstraming	*) LEI-boekhoudnet	CBS-oogstraming	LEI-boekhoudnet
1977	33.7	33.7	46.5	48.9
1978	42.2	40.1	48.5	50.4
1979	42.8	43.7	44.5	47.2
1980	35.5	36.9	49.0	52.0
1981	42.8	44.5	54.0	56.2
1982	35.8	39.4	59.5	62.1
1983	36.1	35.4	46.5	46.7
1984	43.0	42.5	54.0	53.6
1985	42.0	42.0	48.5	48.9
1986	40.4	41.2	56.0	56.6
Gemiddeld	39.4	39.9	50.7	52.3

*) Exclusief pootaardappelen bij een onderwatergewicht van 400 gram.

Bron: Landbouwcijfers, tabel 51-c, diverse jaren; Poppe, 1987, tabel 2.4; Bedrijfsuitkomsten, tabel 2.2, diverse jaren.

voor de hand. De indruk bestaat echter dat indien daarvoor gecorrigeerd wordt, de LEI-cijfers in het algemeen nog aan de lage kant zijn. Daaruit kunnen echter moeilijk conclusies getrokken worden, omdat verschillen in rassen - met daarbij behorende verschillen in kilogram- en geldopbrengsten - aan dit verschil ten grondslag kunnen liggen.

11.4.4 Intensieve veehouderij

In de pluimvee- en varkenssector worden ook buiten de LEI-staekproef om veel technische en economische kengetallen geproduceerd. Het LEI verwerkt jaarlijks - in samenwerking met het CAD - de gegevens van enkele duizenden bedrijven die meedoen aan de

Technisch-Economische Administratie (TEA) van de Consulentschappen. Toch is het moeilijk om de steekproefuitkomsten naast andere cijfers te leggen, om zo te komen tot een beoordeling van de representativiteit van de steekproef. De oorzaak daarvan ligt niet in de uitsluiting van de kleine bedrijven in de steekproef; deze komen in het algemeen ook nauwelijks voor in de TEA-cijfers. Het grootste probleem ligt in de selectie van de bedrijven die deelnemen aan de TEA. Terwijl de LEI-bedrijven op steekproefbasis gekozen worden, gebeurt deelname aan de TEA op eigen initiatief. Als gevolg daarvan is het mogelijk dat de TEA-administraties een te gunstige voorstelling van de stand van zaken in de intensieve veehouderij weergeven.

Desondanks is het de moeite waard om tot een vergelijking te komen. Voor de varkenssector zijn de mogelijkheden daartoe aanwezig. Voor de overige sectoren in de intensieve veehouderij (pluimveehouderij, kalvermesterij) wordt een vergelijking moeilijker, vanwege de kleine steekproefaantallen. Uit tabel 11.1 bleek al dat voor sommige groepen binnen de intensieve veehouderij geldt dat de bedrijven in de steekproef gemiddeld kleiner zijn dan in de populatie. Dat zal zeker invloed hebben op de technische en economische resultaten. De steekproefuitkomsten geven daardoor waarschijnlijk een iets te ongunstig beeld van de situatie in deze sectoren.

Voor bedrijven met varkens kan een globale vergelijking van cijfers uit de twee bronnen gemaakt worden door de TEA-bedrijven wegingsfactoren te geven die ervoor zorgen dat een goede representatie van de nationale boerderij ten aanzien van regio en grootte van bedrijven wordt verkregen. Na herberekening van de TEA-kengetallen op een manier die vergelijking met LEI-kengetallen mogelijk maakt, en na herleiding van de cijfers op LEI-boekjaren, zijn de volgende conclusies te trekken. Het zijn voorlopige conclusies op basis van lopend onderzoek door de sectie Prognoses, Rentabiliteit en Financiering (PRF) binnen de afdeling Landbouw van het LEI.

Gemiddeld over de boekjaren 1983-1986 komen de technische en economische resultaten op LEI-steekproefbedrijven iets ongunstiger uit dan op TEA-bedrijven. De voederconversie bij mestvarkens ligt iets hoger, de groei per dier per dag iets lager. Een en ander leidt ertoe dat het jaarsaldo per gemiddeld aanwezig mestvarken op LEI-bedrijven ruim 10 gulden lager ligt dan op TEA-bedrijven. De ontwikkelingen van jaar op jaar in technische en economische kengetallen lopen in veel gevallen parallel.

Bij de bedrijven met fokzeugen zijn soortgelijke conclusies te trekken. Het aantal grootgebrachte biggen per worp verschilt nauwelijks tussen LEI- en TEA-bedrijven. Het aantal worpen per zeug per jaar ligt op de steekproefbedrijven iets lager dan op de TEA-bedrijven en dientengevolge geldt hetzelfde voor het aantal grootgebrachte biggen per zeug per jaar. De technische resultaten zorgen er samen met een hoger voerverbruik per zeug voor dat het

jaarsaldo per fokzeug op LEI-bedrijven in de periode 1983-1986 gemiddeld ruim 50 gulden lager lag dan op de TEA-bedrijven.

Opvallend is verder dat de verschillen zich met name voordoen bij vergelijkbare groepen kleinere bedrijven. Dat geldt zowel voor de bedrijven met mestvarkens als voor de bedrijven met fokvarkens.

Gezien de aard van de gegevensverzameling ligt het enigszins voor de hand dat bedrijven met een TEA-administratie tot betere resultaten komen dan LEI-steekproefbedrijven. De zelfselectiviteit van TEA-bedrijven en de leereffecten bij langere deelname aan deze administratie, zullen er waarschijnlijk voor zorgen dat de "betere" bedrijven oververtegenwoordigd zijn in de TEA-cijfers. Of de LEI-steekproef een te positief of een te negatief beeld geeft van de situatie in de varkenshouderij, valt niet met zekerheid te zeggen, maar de vergelijking met de TEA-cijfers doet vermoeden dat de representativiteit vrij goed is.

11.5 Vergelijking 32 en 275 wegingsfactoren

In hoofdstuk 9 is aangegeven dat bij een stratificatie achteraf de 32 basisstrata die gevormd worden bij de bedrijfskeuze in elk geval weer aan bod moeten komen. Omdat binnen de basisstrata dezelfde trekkingspercentages zijn gehanteerd voor substrata, ligt het niet meteen voor de hand om bij een stratificatie achteraf ook substrata te vormen binnen de basisstrata.

Voor de nationale boerderij is onderzocht in welke mate een reductie tot 32 wegingsfactoren - één wegingsfactor per basisstratum - leidt tot andere steekproefuitkomsten. Tabel 11.10 geeft de resultaten.

De verschillen zijn in het algemeen vrij klein. De oppervlakte en het sbe-aantal worden iets groter bij toepassing van slechts 32 wegingsfactoren, terwijl de gemiddelde leeftijd van de ondernemer iets daalt. De oppervlakte marktbaar gewassen neemt met bijna 5% toe. Voor variabelen in de kosten- en opbrengst sfeer maakt het weinig verschil of uitgegaan wordt van 275 of van 32 wegingsfactoren.

Uiteraard zegt het niet alles indien de gemiddelde waarde van een variabele niet verandert bij een vermindering van het aantal wegingsfactoren. Een te geringe representatie van oudere ondernemers bij 32 factoren kan gedeeltelijk gecompenseerd worden door een eveneens te gering aantal jongere ondernemers, waardoor de gemiddelde leeftijd maar weinig verandert.

De indruk bestaat dat door de verdergaande stratificatie naar 275 strata, met name de spreiding van de bedrijfskenmerken en van de resultaten beter gerepresenteerd wordt.

Bij toepassing van 32 gewichten bij berekeningen voor BUL-groepen, kunnen veel grotere verschillen met de normale bereke-

ningen voorkomen. Dat zal zich met name voordoen in BUL-groepen die samengesteld zijn uit delen van strata die een afwijkend trekkings- of uitwerkingspercentage hebben ten opzichte van ande-

Tabel 11.10 Steekproefuitkomsten voor de gemiddelde nationale boerderij (BUL) bij 32 en bij 275 wegingsfactoren (gemiddeld voor de boekjaren 1982-1986)

Variabele	Geschatte waarde bij		Verschil in %
	275 wegings- factoren	32 wegings- factoren	
Opp. cult. grond (ha)	23,69	24,09	1,7
Opp. blijvend grasland	13,83	13,88	0,4
Opp. kunstw. en v.gew.	1,45	1,41	- 2,8
Opp. markt. gewassen	8,37	8,76	4,7
Opp. boomgaard en uitl.	0,04	0,04	0,0
Bal.waarde (* fl.000,-)			
- grond en gebouwen	523	528	1,0
- vee	146	145	- 0,7
- werktuigen	79	79	0,0
Aantal			
- melkkoeien	34,0	34,0	0,0
- ov. weidevee (oke)	12,7	12,6	- 0,8
- fokzeugen	14,1	14,3	1,4
- mestvarkens	57,1	56,2	- 1,6
- leghennen	516	535	3,7
- slachtkuikens	490	461	- 5,9
Sbe totaal (gem.)	209	210	0,5
Sbe intensieve veeh.	37	37	0,0
Sbe totaal (mei)	204	205	0,5
Leeftijd ondernemer	48	47	- 2,1
Volw. arbeidskrachten	1,5	1,5	0,0
Melkproductie (kg.)	186.839	186.576	- 0,1
Factorkosten			
- arbeid	92.472	92.179	- 0,3
- rente	17.273	17.290	0,1
- netto pacht	18.640	18.784	0,8
- werk door derden	4.159	4.030	- 3,1
- totaal	132.544	132.283	- 0,2
Non-factorkosten	219.161	217.708	- 0,7
Totale kosten	351.705	349.991	- 0,5
Totale opbrengsten	324.735	323.264	- 0,5
Netto overschot	- 26.970	- 26.726	+ 0,9
Arbeidsopbrengst	65.502	65.452	- 0,1
Factoropbrengst	105.573	105.556	- 0,0

re strata binnen het basisstratum waartoe ze behoren. In de akkerbouwsector zorgen verschillende steekproefpercentages voor verschillende regio's binnen een basisstratum - voortvloeiend uit gerichte regionale aanvullingskeuzes - voor vrij grote verschillen in de twee berekeningswijzen.

Bedacht moet worden dat de verschillen tussen de twee berekeningswijzen gerelateerd moeten worden aan de standaardfouten van de steekproefschattingen. Bij een relatieve standaardfout van bijvoorbeeld 1% verdient een systematische fout van 2% in het algemeen meer aandacht dan bij een relatieve standaardfout van 20%.

Verder is het ook weer niet zo verrassend dat er maar kleine verschillen zijn tussen de berekening op basis van 32 en die op basis van 275 gewichten. In principe hebben alle bedrijven binnen een basisstratum immers gelijke kansen om in de steekproef terecht te komen. Verschillen in non-respons tussen oppervlakte-

Tabel 11.11 Overzicht van steekproefuitkomsten voor de gemiddelde nationale boerderij (FIP) bij 32 en bij 275 wegingsfactoren (gemiddeld voor de boekjaren 1982-1986)

Variabele	Geschatte waarde bij		Verschil in %
	275 wegingsfactoren	32 wegingsfactoren	
Netto overschot	- 38.000	- 37.594	1,1
Loon bedrijfshoofd	64.899	65.342	0,7
Arbeidsopbrengst bedr.h.	26.899	27.747	3,2
Ondernemersinkomen	39.179	39.726	1,4
Tot.ber.loon gez.led.	27.867	27.036	- 3,0
Gezinsinkomen uit bdr	67.273	67.028	- 0,4
Tot. ink. buiten bedr	13.519	13.338	- 1,0
Totaal gezinsinkomen	80.792	80.416	- 0,5
Betaalde belastingen	8.083	7.888	- 2,4
Betaalde premies	8.188	8.216	0,3
Besteedbaar inkomen	64.521	64.313	- 0,3
Totale gezinsbest.	50.191	49.757	- 0,9
Afschrijvingen	23.736	23.783	0,2
Besparingen	14.331	14.556	1,6
Verm. subsidies	1.056	1.058	0,2
Boekwinsten	774	653	-15,7
Totaal eigen middelen	44.131	44.690	1,3
Totaal vreemde midd.	28.775	29.766	3,4
Totaal middelen	72.907	74.455	2,1

Tabel 11.11 (vervolg)

Variabele	Geschatte waarde bij		Verschil in %
	275 wegings- factoren	32 wegings- factoren	
Bruto inv. totaal	47.560	48.162	1,3
Aflossingen leningen	19.705	20.186	2,4
Mut. veldinv.+ voorr.	2.419	2.569	6,2
Grond in eigendom	408.364	406.788	- 0,4
Gebouwen in eigendom	177.963	177.101	- 0,5
Pachtersinvesteringen	11.887	12.110	1,9
Dode inventaris	88.462	88.726	0,3
Veestapel	140.304	140.839	0,4
Veldinventaris/voorr.	24.907	25.610	2,8
Totaal produktiemidd.	851.888	851.174	- 0,1
Totale liquiditeit	95.876	95.458	- 0,4
Tot. vorderingen lang	15.358	15.315	- 0,3
Totaal ov bezittingen	106.128	104.091	- 1,9
Balanstotaal	1.069.250	1.066.037	- 0,3
Lang vr. verm. totaal	252.687	255.923	1,3
Totaal kort vr. verm.	21.504	22.401	4,2
Eigen vermogen	795.059	787.714	- 0,9
Ha. cultuurgrond	22,88	23,05	0,7
Sbe (gem. boekjaar)	206	206	0,3
Sbe (meitelling)	204	205	0,1
Leeftijd ondernemer	48	47	- 1,3
Melkproduktie (kg.)	180.102	180.522	0,2

klassen, leeftijdsklassen en regio's die ervoor zouden kunnen zorgen dat bij de werving van bedrijven bepaalde soorten bedrijven een grotere kans krijgen om in de steekproef terecht te komen, worden opgevangen door het principe van vervangende bedrijven. Verschillen in non-respons bij de uitwerking van boekhoudingen zouden roet in het eten kunnen gooien, maar deze verschillen vertonen maar een lichte samenhang met de drie criteria die de 32 basisstrata verdelen in 275 substrata (paragraaf 10.4).

Er resteren dan twee factoren die zouden kunnen zorgen voor de noodzaak om binnen een basisstratum verder te stratificeren. In de eerste plaats is er de al genoemde keuze van extra bedrijven in bepaalde regio's, ten behoeve van bepaalde regionale BUL-groepen. Deze bedrijven krijgen dus opzettelijk een grotere kans om in de steekproef terecht te komen. In de tweede plaats is toepassing van het principe van vervangende bedrijven die sterk lij-

ken op de "eerste keus"-bedrijven, niet altijd voor 100% succesvol. In sommige jaren lukt het bijvoorbeeld niet om voor 1 mei voor bepaalde leeftijds- of oppervlakte-klassen binnen een basisstratum de gewenste aantallen steekproefbedrijven tot deelname te bewegen (zie paragraaf 10.3.10 en 10.3.11). Stratificatie achteraf binnen de basisstrata met behulp van de betreffende criteria, zorgt er dan voor dat dit niet tot systematische afwijkingen in de uitkomsten leidt.

Dit alles leidt tot de conclusie dat het zinvol kan zijn om - indien daarvoor binnen een basisstratum de mogelijkheid bestaat - bij de stratificatie achteraf gebruik te maken van de criteria oppervlakte, leeftijd en regio. Daarbij zal de regio als stratificatie-factor meer prioriteit dienen te krijgen in basisstrata waar regionale aanvullingen worden gekozen. Van de substratificatie zal echter maar een beperkte invloed uitgaan op de representativiteit.

In tabel 11.11 wordt de exercitie die is uitgevoerd voor de BUL-bedrijven herhaald voor de FIP-bedrijven. Ook hier blijken de gemiddelden bij 32 en bij 275 wegingsfactoren weinig uiteen te lopen.

11.6 Conclusies

Vergelijking van de meetingsgegevens op basis van de steekproefbedrijven met die op basis van alle bedrijven in de populatie, leert dat de steekproef op de belangrijkste punten nauwelijks afwijkt van de populatie. Waar wel afwijkingen zijn te constateren, gaat het in vrijwel alle gevallen om relatief weinig voorkomende diersoorten en gewassen als mestkalveren, schapen, geiten, paarden, moederdieren, slachtkuikens, kalkoenen, tuinakergewassen, landbouwzaden en peulvruchten. In de meeste gevallen is de onder- of oververtegenwoordiging te verklaren uit verschillen in respons bij de bedrijfskeuze of bij de uitwerking van boekhoudingen. Bij de opzet van de steekproef en bij de berekening van steekproefuitkomsten door middel van poststratificatie wordt wel geprobeerd om zoveel mogelijk de gevolgen van selectieve non-respons te ondervangen, maar het is onmogelijk om daarbij op alle factoren te letten.

Vergelijking van steekproefuitkomsten met gegevens uit externe bronnen leidt tot de conclusie dat er op enkele punten wel lichte verschillen zijn - bijvoorbeeld dat er op LEI-bedrijven een grotere deelname aan de melkcontrole is - maar dat daaruit moeilijk valt af te leiden in hoeverre er verschillen in bedrijfsuitkomsten zijn tussen de steekproef en de populatie.

De kilogram-opbrengsten per hectare gewas en de melkproductie per koe liggen in de LEI-steekproef in het algemeen slechts enkele procenten hoger dan in andere bronnen. Helaas is moeilijk te achterhalen of er op LEI-bedrijven ook hogere kosten worden

gemaakt. Een vertaling van verschillen in termen van bedrijfsuitkomsten is daardoor moeilijk te maken. Uit de vergelijking van meetellingsvariabelen die te maken hebben met de bedrijfsvoering kan wel worden afgeleid hoe het zit met de verkaveling op steekproefbedrijven, of met de berekening, of met de aanwezigheid van beregeningsmogelijkheden, maar die informatie is te fragmentarisch. Een onderzoek van verschillen in mechanisatie op basis van de door het CBS in 1979 en 1985 uitgevoerde mechanisatie-tellingen (Landbouwcijfers, 1987) kan hier wellicht in de toekomst meer licht op werpen.

De algemene conclusie van dit hoofdstuk kan zijn dat er sterke aanwijzingen zijn dat de bedrijfsuitkomsten op basis van het boekhoudnet een grote mate van representativiteit bezitten. Representativiteit is echter maar één kant van de medaille. De andere kant wordt gevormd door de betrouwbaarheid van de steekproefuitkomsten. De volgende hoofdstukken gaan daar op in.

12. Betrouwbaarheid op nationaal niveau

12.1 Inleiding

In hoofdstuk 4 is aangegeven dat een zo groot mogelijke betrouwbaarheid van steekproefuitkomsten voor de nationale boerderij, de eerste doelstelling van het boekhoudnet is. In dit hoofdstuk wordt ingegaan op die betrouwbaarheid van nationale cijfers. Ook wordt aangegeven in welke mate de stratificatie de betrouwbaarheid van deze cijfers beïnvloedt. Onderzocht is of de optimale stratificatie en allocatie bij de opzet van de steekproef inderdaad leidt tot de maximaal haalbare reductie in standaardfouten.

12.2 De nationale boerderij

De steekproefpopulatie wordt gevormd door alle landbouwbedrijven die groter zijn dan een bepaalde minimumomvang en kleiner zijn dan 2.000 sbe. Deze bedrijven samen worden de nationale boerderij genoemd. Op deze "boerderij" zijn ruim twee miljoen melkkoepen aanwezig, wordt ruim een half miljoen hectare cultuurgrond bebouwd met marktbaar gewassen en zijn ruim 100.000 volwaardige arbeidskrachten aan het werk (Bedrijfsuitkomsten, 1988:20).

Naast de nationale boerderij is er de gemiddelde nationale boerderij. Deze - alleen boekhoudkundig bestaande - boerderij wordt verkregen door de gegevens van alle populatiebedrijven te middelen. Op deze "boerderij" zijn bijvoorbeeld 33 melkkoepen, 17 fokzeugen en 68 mestvarkens aanwezig. Van de bijna 25 hectare cultuurgrond wordt ruim een derde deel bebouwd met marktbaar gewassen. Er is 1,5 volwaardige arbeidskracht aanwezig. De bedrijfsuitkomsten van de gemiddelde nationale boerderij worden berekend door de aggregaten van de nationale boerderij te delen door het totaal aantal individuele bedrijven in de populatie (Bedrijfsuitkomsten, 1988:23-25).

Naast de op bovengenoemde wijze gedefinieerde nationale boerderij, wordt in LEI-publikaties gewerkt met een aggregaat van bedrijven dat een deelverzameling vormt van de nationale boerderij, en waarvoor eveneens de naam nationale boerderij wordt gebruikt (Aukema en Overgaw, 1988; Poppe, 1988). Deze "boerderij" - waarvan een groot deel van de NV's, BV's en maatschappen is uitgezonderd - komt in paragraaf 12.6 aan de orde bij de bespreking van de betrouwbaarheid van steekproefuitkomsten in de financiële sfeer.

12.3 Betrouwbaarheid

In dit hoofdstuk komt de statistische betrouwbaarheid van de steekproefuitkomsten aan de orde. Daarbij gaat het om een kwantificering van de onzekerheid die nu eenmaal verbonden is aan een steekproef. Met behulp van een dergelijke kwantificering kunnen uitspraken gedaan worden over de kans dat toevallige afwijkingen tussen steekproef en populatie niet groter zijn dan een bepaalde waarde (paragraaf 6.2). Aan eventuele systematische afwijkingen (vertekeningen) tussen steekproef en populatie wordt hier voorbijgegaan. In hoofdstuk 11 is daaraan aandacht besteed. Dat wil niet zeggen dat systematische en toevallige afwijkingen niet in samenhang moeten worden bekeken.

Bij schattingen met vertekeningen is de nauwkeurigheid van een schatting - gemeten door de gemiddelde kwadratische afwijking - een betere maatstaf dan de precisie, die gemeten wordt door de standaardfout. De gemiddelde kwadratische fout van een schatting is gelijk aan de som van de variantie en het kwadraat van de vertekening van de schatting. Bij de constructie van betrouwbaarheidsintervallen zou de variantie dus vervangen moeten worden door de gemiddelde kwadratische fout. Deze is echter in het algemeen onbekend. Uit hoofdstuk 11 kunnen wel aanwijzingen verkregen worden over de grootte van de vertekeningen van sommige variabelen, zodat betrouwbaarheidsintervallen eventueel aangepast kunnen worden aan de vertekeningen. Voor een verdere behandeling van dit onderwerp, zie Bethlehem en Kersten (1986). Daarin wordt aangetoond dat de verhouding tussen de vertekening en de standaardfout bepaalt in hoeverre betrouwbaarheidsintervallen op basis van standaardfouten bruikbaar zijn in geval van vertekening.

12.4 De doelvariabelen

In een steekproef zullen altijd toevallige afwijkingen ten opzichte van de populatie kunnen voorkomen. Het doel van de steekproefopzet is om de kans op grote toevallige fouten in de geschatte waarden van variabelen te reduceren. Als doelvariabelen voor de LEI-steekproef zijn de normatieve toegevoegde waarde per bedrijf en de arbeidsopbrengst van de ondernemer gekozen (paragraaf 4.3). In tabel 12.1 is weergegeven hoe groot de standaardfout van deze grootheden was voor de boekjaren 1979-1986. De standaardfouten hebben betrekking op de gemiddelde nationale boerderij. Ze zijn berekend met behulp van vergelijking (6.4.5). De standaardfouten van de totale nationale boerderij zijn een factor N (ongeveer 70.000) groter (zie vergelijking (6.4.6)).

Uit deze tabel kan geconcludeerd worden dat de sbe-schatting een standaardfout van ongeveer 1,5 sbe heeft. Ten opzichte van een niet-gestraficeerde steekproef is de standaardfout 50 à 60% gereduceerd. De stratificatie naar "sbe in metelling" draagt

Tabel 12.1 Grootte van de standaardfouten van de doelvariabelen voor de gemiddelde nationale boerderij naar boekjaar

Boekjaar	Sbe (gemiddeld in het boekjaar) *		Arbeidsopbrengst van de ondernemer(s) **)		
	gem.	standaardfout	gem.	standaardfout	
1979	192	2,0 (3,8)			
1980	196	1,6 (3,3)			
1981	203	1,5 (3,4)			
1982	209	1,5 (3,6)	32.940	1.656	(1.667)
1983	211	1,5 (3,7)	30.774	1.638	(2.117)
1984	216	1,7 (3,9)	23.903	1.618	(1.767)
1985	218	1,6 (3,9)	24.566	1.554	(1.759)
1986	218	1,5 (4,0)	22.313	1.573	(1.619)

N.B.: Tussen haakjes staat de standaardfout die gerealiseerd zou zijn bij een enkelvoudige aselechte steekproef en eenzelfde aantal waarnemingen.

- *) De standaardbedrijfseenheid (sbe) is - op een schaalfactor na - gelijk aan de normatieve toegevoegde waarde. De schatting heeft betrekking op de BUL-bedrijven.
- ***) Dit begrip wordt gebruikt in de FIP. De schatting heeft betrekking op de FIP-bedrijven. In de BUL wordt gewerkt met de "arbeidsopbrengst per ondernemer" (BUL-gebieden) en met de "arbeidsopbrengst per bedrijf" (nationale boerderij).

daar het grootste deel aan bij 1). De reductie in de standaardfout van de geschatte gemiddelde arbeidsopbrengst van de ondernemer(s) ligt bij de gekozen stratificatie op ongeveer 10%. Daarvoor is opnieuw de stratificatie-variabele "sbe in metelling" verantwoordelijk, al draagt ook de stratificatie naar bedrijfstype iets bij aan de reductie.

De standaardfout van de arbeidsopbrengst per bedrijf is in het algemeen kleiner dan die van de arbeidsopbrengst van de ondernemer(s). De reductie door de stratificatie is daarbij ongeveer 30% (zie tabel 12.2).

- 1) De standaardfout van de schatting van de gemiddelde waarde van het aantal sbe in de metelling wordt uiteraard met een nog groter percentage verminderd, namelijk met 65 à 70%. Deze standaardfout bedraagt 1,2 sbe.

12.5 Bedrijfseconomische grootheden

12.5.1 Standaardfouten

Veel schattingen in de bedrijfsgruotte-, kosten- en opbrengsten-sfeer laten overeenkomstige reducties in standaardfouten

Tabel 12.2 Steekproefschattingen voor het gemiddelde en de standaardfout voor een aantal variabelen voor de gemiddelde nationale (BUL-)boerderij (gemiddelde van de boekjaren 1979-1986)

Variabele	Gemiddelde	Standaardfout van het gemiddelde		Reductie door stratificatie in %
		abso- luut	in %	
Totale kosten	343.889	4.029	1,2	41
Totale opbrengsten	308.538	4.139	1,3	43
Faktorkosten	130.578	968	0,7	39
Faktoropbrengsten	95.226	1.295	1,4	42
Netto bedr. resultaat	- 35.351	1.301	3,7	14
Arbeidsopbrengst	55.988	1.183	2,1	31
Arbeidkosten	91.340	830	0,9	19
Rente	16.419	156	1,0	50
Netto pacht	18.623	254	1,4	40
Werk door derden	4.149	92	2,2	37
Nonfaktorkosten tot.	213.311	3.572	1,7	39
Sbe totaal (boekjaar)	208	2	0,8	56
Sbe totaal (mei)	205	1	0,6	67
Opp. cultuurgrond (ha)	23,44	0,20	0,8	65
- blijvend grasland	13,82	0,17	1,2	58
- marktbaar gewassen	8,10	0,15	1,8	74
- kunstweide en vg.	1,48	0,08	5,7	18
- boomgaard en uitl.	0,04	0,01	29,5	- 3
Volwaardige arbeidskr.	1,5	0,01	0,9	17
Leeftijd ondernemer	48	0,3	0,6	17
Aantal melkkoeien	32,4	0,4	1,1	64
Overig weidevee (oke)	12,7	0,3	2,4	30
Aantal fokzeugen	152	8	5,3	41
Aantal mestvarkens	621	54	8,6	11
Aantal leghennen	481	65	13,6	44
Aantal slachtkuikens	378	103	27,2	- 3

zien als die voor het aantal sbe en de arbeidsopbrengst. Voor een aantal variabelen is in tabel 12.2 weergegeven wat - gemiddeld gedurende de boekjaren 1979-1986 - de schatting van het gemiddelde en de daarbij behorende standaardfout was. In de laatste kolom staat de reductie in de standaardfout ten opzichte van de standaardfout die hoort bij een enkelvoudige steekproef. Een overzicht van standaardfouten van quotiënten wordt gegeven in paragraaf 12.7.

Uit de tabel wordt duidelijk dat de standaardfouten van de kosten-, opbrengsten- en inkomens-grootheden slechts enkele procenten van de schattingen bedragen. Dat is het resultaat van een vrij groot aantal steekproefbedrijven - ruim 1.000 per jaar - en van de toepassing van stratificatie bij de steekproef. In de laatste kolom is aangegeven hoeveel de standaardfouten zijn verminderd ten opzichte van toepassing van een enkelvoudige steekproef - bij gelijke aantallen bedrijven. De grootste reducties worden uiteraard bereikt voor grootheden die nauw samenhangen met de stratificatie-variabelen. De standaardfout van het aantal melkkoeien bijvoorbeeld, zou ongeveer driemaal zo groot zijn bij een enkelvoudige steekproef. Voor veel grootheden in de opbrengsten- en kostensfeer wordt de standaardfout bijna gehalveerd. De reductie is vrij klein bij het netto bedrijfsresultaat (netto overschot). De samenhang van deze variabele met de stratificatie-variabelen is kennelijk minder groot.

Voor sommige grootheden levert de steekproef ondanks de stratificatie wat minder betrouwbare uitkomsten. De schatting van het aantal slachtkuikens per bedrijf komt op gemiddeld 378 met een standaardfout van 103. Dat wil zeggen dat de kans dat het gemiddelde aantal slachtkuikens per bedrijf in de populatie kleiner is dan 275 of groter is dan 480 ruim 30% is. De oorzaak van deze hoge standaardfout is de grote spreiding in het aantal slachtkuikens per bedrijf. Veel bedrijven hebben geen slachtkuikens en andere hebben grote aantallen. Bovendien zijn deze bedrijven over een groot aantal strata verspreid, waardoor de stratificatie de standaardfout niet gunstig beïnvloedt.

12.5.2 De invloed van de stratificatie

In hoofdstuk 5 kwam naar voren dat de populatie gestratificeerd wordt aan de hand van vijf variabelen. Eerst wordt gestratificeerd naar bedrijfstype en sbe-klasse. Daarna wordt binnen de zo ontstane basisstrata gestratificeerd naar achtereenvolgens bedrijfsoppervlakte, leeftijd van de ondernemer en regio. Eén van de redenen om te stratificeren is het reduceren van standaardfouten van schattingen. In deze subparagraaf wordt ingegaan op de vraag wat elk van de afzonderlijke stratificatie-variabelen bijdraagt aan de reductie.

Tabel 12.3 geeft aan met welke percentages de standaardfouten van een aantal schattingen in 1986 zouden zijn afgenomen indien een bepaalde vorm van stratificatie zou zijn toegepast.

Achtereenvolgens gaat het daarbij om een stratificatie naar bedrijfstype (8 strata), naar bedrijfstype en sbe-klasse (32 strata) en de werkelijk toegepaste stratificatie (275 strata).

De stratificatie naar bedrijfstype levert voor een beperkt aantal variabelen een reductie van de standaardfout op. De opper-

Tabel 12.3 Percentage waarmee de standaardfout van een schatting op basis van een enkelvoudige steekproef gereduceerd wordt door stratificatie, naar de mate van stratificatie (boekjaar 1986)

Variabele	Stratificatie naar:		
	bedrijfs- type	bedrijfs- type/ sbe- klasse	bedrijfstype/ sbe-klasse/ oppervlakte/ leeftijd/ regio
Totale kosten	5	47	49
Totale opbrengsten	4	48	50
Factorkosten	1	47	45
Factoropbrengsten	1	46	46
Netto bedr. resultaat	3	14	11
Arbeidsopbrengst	1	33	33
Arbeidskosten	3	29	23
Rente	-1	50	52
Netto pacht	-1	40	39
Werk door derden	26	43	41
Non-factorkosten tot.	8	43	45
Sbe totaal (boekjaar)	2	60	62
Sbe totaal (mei)	1	68	68
Opp. cultuurgrond (ha)	23	57	67
- blijvend grasland	26	49	62
- marktbaar gewassen	48	70	75
- kunstweide en vg.	-1	21	27
- boomgaard en uitl.	-4	-17	-60
Volwaardige arbeidskr.	1	25	21
Leeftijd ondernemer	-1	- 9	22
Aantal melkkoeien	23	64	64
Overig weidevee (oke)	7	28	23
Aantal fokzeugen	25	48	43
Aantal mestvarkens	6	13	12
Aantal leghennen	27	49	50
Aantal slachtkuikens	3	10	1

vlakke cultuurgrond, het aantal melkkoeien, het aantal fokzeugen en het aantal leghennen per bedrijf zijn daar voorbeelden van.

De 32 basisstrata geven voor de meeste variabelen vrij forse verminderingen van de standaardfout. Alleen variabelen waarbij het verband met het bedrijfstype en de bedrijfsgrootte nauwelijks aanwezig is, profiteren niet van deze stratificatie. Het gaat daarbij om bijvoorbeeld de leeftijd van de ondernemer, de oppervlakte boomgaard en uitloop, het aantal mestvarkens en het aantal slachtkuikens.

Opvallend is dat de stratificatie die verder gaat dan die naar 32 basisstrata voor veel variabelen nauwelijks winst geeft ten aanzien van de standaardfout. Voor sommige variabelen schiet de stratificatie te ver door waardoor een grotere standaardfout wordt gerealiseerd dan bij 32 basisstrata. Variabelen waarvoor de verdergaande stratificatie wel tot betere resultaten leidt, hebben een vrij sterke samenhang met de stratificatie-variabelen die daarbij gebruikt worden. Het gaat om de oppervlakte cultuurgrond en de uitsplitsing daarvan over de verschillende aanwendingsvormen en de leeftijd van de ondernemer.

Het geheel overziende kan geconcludeerd worden dat vanuit het oogpunt van de betrouwbaarheid van schattingen voor de nationale boerderij, de stratificatie naar oppervlakte, leeftijd en regio weinig toevoegt aan de basis-stratificatie.

12.6 Financiële grootheden

12.6.1 Standaardfouten

Voor een aantal financierings-variabelen is in tabel 12.4 weergegeven wat - gemiddeld gedurende de boekjaren 1982-1986 - de schatting van het gemiddelde en de daarbij behorende standaardfout was. In de laatste kolom staat de reductie in de standaardfout ten opzichte van de standaardfout die hoort bij een ongestratificeerde steekproef. In bijlage 4 wordt deze tabel met een groot aantal variabelen uitgebreid. Een overzicht van standaardfouten van quotiënten is te vinden in paragraaf 12.7.

De relatieve standaardfouten zijn in het algemeen iets hoger dan bij de bedrijfsuitkomsten. De reducties door de stratificatie zijn ook minder groot. Toch is de standaardfout in veel gevallen slechts enkele procenten van het geschatte gemiddelde, vooral voor de belangrijkste variabelen. Bij belastingen, besparingen en investeringen is de relatieve standaardfout iets groter. Verder ligt het voor de hand dat een variabele als "erfenissen" een grote spreiding heeft die min of meer losstaat van de stratificatie. Dat leidt tot een hoge relatieve standaardfout.

Tabel 12.4 Steekproefschattingen voor het gemiddelde en de standaardfout voor een aantal variabelen voor de gemiddelde nationale (FIP-)boerderij (gemiddelde van de boekjaren 1982-1986)

Variabele	Gemiddelde	Standaardfout gemiddelde		Reductie door stratificatie in %
		abso- luut	in %	
Netto bedr. resultaat	- 38.000	1.558	4,1	10
Loon bedrijfshoofd	64.899	686	1,1	0
Arbeidsopbrengst bedr.h.	26.899	1.608	6,0	10
Ondernemersinkomen	39.179	1.312	3,0	13
Tot.ber.loon gez.leden	27.867	952	3,4	5
Gezinsinkomen uit bedr	67.273	1.492	2,2	22
Tot. ink. buiten bedr.	13.519	448	3,3	- 8
Totaal gezinsinkomen	80.792	1.574	2,0	20
Betaalde belastingen	8.083	534	6,6	16
Betaalde premies	8.188	300	3,7	- 4
Besteedbaar inkomen	64.521	1.411	2,2	17
Totale gezinsbest.	50.191	733	1,5	6
Afschrijvingen	23.736	340	1,4	38
Besparingen	14.331	1.397	9,8	10
Erfenissen	4.314	794	18,4	- 4
Boekwinsten	774	268	34,6	- 17
Totaal eigen middelen	44.131	1.690	3,8	13
Totaal vreemde midd.	28.775	3.434	11,1	2
Totaal middelen	72.907	3.898	5,4	7
Bruto inv. totaal	47.560	3.511	7,4	1
Grond in eigendom	408.364	10.336	2,5	19
Gebouwen in eigendom	177.964	4.024	2,3	23
Veestapel	140.304	1.866	1,3	50
Totaal produktiemidd.	851.888	12.571	1,5	32
Totale liquiditeit	95.876	3.428	3,6	5
Balanstotaal	1.069.250	14.328	1,3	32
Totaal vreemd verm.	274.191	7.194	2,6	24
Eigen vermogen	795.059	13.652	1,7	20
Ha. cultuurgrond	22,88	0,19	0,8	64
Sbe (gem. boekjaar)	206	1,5	0,7	58
Sbe (meitelling)	204	1,1	0,6	67
Leeftijd bedrijfshoofd	48	0,3	0,6	20

Tabel 12.5 Percentage waarmee de standaardfout van een schatting op basis van een enkelvoudige steekproef gereduceerd wordt door stratificatie, naar de mate van stratificatie (nationale FIP-boerderij, boekjaar 1986)

Variabele	Stratificatie naar:		
	bedrijfs- type	bedrijfs- type/sbe- klasse	bedrijfstype/ sbe-kl./opp. leeftijd/regio
Netto bedr. resultaat	3	11	- 3
Loon bedrijfshoofd	4	- 3	2
Arbeidsopbrengst bedr.h.	1	9	3
Ondernemersinkomen	0	11	2
Tot.ber.loon gez.leden	0	10	4
Gezinsinkomen uit bedr	1	24	16
Tot. ink. buiten bedr.	2	- 5	- 2
Totaal gezinsinkomen	1	23	15
Betaalde belastingen	0	23	21
Betaalde premies	0	8	- 3
Besteedbaar inkomen	2	18	10
Totale gezinsbest.	- 1	12	11
Afschrijvingen	1	40	38
Besparingen	2	12	1
Erfenissen	5	- 5	0
Boekwinsten	- 4	- 5	-43
Totaal eigen middelen	2	17	7
Totaal vreemde midd.	2	10	5
Totaal middelen	2	14	5
Bruto inv. totaal	0	9	2
Grond in eigendom	7	22	21
Gebouwen in eigendom	5	25	22
Veestapel	13	54	50
Totaal produktiemidd.	2	34	32
Totale liquiditeit	5	3	- 7
Balanstotaal	1	33	31
Totaal vreemd verm.	1	28	28
Eigen vermogen	3	20	17
Ha. cultuurgrond	24	55	63
Sbe (gem. boekjaar)	1	61	62
Sbe (meitelling)	1	70	69
Leeftijd bedrijfshoofd	0	- 5	24

12.6.2 De invloed van de stratificatie

Tabel 12.5 geeft aan met welke percentages de standaardfouten van een aantal schattingen in 1986 zouden zijn afgenomen in-

dien een bepaalde vorm van stratificatie zou zijn toegepast. Evenals in tabel 12.3 gaat het daarbij achtereenvolgens om een stratificatie naar bedrijfstype (8 strata), naar bedrijfstype en sbe-klasse (32 strata) en de werkelijk toegepaste stratificatie (275 strata).

De conclusies die getrokken werden ten aanzien van de bedrijfseconomische variabelen, zijn in grote lijnen ook van toepassing op de financiële variabelen. De meest gedetailleerde stratificatie levert in de meeste gevallen grotere standaardfouten dan de basis-stratificatie. Hier geldt nog sterker dat door de verdergaande stratificatie geen reductie in standaardfouten optreedt. Opvallend is dat ook voor variabelen waar een samenhang met de leeftijd van de ondernemer voor de hand ligt, de 275 strata niet beter scoren dan de 32 strata. Dat kan een aanwijzing zijn dat de indeling in leeftijdsklassen (jonger dan 35 jaar, 35 tot 60 jaar, ouder dan 60 jaar) niet helemaal optimaal is, of dat de leeftijds-variabele te weinig aan bod komt in de stratificatie. Nader onderzoek op dit punt zal dit moeten uitwijzen.

De reducties in de standaardfouten liggen in de orde van grootte van 10% tot 20%.

12.7 Quotiënten

Quotiënten die een verhoudingsgetal weergeven voor een totale populatie (bijvoorbeeld kg melk per koe) kunnen op twee manieren berekend worden. De ene manier is om per bedrijf de waarde van het quotiënt te berekenen waarna het quotiënt voor de gehele populatie berekend wordt als het gemiddelde van de individuele quotiënten ("mean per unit estimate"). De andere berekeningswijze komt er op neer dat voor alle bedrijven de waarde van de grootte in de teller wordt opgeteld en gedeeld wordt door de som van de waarden in de noemer ("ratio estimate"). Deze laatste berekening van quotiënten wordt ook toegepast voor veel kengetallen uit het boekhoudnet (Bedrijfsuitkomsten, 1988:36-37).

Te bewijzen valt (Cochran, 1977:155) dat de steekproefvariantie van de schatting van een totaal-quotient \hat{R} ($=\sum y / \sum x = y/\bar{x}$) op basis van een enkelvoudige aselechte steekproef bij benadering gelijk is aan:

$$\text{Var}(\hat{R}) \approx \frac{1 - (n/N)}{n \cdot (\bar{X}^2)} * \sum_{i=1}^N (y_i - R \cdot x_i)^2 / (N-1) \quad (12.5.1)$$

Deze quotiëntschatter is geen zuivere schatter, maar bij grote steekproeven is de onzuiverheid te verwaarlozen; deze is van de orde $(1/n)$.

Voor gestratificeerde steekproeven met aselechte trekking binnen de strata, is de variantie van de quotiëntschatter $\hat{R}_{st} =$

$(\bar{y}_{st}/\bar{x}_{st})$ te berekenen als (analoog aan 12.5.1):

$$\text{var}(\hat{R}_{st}) = \frac{1 - (n/N)}{n * \bar{x}_{st}^2} * \sum_{i=1}^n \frac{W_i * (y_i - \hat{R}_{st} * x_i)^2}{(W_i/n) * (n-1)} \quad (12.5.2)$$

waarbij

\hat{R} = de steekproefschatting van R
 \bar{x}_{st} = de steekproefschatting van X

In deze berekening wordt weliswaar rekening gehouden met de wegingsfactoren W_i , maar niet expliciet met de indeling in strata. Vergelijking 12.5.3 houdt daar wel rekening mee. De betreffende quotiëntschatting wordt de "combined ratio estimator" (\hat{R}_C) genoemd (Cochran, 1977:166).

$$\hat{R}_C = \frac{\bar{y}_{st}}{\bar{x}_{st}}$$

De variantie van \hat{R}_C kan geschat worden als:

$$\text{var}(\hat{R}_C) = \sum_{h=1}^H \frac{N_h * (N_h - n_h)}{n_h * \bar{x}_{st}^2 * (n_h - 1)} * \sum_{i=1}^{n_h} [(y_{hi} - \bar{y}_h) - \hat{R}_C * (x_{hi} - \bar{x}_h)]^2 \quad (12.5.3)$$

Als benadering voor de standaardfout van een totaal-quotiënt wordt in de praktijk wel de standaardfout gebruikt van de grootheid die berekend wordt door *per bedrijf* het quotiënt te bepalen waarna het gemiddelde van al deze quotiënten ("mean per unit") de uitkomst bepaalt (vergelijking 6.4.7). Deze berekening kan ook worden uitgevoerd door rekening te houden met de stratificatie (vergelijking 6.4.5).

In de volgende tabel worden deze laatste berekeningswijzen - die op zich erg aantrekkelijk zijn omdat in principe dezelfde berekening als bij absolute grootheden wordt uitgevoerd - vergeleken met een berekening volgens vergelijking 12.5.2 en vergelijking 12.5.3. Gekozen is voor een aantal quotiënten waarover jaarlijks in de BUL en de FIP wordt gerapporteerd.

Toepassing van de juiste formule voor de berekening van de standaardfout (laatste kolom), blijkt in de meeste gevallen te leiden tot standaardfouten die veel kleiner zijn de standaardfouten die berekend worden met de formules die bedoeld zijn voor absolute grootheden (kolommen 3 en 4).

Ook het al dan niet rekening houden met het feit dat het gaat om een gestratificeerde steekproef, blijkt de uitkomsten sterk te beïnvloeden. Bij vergelijking van de laatste twee kolom-

men blijkt dat door de stratificatie de standaardfout van een quotiënt in een aantal gevallen met meer dan 50% verminderd wordt ten opzichte van een enkelvoudige steekproef.

Een conclusie uit tabel 12.6 kan zijn dat bij schattingen voor quotiënten een algemene toepassing van formules van standaardfouten voor absolute grootheden weinig zin heeft. Het hangt van de variatie in de teller en de noemer af of deze toepassing een goede benadering geeft voor de standaardfout.

Tabel 12.6 Overzicht van standaardfouten van geschatte quotiënten naar berekeningsmethode van de standaardfout (boekjaar 1986)

Variabele	Quotiënt schatting (\hat{R}_C)	Berekeningsbasis			
		gemiddelde van bedrijfsquotiënten		totaal- quotiënt	
		rekening houdend met stratum- structuur? (+formule)			
		nee	ja	nee	ja
*)		6.4.7	6.4.5	12.5.2	12.5.3
BUL:					
- % marktbaar gewassen	22	1,13	0,29	1,43	0,30
- % rundvee + voedergew.	57	1,36	0,41	1,61	0,43
- % int. veehouderij(i.v.)	20	1,03	0,37	1,22	0,35
- sbe / ha	8,8	146	277	1,6	0,6
- sbe (excl.i.v.)/ ha	7,0	0,9	1,0	0,8	0,5
- sbe mg. / ha mg.	5,4	1,5	1,6	1,0	0,5
- sbe rv. / ha vg.	7,8	1,3	0,9	0,8	0,6
- opbr./f 100,- kosten	88	0,47	0,43	0,52	0,36
- arb.opbr./f 100,- k.	57	1,49	1,40	1,67	1,17
- arbeidsopbr. /vak	38.295	936	900	1.032	781
- stikstof /ha	271	4,3	3,4	4,7	2,9
- fosfaat /ha	49	1,2	1,1	1,6	1,2
- kali /ha	57	1,8	1,5	2,8	1,6
- waarde werktuigen/ha	9.434	7.261	10.019	189	113
- arbeidskosten / ha	4.123	7.930	15.093	96	43
- werktuigkosten / ha	1.601	920	1.352	28	18
- werk door derden/ha	440	416	698	11	8
- bewerkingskosten/ha	6.035	8.742	16.413	113	52

Tabel 12.6 (1e vervolg)

Variabele	Quotiënt schatting (\hat{R}_c)	Berekeningsbasis			
		gemiddelde van bedrijfsquotiënten			totaal- quotiënt
		6.4.7	6.4.5	12.5.2	12.5.3
- pacht / ha	1.605	2.246	4.201	34	22
- kosten grond en gebouwen / ha	1.703	2.500	4.695	36	23
- overige niet-toege- rekende kosten / ha	671	1.296	2.475	14	8
- totaal niet-toege- rekende kosten / ha	8.409	12.501	23.539	153	67
- opbrengsten / ha	13.448	55.156	106.664	378	166
- toeger. kosten/ha	6.813	44.641	86.425	276	129
- saldo / ha	6.635	11.042	20.927	118	62
- netto bedr.res./ha	- 1.774	1.920	3.060	89	57
- bewerkingseenh./vak	32.580	535	471	473	319
- opbr. / koe	5.500	522	599	45	38
- kosten / koe	1.973	618	715	31	27
- saldo / koe	3.527	127	131	31	26
- opbr. / ha v.opp	10.940	220	172	162	125
- kosten / ha v.opp	4.585	136	130	93	74
- saldo / ha v.opp	6.355	113	73	84	65
- opbrengsten intens. veeh. / sbe i.v.	2.363	49	51	67	67
- kosten (idem)	1.753	41	45	60	61
- saldo (idem)	610	13	13	19	20
- opbr. mg./ ha mg.	6.214	113	97	126	68
- kosten (idem)	1.870	32	27	50	28
- saldo (idem)	4.344	89	80	98	59
- arbeidskosten / sbe	469	7,0	6,2	6,6	4,2
- rente /sbe	78	0,6	0,7	0,6	0,5
- netto pacht /sbe	108	1,4	1,7	1,3	1,3
- werk door derden/sbe	23	0,6	0,5	0,7	0,4
- factorkosten / sbe	678	7,0	6,5	6,9	4,4
- non-factorkosten/sbe	1.052	18,2	19,5	19,0	13,0
- totale kosten / sbe	1.730	19,8	21,4	20,3	14,3
- opbrengsten / sbe	1.528	20,1	23,0	19,8	14,6
- netto bedr.res. /sbe	- 202	8,8	8,5	9,3	6,3
- arbeidsopbrengst/sbe	267	5,7	6,7	6,0	5,1
- factoropbrengst /sbe	476	5,9	7,1	5,9	5,2

Tabel 12.6 (2e vervolg)

Variabele	Quotiënt schatting (\hat{R}_c)	Berekeningsbasis			
		gemiddelde van bedrijfsquotiënten		totaal- quotiënt	
		6.4.7	6.4.5	12.5.2	12.5.3
FIP:					
(Percentages)					
- besparingen/ gezinsbestedingen	23,0	58,0	55,0	1,8	1,8
- eigen vermogen/ totaal vermogen	73,3	0,7	0,7	0,7	0,6
- opbrengst vermogen/ gemiddeld vermogen	- 0,3	0,3	0,3	0,2	0,1
- opbrengst eig.verm./ gem. eigen vermogen	- 3,0	1,5	1,9	0,2	0,2

*) Voor een omschrijving van de gebruikte kengetallen, zie Bedrijfsuitkomsten, 1988, tabel 2.2 en paragraaf 3.4, en Aukema en Overgaw, 1988.

12.9 Evaluatie van het allocatie-mechanisme

12.9.1 Inleiding

Met het oog op een zo groot mogelijke betrouwbaarheid van de schattingen van bepaalde grootheden, vindt er bij de opzet van de steekproef een optimale stratificatie plaats, gevolgd door een optimale allocatie van de steekproefbedrijven over de strata (zie paragraaf 6.5 en 6.6). In grote lijnen komt dat er op neer dat de bedrijven in de metelling worden ingedeeld naar acht bedrijfstypen en dat vervolgens binnen ieder type vier sbe-klassen worden gevormd op een zodanige manier dat per bedrijfstype geldt dat aan een bepaalde voorwaarde is voldaan. Die voorwaarde is dat het produkt van het aantal bedrijven in een sbe-klasse en de standaarddeviatie van de variabele $(sbe)^2$ in die sbe-klasse voor alle vier sbe-klassen gelijk is. De allocatie die bij deze optimale stratificatie hoort, verdeelt de circa 1.100 steekproefbedrijven evenredig aan deze standaarddeviatie over de gevormde 32 klassen. Gegeven de genoemde voorwaarde houdt dit in dat per bedrijfstype in elk van de vier sbe-klassen een gelijk aantal bedrijven gekozen wordt.

Het idee achter deze wijze van verdelen van de steekproefbedrijven over delen van de populatie is dat de standaarddeviatie van de variabele (sbe)^{*} een goede maat is voor de standaarddeviaties van de doelvariabelen (zie paragraaf 6.5.2). Voor de jaren waarvoor steekproefuitkomsten beschikbaar zijn, is nagegaan hoe geschikt de gekozen maat is geweest. In de volgende sub-paragraaf wordt met een voorbeeld aangegeven hoe dat onderzocht kan worden. Sub-paragraaf 12.9.3 geeft een evaluatie voor de gehele populatie op basis van informatie over meerdere boekjaren.

12.9.2 Een voorbeeld

In tabel 12.7 wordt weergegeven wat de standaarddeviaties van de hulpvariabele en de doelvariabelen waren in de vier sbe-klassen binnen het bedrijfstype "rundveehouderij" in boekjaar 1984. De tweede kolom in de tabel geeft de standaarddeviaties van de hulpvariabele (sbe)^{*}. Deze hebben betrekking op de metelling van 1982 die gebruikt is bij de keuze van bedrijven voor boekjaar 1984. De verhouding in steekproefpercentages tussen de vier sbe-klassen wordt bepaald door de verhouding in de standaarddeviaties van de hulpvariabele (zie paragraaf 6.5.2). De steekproefpercentages in de rundveehouderij verhieldden zich bij de keuze voor boekjaar 1984 dus als 9:10:13:35. De vraag is nu of de verhoudingen tussen de standaarddeviaties een goede afspiegeling vormen van de verhoudingen tussen de standaarddeviaties per sbe-klasse voor de *doelvariabelen* en of ze dus als leidraad kunnen dienen bij de allocatie.

Voor de doelvariabelen genormaliseerde toegevoegde waarde (aantal sbe) en arbeidsopbrengst van de ondernemer zijn de verhoudingen in standaarddeviatie tussen de sbe-klassen volgens tabel 12.7 respectievelijk 26:33:42:140 en 27:38:45:67. Voor beide doelvariabelen komen de verhoudingen globaal overeen met de verhoudingen bij de hulpvariabele. Niveau-verschillen in de stan-

Tabel 12.7 Standaarddeviaties van de allocatie-variabele bij de bedrijfskeuze en de doelvariabelen van het boekhoudnet naar sbe-klasse (boekjaar 1984; rundveebedrijven)

Sbe- klasse	Metelling 1982 (sbe) [*]	Boekhoudnet 1984	
		sbe gem. tijdens boekjaar	arbeidsopbrengst van de ondernemer (*1.000)
1	9	26	27
2	10	33	38
3	13	42	45
4	35	140	67

daarddeviaties beletten echter een scherp zicht hierop. Een enigszins andere presentatie van deze cijfers in tabel 12.8 laat beter zien in welke mate de verhoudingen parallel lopen. In deze tabel worden de standaarddeviaties van de doelvariabelen gedeeld door de standaarddeviatie van de hulpvariabele.

Tabel 12.8 laat zien dat de standaarddeviatie van de doelvariabelen in verhouding tot de standaarddeviatie van de hulpvariabele niet in alle sbe-klassen gelijk is. Bij de doelvariabele "sbe" hebben de grotere bedrijven een hogere standaarddeviatie dan op grond van de verhoudingen bij de hulpvariabele verwacht zou mogen worden. Bij de arbeidsopbrengst van de ondernemer heeft de hoogste sbe-klasse daarentegen een relatief lage standaarddeviatie. "Gemiddeld" over de twee doelvariabelen is het quotiënt van de standaarddeviaties echter vrij constant.

In de volgende sub-paragraaf wordt nagegaan in hoeverre deze voorlopige conclusies ook opgaan voor andere bedrijfstypen en voor andere boekjaren.

Tabel 12.8 Quotiënt van de standaarddeviatie van de doelvariabelen en de allocatie-variabele naar sbe-klasse (boekjaar 1984; rundveebedrijven)

Sbe-klasse	Sbe gem. tijdens boekjaar (*0,1)	Arbeidsopbrengst van de ondernemer (*1.000)
1	29	30
2	33	38
3	32	35
4	40	19

12.9.3 De boekjaren 1981-1986

In bijlage 5 wordt een tabel gegeven waarin voor alle basisstrata de standaarddeviaties worden gegeven voor allocatie- en doelvariabelen. Het zijn gemiddelden over de boekjaren 1981-1986. In tabel 12.9 worden de standaarddeviaties uit bijlage 5 weergegeven in verhouding tot de standaarddeviatie van de hulpvariabele (sbe)². Het is een uitgebreide versie van tabel 10.8. Er zijn vier doelvariabelen in opgenomen.

De eerste conclusie die uit tabel 12.9 getrokken kan worden is dat de standaarddeviatie van (sbe)² in de metelling in het algemeen een goede maatstaf is voor de standaarddeviatie van de doelvariabelen. Per doelvariabele variëren de quotiënten immers maar weinig tussen de basisstrata. De standaarddeviatie van het aantal sbe is gemiddeld ongeveer 3,5 maal zo groot als de stan-

daarddeviatie van de hulpvariabele (kolom 2). Omdat de meetlingshulpvariabele "(sbe)³" is en niet "sbe" zou verwacht mogen worden dat de verhouding tussen de standaarddeviaties bij de laagste sbe-klassen lager zou zijn dan bij de hoogste sbe-klassen. Dat is echter nauwelijks het geval. De verklaring daarvoor is te vinden door vergelijking van de kolommen 3 en 4 in tabel B5.1. Uit deze tabel blijkt dat de variantie in (sbe)³ gedurende het boekjaar in veel gevallen aanmerkelijk groter is dan de variantie in (sbe)³ in de metelling (vergelijking kolom 3 en 4). Bij de (kleinere) intensieve bedrijven zit daar een factor twee tussen. De klassen met kleinere bedrijven laten bij alle bedrijfstypen de grootste verschillen zien 1).

Tabel 12.9 *Verhouding tussen de standaarddeviaties van doelvariabelen in de boekhouding en de hulpvariabele (sbe)³ in de metelling, naar basisstratum en doelvariabele (gemiddeld over de boekjaren 1981-1986)*

Type/ sbe- klasse	Variabele			
	sbe gem. tijdens boekjaar (*0,1)	arbeids- opbrengst ondernemers (*1000)	*) arbeids opbrengst per ondern. (*1000)	factor opbrengst (*1.000)
1- 1	33	29	29	32
2	43	44	47	51
3	43	37	43	47
4	43	19	23	27
2- 1	31	28	29	24
2	30	34	35	28
3	31	29	33	29
4	36	14	18	23
3- 1	33	30	30	28
2	32	37	38	41
3	31	37	37	38
4	39	29	31	34

- 1) Dat ligt ook voor de hand. Doordat bij de metelling weinig bedrijven méér dieren en gewassen zullen opgeven dan aanwezig en meer bedrijven minder zullen opgeven dan aanwezig, zullen vooral de laagste sbe-klassen daarvan de gevolgen ondervinden. In de intensieve veehouderij speelt incidentele leegstand van de hokken op 1 mei een grote rol.

Tabel 12.9 (vervolg)

Type/ sbe- klasse	Variabele				
	sbe gem. tijdens boekjaar (***) (*0,1)	arbeids- opbrengst ondernemers (*1000)	*)	arbeids- opbrengst per ondern. (*1000)	***) factor- opbrengst (*1.000)
4- 1	50	44		45	50
2	46	33		33	38
3	42	24		30	32
4	51	29		35	39
5- 1	42	34		36	32
2	36	46		48	44
3	41	28		29	30
4	33	22		27	34
6- 1	23	29		29	25
2	27	25		26	18
3	27	39		42	39
4	30	18		22	29
7- 1	26	27		27	19
2	43	36		39	36
3	31	38		39	33
4	36	19		23	24
8- 1	37	35		35	32
2	47	34		37	37
3	38	31		32	32
4	32	22		31	29

*) Arbeidsopbrengst van de ondernemer(s) (FIP).

***) Arbeidsopbrengst van de ondernemer per ondernemer (BUL).

****) 1=akkerbouw; 2=rundvee; 3=varkens; 4=pluimvee; 5= gem. akk.
6=gem.rundvee/akk.; 7=gem.rundvee/intensief; 8=gem. intens.

De standaarddeviatie van $(sbe)^{\%}$ in de metelling op het moment van de bedrijfskeuze geeft voor sbe-klasse 4 wel een goede benadering van de variantie in $(sbe)^{\%}$ gedurende het boekjaar. De variantie in sbe-klasse-3 wordt ongeveer 30 à 40% onderschat, die in sbe-klasse-2 ongeveer 40 à 50%, terwijl de onderschatting in sbe-klasse-1 nogal varieert per bedrijfstype. De onderschatting is in het algemeen minder op rundveebedrijven dan op de andere bedrijven.

Samenvattend kan gezegd worden dat het vooral aan het verschil tussen het aantal sbe in de metelling en het aantal sbe in

de boekhouding te danken is dat $(sbe)^{\frac{1}{2}}$ een goede hulpvariabele is voor de optimale allocatie ten behoeve van de doelvariabele "sbe gemiddeld tijdens boekjaar". Gemiddeld over de acht bedrijfstypen is het quotiënt voor deze doelvariabele voor de vier sbe-klassen respectievelijk 34, 38, 36 en 38. Daarbij valt wel op dat in de pluimveehouderij de quotiënten enigszins hoger liggen. Incidentele leegstand van hokken per 1 mei zorgt er voor dat er binnen een sbe-klasse een grotere heterogeniteit in termen van bedrijfs-grootte aanwezig is dan op grond van de meetingsgegevens vermoed zou kunnen worden.

Ten aanzien van de arbeidsopbrengst van de ondernemers (kolom 3 in tabel 12.9) kan geconcludeerd worden dat de quotiënten van de standaarddeviaties van de hulpvariabele en deze doelvariabele meer variatie tussen de basisstrata vertonen dan die bij het aantal sbe. In de vier sbe-klassen is het quotiënt gemiddeld over de acht bedrijfstypen respectievelijk 33.000, 38.000, 36.000 en 26.000. Het quotiënt in sbe-klasse 4 is dus fors lager. Dat geeft aan dat de variatie in $(sbe)^{\frac{1}{2}}$ voor deze klasse een minder goede indicatie geeft voor de variatie in de arbeidsopbrengst.

Worden de verhoudingsgetallen van de doelvariabelen (kolommen 2 en 3) gemiddeld en vervolgens per sbe-klasse ongewogen gemiddeld over de bedrijfstypen dan is het gemiddelde in de vier sbe-klassen achtereenvolgens 33, 38, 36 en 32. De laagste en de hoogste sbe-klasse hebben een iets lagere verhouding tussen standaarddeviaties van doelvariabelen en hulpvariabele dan de tussenliggende klassen. De verschillen zijn echter verrassend klein. De variabele $(sbe)^{\frac{1}{2}}$ is gekozen als een compromis tussen "sbe" en "de wortel uit sbe". Dat blijkt een goede keus te zijn geweest. Ook voor een aantal andere doelvariabelen geeft deze hulpvariabele in grote lijnen goed aan wat de verhoudingen in de standaarddeviaties zijn. Bij de factoropbrengst (kolom 4) zijn de gemiddelde verhoudingen tussen de sbe-klassen respectievelijk 30, 37, 35 en 30. Bij de arbeidsopbrengst per ondernemer (kolom 5) lopen de quotiënten nog iets meer uiteen dan bij de arbeidsopbrengst van alle ondernemers op een bedrijf. Gezien de aard van deze variabele ligt dat ook voor de hand. Door de arbeidsopbrengst te delen door het aantal ondernemers, wordt gedeeltelijk gecorrigeerd voor de grootte van de bedrijven.

Een ander interessant punt uit de tabel is de gemiddelde verhouding per bedrijfstype. Middeling van alle getallen in de kolommen 2 en 3 geeft voor de rundveebedrijven een verhouding van 30. Bij de akkerbouwbedrijven en de gemengde akkerbouwbedrijven is dit respectievelijk 38 en 37. De varkenshouderijbedrijven laten een verhouding van 34 zien, de pluimveebedrijven 41, en de gemengd intensieve bedrijven 36. De gemengde groepen met rundveehouderij komen uit in de buurt van de zuivere rundveebedrijven, namelijk op 28 en 33, waarbij het laatste percentage betrekking heeft op de bedrijven met relatief veel intensieve veehouderij.

Indien bij de allocatie alleen zou worden afgegaan op de standaarddeviatie van de hulpvariabele dan zouden meer rundveebedrijven in de steekproef terechtkomen dan op grond van betrouwbare schatting van de doelvariabelen gewenst is. De regel dat van de gealloceerde rundveebedrijven er 100 worden toegewezen aan andere bedrijfstypen (paragraaf 6.6.4) sluit dus uitstekend aan bij bovenstaande cijfers.

Verder kan geconcludeerd worden dat vanuit het gezichtspunt van een betrouwbare schatting van gemiddelden en totalen voor de nationale boerderij, het steekproefpercentage voor de pluimvee- en de akkerbouwbedrijven bij de huidige allocatie iets aan de lage kant is, terwijl dat voor gemengde rundveebedrijven iets aan de hoge kant is.

12.9.4 Enkele kanttekeningen

De conclusies uit de vorige sub-paragraaf hebben op één punt nog een nadere toelichting nodig. Er is geconstateerd dat de standaarddeviatie van $(sbe)^{\frac{1}{2}}$ in de meitelling in de meeste basisstrata lager ligt dan de standaarddeviatie van deze variabele zoals die wordt gemeten in het boekhoudnet. Vooral de laagste sbe-klassen worden daarmee geconfronteerd. Daarvoor zijn ook redenen aangegeven. In sbe-klasse 4 is er nauwelijks sprake van een grotere standaarddeviatie in het boekhoudnet. Dat laatste is echter voor een deel maar schijn. De oorzaak daarvan ligt in het feit dat de zeer grote bedrijven enigszins ondervertegenwoordigd zijn in de steekproef, met name in de intensieve veehouderij. Tabel B5.1 in bijlage 5 maakt dat duidelijk. De variantie in $(sbe)^{\frac{1}{2}}$ in de meitelling (kolom 2) blijkt in de betreffende strata te worden onderschat door de steekproefbedrijven (kolom 3). Een duidelijk voorbeeld daarvan is te vinden in sbe-klasse 4 in de pluimveehouderij. De meitelling komt op een standaarddeviatie van 79 en het boekhoudnet op een standaarddeviatie van 49 voor dezelfde meitelingsgrootte. Zouden de zeer grote bedrijven wel aanwezig zijn in het boekhoudnet dan zou laatstgenoemde standaarddeviatie hoger zijn, en daarmee ook de standaarddeviatie van de doelvariabelen. De quotiënten in sbe-klasse 4 in tabel 12.9 zouden daarmee ook hoger worden en de geschiktheid van de hulpvariabele voor de allocatie zou daardoor nog toenemen. De quotiënten lagen bij deze sbe-klassen immers iets aan de lage kant.

12.10 Bedrijfstypering en betrouwbaarheid

Niet iedere bedrijfstypering is even geschikt voor het vormen van homogene groepen bij de stratificatie. Het is de vraag in hoeverre de 67%-grens bij de huidige stratificatie en de indeling aan de hand van vier "sectoren" akkerbouw, rundveehouderij, varkenshouderij en pluimveehouderij, evenwichtig is vanuit het gezichtspunt van homogeniteit ten aanzien van de doelvariabelen (paragraaf 6.8).

Voor een aantal alternatieve typeringen is nagegaan wat de invloed ervan is op de betrouwbaarheid van de steekproefuitkomsten. Daarbij bleek dat de wijze van definiëren van de sectoren weinig invloed heeft op de homogeniteit en de daaruit voortvloeiende standaardfouten voor schattingen voor de gemiddelde nationale boerderij; althans indien de genoemde vier sectoren maar in hoofdlijnen in beeld komen. Afsplitsing van enkele kleinere onderdelen van bedrijfstypen (bijvoorbeeld slachtkuikens- en mestkalverenbedrijven) levert wel een iets grotere interne homogeniteit, maar gezien de geringe aantallen bedrijven waar het daarbij om gaat, tikt dat nauwelijks door in de standaardfouten op nationaal niveau.

De grens tussen zuivere en gemengde bedrijven heeft evenmin veel invloed op de standaardfouten. Een grens van 60%, 80% of 90% blijkt in het algemeen hogere standaardfouten te geven dan de nu gehanteerde grens van 67% of een grens van 70%. Bij grenzen van 80% en 90% komt dat merkwaardig genoeg niet alleen door een grotere heterogeniteit binnen de gemengde groepen, maar ook door het heterogener worden - in termen van de doelvariabelen - van de groepen zuivere bedrijven. Handhaving van de huidige grens van 66 2/3% verdient dus de voorkeur.

Afhankelijk van de gebruikte bedrijfstypering heeft de stratificatie naar bedrijfsoppervlakte binnen de basisstrata een grotere of kleinere invloed. Bij de huidige basisstratificatie levert het oppervlakte-criterium echter nauwelijks een wezenlijke bijdrage aan de vermindering van de onbetrouwbaarheid. Wordt "regio" als derde stratificatie-variabele gehanteerd, dan resulteert vaak een grotere betrouwbaarheid op nationaal niveau dan bij de huidige volgorde. De verschillen zijn echter minimaal.

13. Betrouwbaarheid op groeps-niveau

13.1 Inleiding

In hoofdstuk 7 is aangegeven dat de tweede doelstelling van het boekhoudnet betrekking heeft op het verkrijgen van "zo betrouwbaar mogelijke gemiddelde uitkomsten (...) van groepen bedrijven, die zijn gevormd op basis van bedrijfsomvang, bedrijfstype en regionale ligging" (Bedrijfsuitkomsten, 1988:9).

In dit hoofdstuk wordt stilgestaan bij de wijze waarop BUL- en FIP-groepen (kunnen) worden gevormd, en welke consequenties dat heeft voor de betrouwbaarheid van uitkomsten voor die groepen.

13.2 Indeling op basis van boekhoudnetgegevens

Bij de stratificatie van de populatie worden strata gevormd op basis van bedrijfskenmerken in de metelling. De zo ontstane strata zijn homogeen in termen van metellingsvariabelen. Dat betekent echter nog niet dat ze ook homogeen zijn in termen van de situatie zoals die gemiddeld gedurende het boekjaar op de bedrijven is. Een bedrijf dat per 1 mei geregistreerd staat als klein akkerbouwbedrijf, kan gedurende het boekjaar een groot pluimveebedrijf blijken te zijn, doordat er op 1 mei toevallig geen dieren op het bedrijf aanwezig waren.

Hoewel er geen grote verschillen zijn tussen een indeling van bedrijven op basis van metellingsgegevens en een indeling op basis van de gemiddelde situatie gedurende het boekjaar, wordt de laatstgenoemde indeling door het LEI geprefereerd voor de presentatie van uitkomsten voor BUL-groepen. Daarbij worden zowel het type als de grootte van de bedrijven bepaald op deze boekhoudnetbasis.

De mogelijkheden om de steekproefbedrijven in te delen in groepen op basis van bedrijfstype, gebied en bedrijfsomvang zijn beperkt. Om voor een afzonderlijke groep betrouwbare cijfers te kunnen publiceren, moeten er voldoende steekproefbedrijven voor die groep beschikbaar zijn. Hoeveel steekproefbedrijven nodig zijn, hangt af van de heterogeniteit van de betreffende BUL-groep. Vanaf 1975 zijn er steeds circa twintig groepen gevormd, waarvoor per groep minimaal 35 tot 40 steekproefbedrijven beschikbaar waren.

In de eerste plaats wordt er een indeling in vijf bedrijfstypen gemaakt. Daarbij worden de in tabel 3.2 vermelde type-grenzen in grote lijnen aangehouden (zie ook Bedrijfsuitkomsten, 1988:10). Vervolgens worden de bedrijven ingedeeld in grotere en

kleinere bedrijven 1). Binnen deze tien groepen wordt nog een verdergaande indeling naar gebieden gemaakt (Bedrijfsuitkomsten, 1988:14). Een klein aantal bedrijven dat wel deel uitmaakt van de nationale boerderij, wordt niet ingedeeld in een bepaalde BUL-groep.

Naast een indeling in twintig min of meer homogene BUL-groepen, worden de bedrijven ook op een andere wijze ingedeeld. Per bedrijfstype wordt een vrij ver gaande indeling gemaakt in grootte-klassen, zonder daarbij te letten op de gebieden waar de bedrijven liggen (Bedrijfsuitkomsten, 1988:13,134-157).

13.3 Van strata naar BUL-groepen

De indeling van bedrijven in strata - bij de stratificatie vooraf en achteraf - vindt plaats op grond van een andere bedrijfstypering, grootte-klasse-indeling en regio-groepering dan de indeling van bedrijven in BUL-groepen bij de presentatie van bedrijfsuitkomsten. Een bedrijf dat bij de steekproeftrekking "akkerbouwbedrijf" heet, kan in de BUL bijvoorbeeld getypeerd worden als "gemengd akkerbouwbedrijf" of als "gemengd intensief bedrijf". Naast deze definitie-verschillen zijn er de in de vorige paragraaf besproken verschillen ten aanzien van de basis waarop de bedrijven worden ingedeeld.

Een en ander heeft tot gevolg dat BUL-groepen betrekking hebben op delen van verschillende strata, waarbij vooraf niet bekend is welke aantallen bedrijven in welke strata tot welke BUL-groep behoren. Indien een stratum 300 populatie-bedrijven bevat en er zijn 15 steekproefbedrijven in dat stratum, waarvan er 6 tot het BUL-type "grotere weidebedrijven in het Zuidelijk Zandgebied" behoren, dan is de steekproef-schatting van het aantal bedrijven in dat stratum met dat BUL-type gelijk aan 120.

In bijlage 9 wordt aangegeven dat voor groepen (deelpopulaties) die delen van strata bevatten, de standaardfout anders berekend moet worden dan voor groepen die zijn samengesteld uit gehele strata. De schattingsfout in het aantal bedrijven komt immers boven op de schattingsfout in de waarde van een variabele. Daardoor heeft een totaalschatting voor een BUL-groep (paragraaf 13.5) een grotere standaardfout dan een schatting van een gemiddelde (paragraaf 13.4).

1) De (sbe-)grens tussen grotere en kleinere bedrijven "komt bij benadering overeen met een produktie-omvang van een man, die gemiddeld bij een doelmatige bedrijfs-grootte wordt gerealiseerd" (Bedrijfsuitkomsten, 1988:11). Het is een grens die verschuift met de arbeidsproductiviteit.

Tabel 13.1 *Standaardfouten (in honderden guldens) volgens formule (B9.27) naar BUL-groep en variabele (gemiddeld over de boekjaren 1981-1983)*

BUL-groep	Variabele					
	netto bedr. result.	arb. opbr. *)	tot. opbr.	totale kosten	fact. opbr.	sbe- bkh. **)
Grotere bedrijven						
Akkerbouw:						
- noordelijk klei	85	94	146	118	94	11
- veenkolonien	76	73	165	173	75	16
- centraal klei	95	96	163	107	110	9
- zuidwestelijk klei	82	85	148	102	101	8
Weide:						
- noordelijk klei	54	62	142	134	58	9
- noord. veenweide	68	70	210	196	80	13
- westelijk weide	67	86	190	170	91	11
- noordelijk zand	65	67	150	136	69	10
- oost./centr. zand	38	50	109	105	39	5
- zuidelijk zand	45	49	136	121	57	7
Akkerbouw gemengd	125	119	372	294	175	20
Rundveehouderij gemengd	49	49	196	182	58	8
Int. veeh. gemengd:						
- nrd/oost/centr. zand	65	70	242	243	73	8
- zuidelijk zand	58	59	446	439	68	12
Kleinere bedrijven						
Akkerbouw:						
- kleigebieden	83	57	80	71	58	4
- veenkol./nrd. zand	56	39	67	83	41	5
Weide:						
- klei/veenweide	32	36	50	59	24	3
- zand	32	34	53	60	23	3
Rundveehouderij gemengd	50	50	117	119	39	4
Int.veeh. gemengd (zand)	39	39	189	188	36	3

*) Arbeidsopbrengst van de ondernemer per ondernemer.

***) Sbe gemiddeld gedurende het boekjaar.

Tabel 13.2 Verhouding tussen standaardfouten berekend volgens formule (6.4.7) en standaardfouten berekend volgens formule (B9.27) naar BUL-groep en variabele (gemiddeld over de boekjaren 1981-1983)

BUL-groep	Variabele					
	netto bedr. result.	arb. opbr.	tot. opbr.	totale kosten	fakt. opbr.	sbe- bkh.
Grotere bedrijven						
Akkerbouw:						
- noordelijk klei	96	96	152	168	125	178
- veenkolonien	111	125	160	160	147	171
- centraal klei	114	115	148	158	139	174
- zuidwestelijk klei	113	114	156	161	135	183
Weide:						
- noordelijk klei	102	107	171	169	159	176
- noord. veenweide	107	115	149	149	144	159
- westelijk weide	91	87	132	134	116	146
- noordelijk zand	105	104	159	159	142	160
- oost./centr. zand	108	107	148	141	139	163
- zuidelijk zand	119	129	144	138	151	178
Akkerbouw gemengd	98	103	97	99	104	115
Rundveehouderij gemengd	113	117	110	110	123	124
Int. veeh. gemengd:						
- nrd/oost/centr. zand	115	116	147	144	134	166
- zuidelijk zand	107	107	112	112	119	133
Kleinere bedrijven						
Akkerbouw:						
- kleigebieden	85	101	120	106	109	128
- veenkol./nrd. zand	105	103	95	93	97	99
Weide:						
- klei/veenweide	110	105	108	106	109	117
- zand	102	99	102	100	104	108
Rundveehouderij gemengd	113	116	116	117	111	111
Int.veeh. gemengd (zand)	110	106	125	126	101	121

13.4 Betrouwbaarheid van uitkomsten van BUL-groepen

13.4.1 Geschatte wegingsfactoren

In bijlage 3 van de BUL wordt voor de onderscheiden BUL-groepen aangegeven wat de standaardfout is van drie variabelen (Bedrijfsuitkomsten, 1988:174). Deze standaardfouten worden berekend met formule (6.4.7). Daarbij wordt er dus van uitgegaan dat de bedrijven weliswaar wegingsfactoren hebben, maar verder wordt er geen gebruik gemaakt van de informatie met betrekking tot de stratificatie. Evenmin wordt rekening gehouden met de schattingsfout van de wegingsfactoren. Een dergelijke vereenvoudiging in de berekening van standaardfouten is aantrekkelijk, maar heeft alleen zin indien bekend is in hoeverre deze berekening een goede graadmeter oplevert voor de orde van grootte van de standaardfouten, berekend volgens formule (B9.27). Is de benadering via de vereenvoudigde formule goed of wordt de standaardfout systematisch onder- of overschat? In deze paragraaf wordt daarop ingegaan.

Tabel 13.1 geeft de grootte van de standaardfouten aan, berekend volgens formule B9.27; in deze formule wordt rekening gehouden met de stratificatie en met de schattingsfouten in de wegingsfactoren.

De standaardfouten zijn in het algemeen groter voor de akkerbouwgroepen dan voor de overige BUL-groepen. Voor een deel wordt dit veroorzaakt door extra hoge standaardfouten in 1983 - een extreem jaar in de akkerbouw -, maar ook in andere jaren zijn de standaardfouten hoger door de grotere heterogeniteit in dit deel van de populatie. Dat geldt in nog sterkere mate voor de groep grotere gemengde akkerbouwbedrijven. De heterogeniteit wordt daar versterkt door het kleine aantal bedrijven.

In tabel 13.2 wordt voor een aantal variabelen weergegeven wat de verhouding is tussen de standaardfout die in de BUL wordt weergegeven (formule 6.4.7) en de "werkelijke" standaardfout (B9.27) 1). Bij de berekening van de standaardfout volgens formule (B9.27) moeten strata waarin slechts één bedrijf tot een bepaalde BUL-groep behoort, worden samengevoegd met andere strata. Strata waarin geen enkel steekproefbedrijf voorkomt dat behoort tot de BUL-groep, worden bij de berekening van de standaardfout buiten beschouwing gelaten.

Uit deze tabel blijkt dat de standaardfout van het netto-bedrijfsresultaat en de arbeidsopbrengst voor vrijwel alle groepen goed geschat wordt door de benaderingsformule (6.4.7). Voor grootheden die sterker samenhangen met de stratificatie-criteria,

1) De werkelijke standaardfouten zijn niet te berekenen. De standaardfouten worden immers uit het steekproefmateriaal berekend en zijn eveneens onderhevig aan schattingsfouten (var (y) is een schatting voor Var (y)).

Tabel 13.3 Verhouding tussen standaardfouten berekend volgens formule (6.4.5) (vaste wegingsfactoren) en standaardfouten berekend volgens formule (B9.27) (geschatte wegingsfactoren) naar BUL-groep en variabele (gemiddeld over de boekjaren 1981-1983)

BUL-groep	Variabele					
	netto bedr. result.	arb. opbr.	tot. opbr.	totale kosten	fact. opbr.	sbe- bkh.
Grotere bedrijven						
Akkerbouw:						
- noordelijk klei	96	96	74	73	85	69
- veenkolonien	97	95	79	82	82	80
- centraal klei	94	94	88	90	88	93
- zuidwestelijk klei	92	91	74	72	83	68
Weide:						
- noordelijk klei	102	99	75	74	86	69
- noord. veenweide	102	100	79	80	82	69
- westelijk weide	101	104	82	83	87	73
- noordelijk zand	103	99	79	75	89	71
- oost./centr. zand	95	93	80	86	86	71
- zuidelijk zand	98	96	88	90	90	76
Akkerbouw gemengd	114	109	103	96	101	83
Rundveehouderij gemengd	96	91	83	85	77	72
Int. veeh. gemengd:						
- nrd/oost/centr.zand	105	109	96	97	105	90
- zuidelijk zand	106	107	98	98	106	94
Kleinere bedrijven						
Akkerbouw:						
- kleigebieden	104	105	96	95	101	87
- veenkol./nrd. zand	102	103	102	103	104	100
Weide:						
- klei/veenweide	109	102	99	105	104	89
- zand	104	104	96	96	103	96
Rundveehouderij gemengd	106	107	101	101	112	96
Int.veeh. gemengd (zand)	107	107	79	80	106	102

is er in het algemeen sprake van een overschatting van de standaardfouten. Daar wordt de vergroting van de standaardfout door het rekening houden met de onzekerheid van de wegingsfactoren, ruim gecompenseerd door de vermindering van de standaardfout als gevolg van het rekening houden met de stratificatie.

De tabel illustreert ook dat voor BUL-groepen die in grote lijnen corresponderen met groepen in de stratificatie, de standaardfouten sterker gereduceerd worden door de stratificatie dan voor BUL-groepen waar dat niet het geval is. Het meest duidelijk komt dit naar voren bij de groepen grotere gemengde bedrijven. Deze groepen bevatten meestal bedrijven die afkomstig zijn uit een groot aantal strata, waar ze slechts een deel van de daar aanwezige bedrijven uitmaken. De stratificatie-winst wordt in die groepen dus gedeeltelijk teniet gedaan door de schattingsfout van de wegingsfactoren.

De reductie bij de kleinere bedrijven is ook minder dan die bij grotere bedrijven. Dit wordt veroorzaakt door het feit dat deze groep voornamelijk afkomstig is uit sbe-klasse-1, terwijl de grotere bedrijven afkomstig zijn uit de overige sbe-klassen.

13.4.2 Vaste wegingsfactoren

Indien de wegingsfactoren geen schattingen waren, maar op de een of andere wijze bekend zouden zijn, zouden de standaardfouten enigszins lager liggen. Tabel 13.3 geeft aan welke reductie in standaardfouten gerealiseerd zou worden indien er geen schattingsfout voor de wegingsfactoren aanwezig zou zijn.

In het algemeen is het te verwachten dat een schatting op basis van vaste wegingsfactoren tot lagere standaardfouten leidt dan een schatting op basis van geschatte wegingsfactoren. De verhoudingen in tabel 13.3 zouden dus theoretisch gezien allemaal kleiner dan 100 moeten zijn. In bijlage 9 wordt aangegeven dat dit bij de gehanteerde formules niet altijd het geval hoeft te zijn. Vergelijking (6.4.5) is immers een exacte formule voor de variantie, terwijl vergelijking (B9.27) een formule is die bij benadering juist is en die bovendien alleen voor grotere steekproeven geldig is. Wanneer in tabel 13.3 getallen voorkomen die groter zijn dan 100 dan duidt dat er in elk geval op dat de onzekerheid in de wegingsfactoren weinig bijdraagt aan de vergroting van de variantie. Hoe lager het verhoudingsgetal, des te meer winst is er te boeken in termen van vermindering van de variantie als er zekerheid zou zijn over de wegingsfactoren.

Een conclusie die getrokken kan worden uit tabel 13.3 is dat de onzekerheid in de wegingsfactoren vooral tot een grotere variantie leidt bij variabelen die sterk samenhangen met de stratificatie-variabelen - vooral de sbe-variabele - en bij de groepen grotere bedrijven. Dat laatste hangt ook weer samen met het feit dat juist deze groepen het meeste profiteren van de stratificatie. Voor de variabelen opbrengsten, kosten, en sbe, ligt een reductie van de standaardfout van 20 tot 25% in het verschiet in-

Tabel 13.4 Verhouding tussen relatieve standaardfouten van totalen (formule B9.23) en gemiddelden (B9.27) naar BUL-groep en variabele (gemiddeld over de boekjaren 1981-1983)

BUL-groep	Variabele					
	netto bedr. result.	arb. opbr.	tot. opbr.	totale kosten	fact. opbr.	sbe- bkh.
Grotere bedrijven						
Akkerbouw:						
- noordelijk klei	101	111	192	240	153	218
- veenkolonien	102	120	165	171	151	171
- centraal klei	106	123	180	215	155	209
- zuidwestelijk klei	103	127	171	201	149	215
Weide:						
- noordelijk klei	118	110	191	225	178	226
- noord. veenweide	107	115	175	200	170	206
- westelijk weide	126	105	206	258	157	254
- noordelijk zand	111	110	186	223	162	207
- oost./centr. zand	129	108	250	284	231	341
- zuidelijk zand	104	125	218	253	191	256
Akkerbouw gemengd	106	117	177	217	154	232
Rundveehouderij gemengd	149	130	238	279	244	370
Int. veeh. gemengd:						
- nrd/oost/centr. zand	109	107	171	180	148	239
- zuidelijk zand	103	117	132	135	153	180
Kleinere bedrijven						
Akkerbouw:						
- kleigebieden	102	124	227	258	174	330
- veenkol./nrd. zand	127	124	292	282	296	388
Weide:						
- klei/veenweide	170	103	259	298	192	382
- zand	144	106	230	259	191	336
Rundveehouderij gemengd	231	103	270	346	245	525
Int.veeh. gemengd (zand)	128	102	144	158	135	317

dien er geen onzekerheid zou zijn met betrekking tot de wegingsfactoren. Dat lijkt in eerste instantie niet zo opzienbarend, maar het betekent wel dat indien voor de realisatie van een bepaalde grootte van de standaardfout nu bijvoorbeeld 50 bedrijven per BUL-groep noodzakelijk zijn, bij vaste wegingsfactoren volstaan zou kunnen worden met circa 30 bedrijven.

13.5 Betrouwbaarheid van totalen van BUL-groepen

Indien steekproefschattingen betrekking hebben op gehele strata of op combinaties van strata, is de relatieve standaardfout van een totaalschatting gelijk aan de relatieve standaardfout van de schatting van het gemiddelde (vergelijking 6.4.6). Bij schattingen voor groepen die dwars door strata heenlopen, is de relatieve standaardfout van het totaal echter groter dan die van het gemiddelde (zie bijlage 9). Dit wordt veroorzaakt door de schattingsfout die wordt gemaakt in het aantal bedrijven per stratum dat deel uitmaakt van de betreffende groep (zie paragraaf 13.8).

In tabel 13.4 wordt voor een aantal variabelen aangegeven hoeveel de relatieve standaardfout van het totaal groter is dan de relatieve standaardfout van het gemiddelde.

Voor variabelen die sterk samenhangen met de stratificatiecriteria en voor groepen bedrijven waarvoor er een grote onzekerheid is in het aantal bedrijven, is de relatieve standaardfout van het totaal fors hoger dan die voor het gemiddelde. Voor de groep kleinere gemengde rundveebedrijven bijvoorbeeld, is de relatieve standaardfout van het gemiddelde ongeveer 3% ¹⁾. De relatieve standaardfout van het totaal bedraagt ongeveer 16%. Dat wil zeggen dat bij de totaalschattingen soms onzekerheidsmarges van enkele tientallen procenten in acht genomen moeten worden. Bij vaststaande wegingsfactoren zouden deze marges veel kleiner zijn.

13.6 Betrouwbaarheid van uitkomsten van FIP-groepen

13.6.1 Geschatte wegingsfactoren

Tabel 13.5 geeft de grootte van de standaardfouten van een aantal grootheden op het gebied van de financiële positie van de ondernemer. De standaardfouten zijn berekend met vergelijking B9.27. Evenals bij de bedrijfsuitkomsten (tabel 11.1) zijn de standaardfouten in het algemeen groter voor de akkerbouwgroepen dan voor de overige FIP-groepen. De standaardfouten van groothe-

1) 4 sbe (zie tabel 13.1) gedeeld door een gemiddelde van ongeveer 120 sbe (zie Bedrijfsuitkomsten, staat B3).

Tabel 13.5 *Standaardfouten volgens formule (B9.27) naar FIP-groep en variabele (gemiddeld over de boekjaren 1982-1983) (in honderden guldens)*

FIP-groep	Variabele					
	arb. opbr. ondern.	gez. in-komen	be-spa-ring	bruto inv.	eigen ver-mogen	bal. to-taal
Grotere bedrijven						
Akkerbouw:						
- noordelijk klei	126	141	117	175	1350	1590
- veenkolonien	80	76	73	283	870	970
- centraal klei	134	130	119	77	1110	1030
- zuidwestelijk klei	84	78	72	104	1170	1330
Weide:						
- noordelijk klei	53	63	58	179	720	850
- noord. veenweide	80	71	62	192	670	770
- westelijk weide	75	73	48	222	600	660
- noordelijk zand	70	55	49	168	500	590
- oost./centr. zand	54	40	38	70	560	590
- zuidelijk zand	51	57	56	160	490	500
Akkerbouw gemengd	134	165	138	221	1160	1360
Rundveehouderij gemengd	51	54	47	101	490	560
Int. veeh. gemengd:						
- nrd/oost/centr.zand	83	85	76	97	450	450
- zuidelijk zand	71	68	74	164	420	570
Kleinere bedrijven						
Akkerbouw:						
- kleigebieden	68	67	58	59	570	630
- veenkol./nrd. zand	53	62	51	119	710	790
Weide:						
- klei/veenweide	34	29	34	78	380	420
- zand	36	25	25	64	250	270
Rundveehouderij gemengd	50	43	39	55	440	430
Int.veeh. gemengd (zand)	44	34	39	89	340	320

Tabel 13.6 Verhouding tussen standaardfouten berekend volgens formule (6.4.7) en standaardfouten berekend volgens formule (B9.27) naar FIP-groep en variabele (gemiddeld over de boekjaren 1982-1983)

FIP-groep	Variabele					
	arb. opbr. ondern.	gez. in- komen	be- spa- ring	bruto inv.	eigen verm. mogen	bal. to- taal
Grotere bedrijven						
Akkerbouw:						
- noordelijk klei	98	93	103	123	102	105
- veenkolonien	92	100	91	100	115	119
- centraal klei	100	109	92	96	111	120
- zuidwestelijk klei	108	119	119	248	107	113
Weide:						
- noordelijk klei	106	124	106	107	112	134
- noord. veenweide	103	102	96	90	98	121
- westelijk weide	86	98	105	92	110	114
- noordelijk zand	105	125	115	89	123	141
- oost./centr. zand	99	117	108	112	111	118
- zuidelijk zand	117	110	93	95	102	120
Akkerbouw gemengd	99	96	94	109	101	104
Rundveehouderij gemengd	116	111	106	99	107	110
Int. veeh. gemengd:						
- nrd/oost/centr.zand	107	109	100	125	109	130
- zuidelijk zand	106	123	108	99	127	114
Kleinere bedrijven						
Akkerbouw:						
- kleigebieden	107	118	113	178	118	113
- veenkol./nrd. zand	94	83	89	92	88	89
Weide:						
- klei/veenweide	121	104	99	121	110	110
- zand	102	104	100	106	105	104
Rundveehouderij gemengd	126	117	113	127	108	111
Int.veeh. gemengd (zand)	106	107	108	105	115	115

den die betrekking hebben op de financiële positie van de ondernemer zijn relatief gezien hoger dan die voor grootheden in de sfeer van de bedrijfsuitkomsten. De grotere heterogeniteit tussen bedrijven ligt daaraan ten grondslag, terwijl ook het kleinere aantal bedrijven een rol speelt.

In tabel 13.6 wordt voor een aantal variabelen weergegeven wat de verhouding is tussen de standaardfout waar in de LEI-praktijk mee wordt gewerkt (vergelijking 6.4.7) en de "werkelijke" standaardfout. De tabel is de FIP-variant van tabel 13.2.

Uit deze tabel blijkt dat de standaardfout van de meeste variabelen vrij goed benaderd wordt door de formule die geen rekening houdt met de stratificatie. Voor grootheden die sterker samenhangen met de stratificatie-criteria, is er in het algemeen sprake van een overschatting van de standaardfouten. Daar wordt de vergroting van de standaardfout door het rekening houden met de onzekerheid van de wegingsfactoren, ruim gecompenseerd door de vermindering van de standaardfout als gevolg van het rekening houden met de stratificatie.

De tabel illustreert ook dat voor FIP-groepen die in grote lijnen corresponderen met groepen in de stratificatie, de standaardfouten sterker gereduceerd worden door de stratificatie dan in FIP-groepen waar dat niet het geval is. Het meest duidelijk komt dit weer naar voren bij de groepen grotere gemengde bedrijven. De stratificatie-winst wordt daar grotendeels teniet gedaan door de schattingsfout van de wegingsfactoren. Als gevolg van het uitsluiten van NV's en BV's - die wel een wegingsfactor krijgen - in de FIP-populatie, speelt het probleem dat iedere groep gevormd wordt uit delen van strata met dientengevolge grotere onzekerheden in de geschatte aantallen bedrijven per groep, hier sterker dan bij de BUL-resultaten. Bij de groep grotere akkerbouwbedrijven in het noordelijk kleigebied is die invloed vrij sterk aanwezig.

13.6.2 Vaste wegingsfactoren

Indien de wegingsfactoren geen schattingen waren, maar op de een of andere wijze bekend zouden zijn, zouden de standaardfouten enigszins lager liggen. Tabel 13.7 geeft aan welke reductie in standaardfouten gerealiseerd zou worden indien er geen schattingsfout voor de wegingsfactoren aanwezig zou zijn.

Evenals in tabel 13.3 komen in tabel 13.7 getallen groter dan 100 voor. Die zijn het gevolg van het feit dat vergelijking (B9.27) een benaderingsformule is (zie paragraaf 13.4.2). Ook hier geldt weer dat hoe lager het verhoudingsgetal, des te meer winst er te boeken is in vermindering van de variantie als er zekerheid zou zijn over de wegingsfactoren.

Een conclusie die getrokken kan worden uit tabel 13.7 is dat de onzekerheid in de wegingsfactoren vooral tot een grotere variantie leidt bij variabelen die sterk samenhangen met de strati-

Tabel 13.7 Verhouding tussen standaardfouten berekend volgens formule (6.4.5) (vaste wegingsfactoren) en standaardfouten berekend volgens formule (B9.27) (geschatte wegingsfactoren) naar FIP-groep en variabele (gemiddeld over de boekjaren 1982-1983)

FIP-groep	Variabele					
	arb. opbr. ondern.	gez. in- komen	be- spa- ring	bruto inv.	eigen ver- mogen	bal. to- taal

Grotere bedrijven						
Akkerbouw:						
- noordelijk klei	101	99	97	94	93	86
- veenkolonien	100	98	101	101	97	97
- centraal klei	105	104	105	104	101	101
- zuidwestelijk klei	93	92	95	91	98	96
Weide:						
- noordelijk klei	101	91	109	100	92	86
- noord. veenweide	100	98	102	102	100	97
- westelijk weide	99	95	91	95	94	89
- noordelijk zand	102	99	104	107	99	86
- oost./centr. zand	93	96	101	104	104	102
- zuidelijk zand	96	97	103	102	97	93
Akkerbouw gemengd	115	110	112	97	100	94
Rundveehouderij gemengd	90	97	97	105	94	85
Int. veeh. gemengd:						
- nrd/oost/centr.zand	110	104	107	101	92	87
- zuidelijk zand	103	107	111	146	107	113
Kleinere bedrijven						
Akkerbouw:						
- kleigebieden	111	97	110	102	110	104
- veenkol./nrd. zand	95	95	95	96	77	70
Weide:						
- klei/veenweide	105	102	105	124	102	107
- zand	100	104	102	101	100	103
Rundveehouderij gemengd	107	105	110	117	116	110
Int.veeh. gemengd (zand)	109	113	108	103	113	117

Tabel 13.8 Verhouding tussen relatieve standaardfouten van totalen (formule B9.23) en gemiddelden (B9.27) naar FIP-groep en variabele (gemiddeld over de boekjaren 1982-1983)

FIP-groep	Variabele					
	arb. opbr. ondern.	gez. in-komen	be-spa-ring	bruto inv.	eigen ver-mogen	bal. to-taal
Grotere bedrijven						
Akkerbouw:						
- noordelijk klei	131	154	110	116	162	178
- veenkolonien	113	134	102	105	137	149
- centraal klei	112	124	106	108	109	114
- zuidwestelijk klei	137	163	117	114	117	119
Weide:						
- noordelijk klei	116	156	108	107	142	163
- noord. veenweide	109	141	106	105	138	151
- westelijk weide	106	143	112	101	186	222
- noordelijk zand	108	162	115	116	154	175
- oost./centr. zand	105	192	119	117	271	191
- zuidelijk zand	131	186	120	108	193	248
Akkerbouw gemengd	114	138	109	113	170	178
Rundveehouderij gemengd	113	195	114	113	240	268
Int. veeh. gemengd:						
- nrd/oost/centr.zand	99	116	100	113	165	220
- zuidelijk zand	113	143	104	105	170	171
Kleinere bedrijven						
Akkerbouw:						
- kleigebieden	111	129	111	107	135	145
- veenkol./nrd. zand	111	132	105	107	140	141
Weide:						
- klei/veenweide	101	137	100	103	145	156
- zand	102	154	106	102	181	198
Rundveehouderij gemengd	99	167	101	122	247	276
Int.veeh. gemengd (zand)	102	122	102	101	151	188

ficatie-variabelen - vooral het balanstotaal en het eigen vermogen - en bij de groepen grotere bedrijven. Dat laatste hangt ook weer samen met het feit dat juist deze groepen het meest profiteren van de stratificatie. Voor de meeste variabelen ligt een reductie van de standaardfout met slechts enkele procenten in het verschiep, indien de onzekerheid met betrekking tot de wegingsfactoren zou worden weggenomen.

13.7 Betrouwbaarheid van totalen van FIP-groepen

In paragraaf 13.5 is aangegeven dat voor groepen bedrijven die dwars door strata heenlopen, de standaardfout van het totaal groter is dan het produkt van het aantal bedrijven en de standaardfout van het gemiddelde. In tabel 13.8 wordt voor een aantal variabelen aangegeven hoeveel de relatieve standaardfout van het totaal groter is dan de relatieve standaardfout van het gemiddelde.

Evenals bij de BUL-groepen is voor variabelen die sterk samenhangen met de stratificatie-criteria en voor groepen bedrijven waarvoor er een grote onzekerheid is in het aantal bedrijven, de relatieve standaardfout van het totaal fors hoger dan die voor het gemiddelde. Omdat in veel gevallen de relatieve standaardfout van het gemiddelde al vrij groot is, betekent dit dat totalen voor FIP-groepen met de nodige voorzichtigheid geïnterpreteerd moeten worden.

13.8 Betrouwbaarheid van schattingen van aantallen bedrijven

De samenstelling van BUL-groepen op grond van eigenschappen van bedrijven gemiddeld gedurende het boekjaar, zorgt ervoor dat de som van de wegingsfactoren - het geschatte aantal bedrijven - in een bepaalde groep, aan het steekproeftoeval onderhevig is. In tabel 13.9 is weergegeven hoe groot de standaardfouten van de geschatte aantallen bedrijven zijn voor de BUL-groepen. Deze standaardfouten zijn berekend met vergelijking (B9.29) uit bijlage 9.

De standaardfouten zijn in het algemeen kleiner dan 10% van de geschatte aantallen. Een uitzondering daarop vormen de groepen gemengde akkerbouwbedrijven en gemengde rundveebedrijven. Ook de groepen met kleine akkerbouwbedrijven laten relatief hoge standaardfouten zien.

Het creëren van 95%-betrouwbaarheidsintervallen voor de geschatte aantallen houdt in dat de intervalsgrenzen gevormd worden door de geschatte aantallen te verminderen c.q. te vermeerderen met de standaardfout maal T, waarbij T ontleent wordt aan een Student-verdeling (T-verdeling). Gezien de geringe aantallen in de BUL-groepen, zal T in het algemeen groter zijn dan 2.0. De 95%-betrouwbaarheidsintervallen omvatten in sommige gevallen dan ook een erg ruim traject. De aantallen bedrijven waarop de resul-

taten van de BUL-groepen betrekking hebben, moeten dan ook met de nodige voorzichtigheid geïnterpreteerd worden.

De kolommen "minimum" en "maximum" in tabel 13.9 geven aan wat de minimale en maximale geschatte aantallen bedrijven in de periode 1981-1986 waren.

Tabel 13.9 Geschatte aantallen bedrijven en bijbehorende standaardfouten naar BUL-groep (boekjaren 1981-1986)

BUL-groep	Geschatte aantal bedrijven		Gemiddelde standaardfout
	minimum	maximum	
Grotere bedrijven			
Akkerbouw:			
- noordelijk klei	952	1.405	113
- veenkolonien	1.368	1.712	166
- centraal klei	2.009	2.479	169
- zuidwestelijk klei	1.991	2.426	187
Weide:			
- noordelijk klei	3.375	3.784	325
- noord. veenweide	1.859	2.226	169
- westelijk weide	2.788	4.210	336
- noordelijk zand	3.181	3.808	310
- oost./centr. zand	4.228	5.578	403
- zuidelijk zand	3.372	5.742	381
Akkerbouw gemengd	1.009	1.471	199
Rundveehouderij gemengd	2.828	4.352	545
Int. veeh. gemengd:			
- nrd/oost/centr.zand	4.135	6.378	473
- zuidelijk zand	4.870	5.596	433
Kleinere bedrijven			
Akkerbouw:			
- kleigebieden	2.029	2.299	257
- veenkol./nrd. zand	761	1.113	181
Weide:			
- klei/veenweide	4.575	6.200	471
- zand	6.871	7.935	611
Rundveehouderij gemengd	1.476	3.148	487
Int.veeh. gemengd (zand)	5.466	6.154	466

Voor de vergelijkbare FIP-groepen zijn de standaardfouten van de geschatte aantallen uiteraard nog groter.

13.9 Betrouwbaarheid van verschillen tussen groepen 1)

Cijfers over de situatie in een bepaalde groep in een bepaald boekjaar worden soms vergeleken met die in een andere groep in datzelfde jaar. Daarbij is het dan de vraag of verschillen in uitkomsten tussen groepen bedrijven zijn toe te schrijven aan het toeval of dat er significante verschillen geconstateerd kunnen worden. De grootte van de standaardfout van het verschil tussen twee steekproefschattingen kan helpen bij de beantwoording van die vraag.

Bij een enkelvoudige steekproef is de variantie van het verschil tussen twee schattingen \bar{x} en \bar{y} gelijk aan:

$$\text{Var}(\bar{x} - \bar{y}) = \text{Var}(\bar{x}) + \text{Var}(\bar{y}) - 2 * \text{cov}(\bar{x}, \bar{y}) \quad (13.9.1)$$

Is er sprake van een gestratificeerde steekproef en worden groepen gevormd die dwars door de strata heenlopen ("crossclasses") dan is de volgende formule van toepassing (Kish and Frankel, 1974):

$$\begin{aligned} \text{var}(\bar{y}_a - \bar{y}_b)_{st} &\approx \sum_{h=1}^H \frac{(N_h - n_h)}{n_h - 1} * \\ &* \frac{1}{\hat{A}^2 * a_h^2} * \sum_{j=1}^{a_h} (y_{ahj} - \bar{y}_a)^2 + \frac{1}{\hat{B}^2 * b_h^2} * \sum_{j=1}^{b_h} (y_{bhj} - \bar{y}_b)^2 \\ &- \frac{N_h}{n^2 h} * \left[\frac{a_h}{\hat{A}} * (\bar{y}_{ah} - \bar{y}_a) - \frac{b_h}{\hat{B}} * (\bar{y}_{bh} - \bar{y}_b) \right]^2 \quad (13.9.2) \end{aligned}$$

waarbij:

- a_h = aantal steekproefbedrijven in groep a in stratum h
- b_h = aantal steekproefbedrijven in groep b in stratum h
- \hat{A}, \hat{B} = schatting van het aantal bedrijven in de populatie in groep a (= $W_h * a_h$) respectievelijk b (= $W_h * b_h$)
- \bar{y}_a, \bar{y}_b = gemiddelde waarde in groep a respectievelijk b
- \bar{y}_{ah} = gemiddelde waarde in groep a in stratum h

1) De betrouwbaarheid van verschillen tussen jaren van een bepaalde groep komen aan de orde in paragraaf 15.3.1.

\bar{y}_{bh} = gemiddelde waarde in groep b in stratum h
 y_{ahj} = waarde van element j in groep a in stratum h

Deze formule zegt dat de variantie van het verschil gelijk is aan de som van de afzonderlijke varianties (B9.27) 1). Voor de berekeningen van standaardfouten van verschillen tussen BUL-groepen kan daarom gebruik gemaakt worden van de standaardfouten in de tabellen 13.1 en 13.5. Kish suggereert dat indien het gaat om kleine "subclasses" binnen strata, de derde term tussen accolades genegeerd kan worden. Berekeningen voor de LEI-steekproef wijzen uit dat weglating van deze term voor BUL-groepen kan leiden tot standaardfouten die enkele tientallen procenten hoger liggen dan die op basis van (13.9.2).

Veel "verschillen" in gemiddelde waarden tussen BUL-groepen zijn terug te voeren op het steekproeftoeval. In 1986 bijvoorbeeld was het netto bedrijfsresultaat op grotere akkerbouwbedrijven in de Veenkoloniën, het Noordelijk, het Centraal en het Zuidwestelijk Kleigebied respectievelijk -38.000, -32.000, -43.000 en -9.900 gulden. Op grond van de standaardfouten in tabel 13.3 kan geconcludeerd worden dat er geen statistisch significante verschillen in netto bedrijfsresultaat aanwezig zijn tussen de eerste drie groepen. Ten aanzien van het Zuidwestelijk Kleigebied is met vrij grote zekerheid te zeggen dat het netto-bedrijfsresultaat daar in 1986 lager lag dan in de overige gebieden.

1) De formule van Kish and Frankel wijkt enigszins af van de formule die ontstaat door sommatie van afzonderlijke varianties volgens (B9.27). Het - in de regel kleine - verschil is

$$\sum_{h=1}^H 2 * \frac{(N_h - n_h)}{N_h} * \frac{1}{n_h - 1} * \frac{\hat{A}_h}{\hat{A}} * (\bar{y}_{ah} - \bar{y}_a) * \frac{\hat{B}_h}{\hat{B}} * (\bar{y}_{bh} - \bar{y}_b)$$

14. Alternatieve schatters

14.1 Inleiding

In het voorafgaande is er stilzwijgend vanuit gegaan dat steekproefuitkomsten geschat worden met een directe schatter. Soms kan het nuttig zijn om gebruik te maken van een andere schatter om daardoor systematische of toevallige fouten te verkleinen. In dit hoofdstuk komen de quotiëntschatting en de regressieschatting aan de orde. Deze schatters maken gebruik van aanvullende informatie die betrekking heeft op alle populatie-elementen.

14.2 Quotiëntschatting

14.2.1 Inleiding

In hoofdstuk 12 zijn berekeningen uitgevoerd voor de schatting van quotiënten. De in paragraaf 12.7 gebruikte quotiëntschatting heeft echter bredere toepassingsmogelijkheden. Ook voor het schatten van een populatie-totaal of -gemiddelde van *niet-quotiënten* (absolute grootheden) kan de quotiëntschatting ("ratio estimator") gebruikt worden. Deze schatter heeft soms voordelen indien van een bepaalde variabele y een gemiddelde \bar{Y} of een totaal Y geschat moet worden en er daarnaast een hulpvariabele x is waarvoor de waarden x_i voor alle elementen in de populatie bekend zijn, of waarvan in ieder geval het totaal X van de waarden x_i bekend is. Vaak gaat het daarbij om een vroegere waarde van y_i uit een volledige telling van de populatie (Cochran, 1977:150-188).

Toepassing van een quotiëntschatting is interessant indien het quotiënt y/x in de populatie minder varieert dan de doelvariabele y . Door eerst het quotiënt \bar{Y}/\bar{X} te schatten en dit te vermenigvuldigen met het bekende populatiegemiddelde \bar{X} , ontstaat een schatting die een grotere nauwkeurigheid kan hebben dan de directe schatting. Of die grotere nauwkeurigheid ook daadwerkelijk gerealiseerd wordt, hangt af van de correlatie tussen y en x .

De quotiëntschatting is in tegenstelling tot de directe schatter, geen zuivere schatter. De onzuiverheid is echter van de orde $(1/n)$ en vermindert dus bij een groter wordende steekproef.

14.2.2 Enkelvoudige steekproeven

De quotiëntschatting wordt in enkelvoudige steekproeven gedefinieerd als:

$$\bar{y}_r = \bar{y} * \left(\frac{\bar{X}}{x}\right) = \hat{R} * \bar{X} \quad (14.2.1)$$

waarbij

$$\hat{R} = \bar{y} / \bar{x}$$

De variantie van de quotiënt-schatter bij een enkelvoudige steekproef kan uit de steekproef geschat worden als:

$$\text{var}(\bar{y}_r) = \frac{1-(n/N)}{n} * \frac{\sum_{i=1}^n (y_i - \hat{R} * x_i)^2}{n-1} \quad (14.2.2)$$

Aangetoond kan worden dat de quotiënt-schatter een kleinere variantie heeft dan de directe schatter indien (Cochran, 1977:158):

$$r > 0,5 * \frac{(S_x / \bar{X})}{(S_y / \bar{Y})} \quad (14.2.3)$$

waarbij:

r = de correlatie-coëfficiënt van y en x
 S_x = de standaarddeviatie van x

14.2.3 Gestratificeerde steekproeven

Bij een gestratificeerde steekproef zijn twee quotiënt-schatters mogelijk, namelijk een schatter waarbij per stratum het quotiënt van y en x wordt berekend (de "separate ratio estimator") en een schatter waarbij voor de hele populatie een quotiënt wordt berekend (de "combined ratio estimator") (Cochran, 1977: 164-169). De separate schatter wordt berekend als:

$$\bar{y}_{rs} = \sum_{h=1}^H W_h * \hat{R}_h * \bar{X}_h \quad (14.2.4)$$

en de gecombineerde schatter als:

$$\bar{y}_{rc} = \sum_{h=1}^H W_h * \hat{R} * \bar{X}_h \quad (14.2.5)$$

De variantie van de gecombineerde schatter is:

$$\text{var}(\bar{y}_{rc}) \approx \frac{1}{N^2} * \sum_{h=1}^H \frac{N_h * (N_h - n_h)}{n_h * (N_h - 1)} * \sum_{i=1}^{n_h} [(y_{hi} - \bar{y}_h) - \hat{R} * (x_{hi} - \bar{x}_h)]^2$$

(14.2.6)

De variantie van de separate schatter is gelijk aan die van de gecombineerde schatter, met uitzondering van de factor \hat{R} , die vervangen wordt door \hat{R}_h (het quotiënt in stratum h).

14.2.4 Quotiëntschattingen voor de nationale boerderij

Tabel 14.1 geeft voor een aantal grootheden de standaardfout van de quotiëntschatter. Daarbij worden de gewogen quotiëntschatter (14.2.1) (kolom 4) en de separate en de gecombineerde schatter (kolommen 5 en 6) vergeleken met de directe schatter. Voor de directe schatter is zowel de standaardfout die rekening houdt met de stratificatie (kolom 3) als de standaardfout die daar geen rekening mee houdt (kolom 2) - en die in de praktijk veel gebruikt wordt - weergegeven.

Tabel 14.1 geeft aan dat gebruikmaking van hulpinformatie via toepassing van een quotiëntschatter (\bar{y}_R), kan leiden tot een grotere betrouwbaarheid van schattingen dan toepassing van de directe schatter (\bar{y}) bij een enkelvoudige steekproef. Met name de hulpvariabele "sbe in metelling" geeft goede resultaten vanwege de hoge correlatie met de te schatten variabelen. De winst van het gebruik van \bar{y}_{rc} of \bar{y}_{rs} ten opzichte van \bar{y}_{st} bij een gestratificeerde steekproef is echter in de meeste gevallen niet groter dan een vermindering van de standaardfout met ongeveer 10%.

Toepassing van de hulpvariabele "hectare cultuurgrond in metelling" levert voor de gemiddelde nationale boerderij weinig winst op ten opzichte van de directe schatter. Blijkbaar is het verband tussen hulp- en doelvariabele te gering voor een bruikbare toepassing van de quotiëntschatter. Toepassing van deze schatter voor bijvoorbeeld het gemiddelde akkerbouwbedrijf leidt vrijwel zeker tot betere resultaten.

De tabel laat ook zien dat toepassing van de quotiëntschatter voor de schatting van de "arbeidskosten" wel leidt tot een grotere betrouwbaarheid bij gebruik van de hulpvariabele "sbe" en niet bij gebruik van "hectare cultuurgrond".

14.2.5 Quotiëntschatters en representativiteit

In de vorige subparagraaf bleek dat een quotiëntschatter een grotere betrouwbaarheid van schattingen kan bewerkstelligen. In sommige gevallen vergroot een quotiëntschatter ook de representativiteit. Dat vindt zijn oorzaak in de correctiefactor (X/\bar{x}) die

Tabel 14.1 Overzicht van standaardfouten van geschatte gemiddelden van een aantal absolute grootheden naar wijze van schatting (gemiddelde nationale boerderij; boekjaar 1986)

Variabele	Schatter				
	directe schatter		quotiënt-schatter		
	\bar{y}	\bar{y}_{st}	\bar{y}_r	\bar{y}_{rc}	\bar{y}_{rs}
Hulpvariabele quotiëntschatter: "sbe in meitelling"					
Arbeidskosten	1.188	903	1.370	884	863
Rente	331	158	145	128	125
Netto pacht	521	320	300	283	280
Werk door derden	163	96	146	88	86
Factorkosten	1.845	1.015	1.420	921	893
Niet-factorkosten	5.941	3.268	4.028	2.893	2.905
Kosten	7.219	3.681	4.173	3.169	3.179
Opbrengsten	7.739	3.874	4.211	3.294	3.354
Netto bedr. res.	1.583	1.409	1.991	1.355	1.305
Arbeidsopbrengst	1.877	1.263	1.332	1.133	1.124
Factoropbrengst	2.525	1.362	1.348	1.159	1.163
VAK*10	0,18	0,15	0,20	0,14	0,14
Cultuurgrond (aren)	62	20	49	18	17
Melkkoeien	1,0	0,4	0,9	0,3	0,3
Hulpvariabele quotiëntschatter: "ha cultuurgrond in meitelling"					
Waarde werktuigen	2.891	1.722	4.803	2.824	3.265
Sbe	4,1	1,5	4,2	1,6	2,5
Sbe (excl. int.v.)	0,9	1,0	2,3	1,3	1,3
Kosten arbeid	1.188	903	2.372	1.031	1.298
Werk door derden	163	96	273	191	196
Pacht	521	320	876	555	667
Opbrengsten	7.739	3.874	9.481	4.047	5.696
Netto bedr. res.	1.583	1.409	2.169	1.385	1.439
Aren cultuurgrond	62	20	14	10	10

wordt toegepast op de directe schatting \bar{y} . Indien y systematisch te groot geschat zou worden door oververtegenwoordiging in de steekproef van bedrijven waar y_i groot is, en er een positief verband is tussen x_i en y_i op een bedrijf, zorgt de correctie-term ervoor dat y op "populatie-niveau" wordt gebracht (zie ook Bethlehem en Kersten, 1986).

Als voorbeeld van deze representativiteit-vergroten- de functie van quotiëntschat- ters kan de schatting van de totale gepachte oppervlakte cultuurgrond dienen. Omdat op steekproefbedrijven relatief meer gepachte grond voorkomt dan op grond van een asele- cte steekproeftrekking verwacht mag worden (paragraaf 11.2.8), wordt de betreffende oppervlakte overschat door de steekproef. Toepas- sing van een quotiëntschatter met als hulpvariabele de oppervlak- te pacht in de metelling, leidt tot een grotere betrouwbaarheid en minder systematische vertekening van de schatting van de ge- pachte oppervlakte volgens het boekhoudnet.

De directe schatter geeft voor boekjaar 1985 aan dat de ge- middeld gepachte hoeveelheid grond 1.200 aren bedraagt. De bijbe- horende standaardfout is 37 aren (zonder stratificatie 50 aren). Toepassing van een quotiëntschatter met als hulpvariabele de hoe- veelheid gepachte grond die is opgegeven in de metelling levert een gemiddelde oppervlakte van $1.200 * (869/912)$ aren met een standaardfout van 22 bij toepassing van de gecombineerde quo- tiëntschatter 1).

In plaats van een 90%-betrouwbaarheidsinterval van ongeveer [1140,1265] levert de quotiëntschatter een interval van ongeveer [1110,1185].

14.2.6 Quotiëntschatters als alternatief voor stratificatie

In paragraaf 14.2.4 is voor een aantal variabelen onderzocht in hoeverre bij een enkelvoudige en bij een gestratificeerde steekproef een quotiëntschatter betere resultaten oplevert dan een directe schatter. Daarbij blijkt ook dat in sommige gevallen een quotiëntschatter bij een enkelvoudige steekproef soms betere resultaten oplevert dan een directe schatter bij een gestratifi- ceerde steekproef. De schatting van de rente-kosten is daar een voorbeeld van. De directe schatter bij een gestratificeerde steekproef levert een standaardfout van 158 gulden terwijl de quotiëntschatter bij een enkelvoudige steekproef een standaard- fout van 145 gulden geeft (tabel 14.1). Het voorbeeld in para- graaf 14.2.5 laat een soortgelijke situatie zien.

Daarmee is aangegeven dat het soms niet noodzakelijk is een (post-)stratificatie uit te voeren, omdat met een quotiëntschat- ter hetzelfde resultaat te bereiken is. De quotiëntschatter heeft als voordeel dat alleen \bar{X} bekend hoeft te zijn voor de populatie

1) Bij toepassing van de quotiëntschatter (14.2.1) is de stan- daardfout 29 en bij toepassing van de separate schatter 26.

en niet de aantallen per stratum. Voor de quotiëntschatter moet daarentegen wel de waarde van de hulpvariabele voor alle steekproefbedrijven bekend zijn.

14.3 Regressieschatters

14.3.1 Inleiding

Evenals de quotiëntschatter maakt de regressieschatter gebruik van een hulpvariabele die hoog gecorreleerd is met de variabele waarvoor een gemiddelde of totaal geschat moet worden. Terwijl de quotiëntschatter geschikt is voor variabelen die in een relatief vaste verhouding tot een hulpvariabele staan, is de regressieschatter ook geschikt voor andere lineaire verbanden, dat wil zeggen voor relaties die niet "door de oorsprong gaan". Bij de regressieschatter wordt namelijk gebruik gemaakt van een additieve correctie-term in plaats van een multiplicatieve correctie-factor (Cochran, 1977:189-204).

14.3.2 Enkelvoudige steekproeven

De lineaire regressie-schatter (\bar{y}_{1r}) van het populatiegemiddelde \bar{Y} is bij een enkelvoudige steekproef gelijk aan:

$$\bar{y}_{1r} = \bar{y} + b * (\bar{X} - \bar{x}) \quad (14.3.1)$$

waarbij:

b = een steekproefschatting van de verandering in y als x een eenheid toeneemt

\bar{X} = het populatie-gemiddelde van de hulpvariabele

\bar{x} = het steekproefgemiddelde van de hulpvariabele

\bar{y} = het steekproefgemiddelde van de te schatten variabele

Deze schatter is evenals de quotiëntschatter onzuiver. De onzuiverheid is van de orde $(1/n)$.

De directe schatter en de quotiëntschatter zijn bijzondere gevallen van de regressieschatter, namelijk voor $b=0$ en $b=(\bar{y}/\bar{x})$.

In een enkelvoudige steekproef geldt dat de variantie van \bar{y}_{1r} minimaal wordt indien voor b de lineaire regressie-coëfficiënt B van y op x in de eindige populatie wordt genomen.

$$B = \frac{\sum_{i=1}^N (y_i - \bar{Y}) * (x_i - \bar{X})}{\sum_{i=1}^N (x_i - \bar{X})^2} \quad (14.3.2)$$

In de praktijk is B meestal onbekend. In paragraaf 14.4 zal worden ingegaan op de situatie waarin b bekend is of bekend verondersteld wordt. In deze paragraaf wordt er van uitgegaan dat B geschat moet worden uit de steekproef. Als steekproefschatting van B wordt de kleinste kwadratenschatting b genomen, waarbij b verkregen wordt door in vergelijking 14.3.2 N te vervangen door n en de populatie-gemiddelden door steekproefgemiddelden. Dan geldt:

$$\text{Var}(\bar{y}_{1r}) = \frac{1-f}{n} * \sum_{i=1}^N (\bar{y}_i - \bar{y})^2 * (1 - r^2) \quad (14.3.4)$$

waarbij:

r = de correlatie-coëfficiënt tussen y en x in de populatie

Voor grote steekproeven kan de variantie van \bar{y}_{1r} geschat worden als (Cochran, 1977:195):

$$\text{var}(\bar{y}_{1r}) = \frac{1-f}{n*(n-2)} * \sum_{i=1}^n [(y_i - \bar{y}) - b*(x_i - \bar{x})]^2 \quad (14.3.5)$$

De regressieschatting is in grote steekproeven precieser dan de quotiëntschatting tenzij de regressielijn van y op x door de oorsprong gaat. De regressieschatting is in grote steekproeven altijd precieser dan de directe schatting.

14.3.3 Gestratificeerde steekproeven

Evenals bij de quotiëntschatting kan bij de regressieschatting in een gestratificeerde steekproef onderscheid worden gemaakt tussen een separate en een gecombineerde schatting. In het eerste geval wordt per stratum een afzonderlijke regressie-schatting gemaakt waarna deze schattingen tot een totaalschatting worden geaggregeerd via de ophoogfactoren. Deze separate schatting yl_{1r} wordt dus berekend als:

$$\bar{y}_{lrs} = \sum_{h=1}^H \frac{N_h}{N} * \bar{y}_{lrh} \quad (14.3.6)$$

waarbij:

$$\bar{y}_{lrh} = \bar{y}_h + b_h * (\bar{x}_h - \bar{x}_h)$$

b_h = de steekproefschatting voor de lineaire regressiecoëfficiënt in stratum h

Dat levert een variantie:

$$\text{var}(\bar{y}_{lrs}) = \text{factor} * \left[\sum_{i=1}^{n_h} (y_{hi} - \bar{y}_h)^2 - b^2_h * \sum_{i=1}^{n_h} (x_{hi} - \bar{x}_h)^2 \right] \quad (14.3.7)$$

waarbij:

$$\text{factor} = \sum_{h=1}^H \left(\frac{N_h}{N} \right)^2 * \left(\frac{1-f_h}{n_h} \right) * \left(\frac{1}{n_h-2} \right)$$

Bij de gecombineerde schatter wordt gebruik gemaakt van de coëfficiënt

$$b_c = \text{factor1} * \frac{\sum_{i=1}^{n_h} (y_{hi} - \bar{y}_h) * (x_{hi} - \bar{x}_h)}{\sum_{h=1}^{n_h} (x_{hi} - \bar{x}_h)^2}$$

waarbij:

$$\text{factor1} = \sum_{h=1}^H \left(\frac{N_h}{N} \right)^2 * \left(\frac{1-f_h}{n_h} \right) * \left(\frac{1}{n_h-1} \right)$$

De variantie van de gecombineerde regressieschatter

$$\bar{y}_{lrc} = \bar{y}_{st} + b_c * (\bar{X} - \bar{x}_{st}) \quad (14.3.8)$$

is te berekenen als (Cochran, 1977):

$$\text{var}(\bar{y}_{lrc}) = \text{factor1} * \left[\sum_{i=1}^{n_h} (y_{hi} - \bar{y}_h) - b_c * \sum_{i=1}^{n_h} (x_{hi} - \bar{x}_h) \right]^2 \quad (14.3.9)$$

14.4 Verschilsschatters

14.4.1 Inleiding

Een schatter die nauw verwant is met de regressie-schatter is de verschilsschatter. Deze schatter is gelijk aan de regressieschatter in vergelijking (14.3.1), maar gaat uit van een vaste coëfficiënt b . De schatter heet daarom ook wel "regression estimator with pre-assigned slope" (Cochran, 1977:190). Deze schatter is zuiver en kan daarom bij kleine steekproeven de voorkeur verdienen boven een quotiënt- of regressieschatter.

14.4.2 Enkelvoudige steekproeven

De variantie van de verschilsschatter

$$\bar{y}_{1r} = \bar{y} + b_0 * (\bar{X} - \bar{x}) \quad (14.4.1)$$

is gelijk aan

$$\text{Var}(\bar{y}_{1r}) = \frac{1-f}{n} * \frac{\sum_{i=1}^N [(y_i - \bar{Y}) - b_0 * (x_i - \bar{X})]^2}{N-1} \quad (14.4.2)$$

Een zuivere steekproefschatting $\text{var}(\bar{y}_{1r})$ voor $\text{Var}(\bar{y}_{1r})$ wordt verkregen door in (14.4.2) N te vervangen door n .

De waarde van b_0 die $\text{Var}(\bar{y}_{1r})$ minimaliseert is de lineaire regressiecoëfficiënt B van y op x in de eindige populatie (zie paragraaf 14.3). De resulterende minimale variantie is gelijk aan:

$$\text{Var}_{\min}(\bar{y}_{1r}) = \frac{1-f}{n} * \sum_{i=1}^N (y_i - \bar{Y})^2 * (1 - r^2) \quad (14.4.3)$$

Hoewel B vaak maar bij benadering bekend is, levert dit in het algemeen geen grote problemen op. De variantie wordt in het algemeen maar weinig vergroot indien in plaats van B een daar iets van afwijkende waarde voor b_0 wordt gekozen.

14.4.3 Gestratificeerde steekproeven

Evenals bij de schatters in de paragrafen 14.2 en 14.3 is er bij gestratificeerde steekproeven sprake van twee schatters \bar{y}_{1rs} en \bar{y}_{1rc} . Als b_h in paragraaf 14.3.3 van tevoren gekozen wordt en niet uit de steekproef geschat worden, dan geldt dat

$$\text{Var}(\bar{y}_{1rs}) = \text{factor} * \left[\sum_{i=1}^{n_h} (y_{hi} - \bar{y}_h) - b_h * (x_{hi} - \bar{x}_h) \right]^2 \quad (14.4.4)$$

De variantie is weer minimaal als b_h gelijk is aan B_h , waarbij B_h de werkelijke regressiecoëfficiënt in stratum h is.

De formule voor de variantie van de gecombineerde schatter is analoog aan (14.4.4), waarbij b_h vervangen wordt door een gemeenschappelijke coëfficiënt b_c voor alle strata. De variantie van deze schatter is minimaal als b_c gelijk is aan een bepaald gewogen gemiddelde B_c van de werkelijke regressiecoëfficiënten B_h in stratum h (Chochran, 1977:201). Bij optimale keuze van b_h en b_c - dat wil zeggen gelijk aan B_h en B_c - heeft de separate schatter een kleinere variantie dan de gecombineerde schatter.

14.5 Algemene regressieschatters

In paragraaf 14.3 is de lineaire regressieschatter besproken. Deze schatter maakt gebruik van een hulpvariabele om zo tot meer zuivere en nauwkeurige schattingen te komen. Deze regressieschatter kan gegeneraliseerd worden door tegelijkertijd informatie van meerdere variabelen te gebruiken.

Is de informatie van de hulpvariabelen kwantitatief van aard dan ontstaat een algemene vorm van de schatter uit paragraaf 14.3. Deze algemene regressieschatter kent op impliciete wijze gewichten toe aan de waarnemingen. Bethlehem en Kersten (1986) tonen aan dat gewichten die worden toegekend aan een steekproefelement, opgebouwd zijn als een lineaire combinatie van gewichtscoëfficiënten die worden bijgedragen door de diverse hulpvariabelen.

Is de informatie van de hulpvariabelen kwalitatief van aard dan kent de gegeneraliseerde regressie-schatter eveneens impliciet gewichten toe aan bedrijven. Het blijkt dat toepassing van de algemene regressieschatter met behulp van kwalitatieve hulpvariabelen overeenkomt met het toepassen van een algemene vorm van post-stratificatie. Bij toepassing van de algemene regressieschatter kunnen meer stratificatie-variabelen aan bod komen dan bij post-stratificatie (Bethlehem en Keller, 1987). Nader onderzoek naar de toepassingsmogelijkheden voor de LEI-steekproef verdient zeker aanbeveling.

Algemene regressie-schatters zijn geen zuivere schatters. Ze zijn onder bepaalde voorwaarden wel consistent 1). De variantie van de algemene regressieschatter is niet exact weer te geven in een formule. Evenals voor de lineaire regressieschatter zijn wel benaderingen voor grote steekproeven bekend.

Naast lineaire regressieschatters zijn ook niet-lineaire regressieschatters ontwikkeld. Deze methode van schatten wordt ook

1) De onzuiverheid en de variantie van een consistente schatter naderen tot 0 als de steekproefgrootte naar oneindig gaat.

wel "raking" of "multiplicatief wegen" genoemd; dit in tegen tegenstelling tot het additief wegen van de algemene lineaire regressieschatter. Het schatten van varianties bij deze schattingsmethode levert echter problemen op (Bethlehem en Kersten (1986); Brackstone and Rao (1979)).

14.6 Schatters voor kleine deelpopulaties

De LEI-landbouw-steekproef heeft een zodanige omvang dat voor een groot aantal groepen binnen de populatie betrouwbare schattingen gegeven kunnen worden (hoofdstuk 13). Voor zeer kleine deelpopulaties waarvan maar enkele bedrijven in de steekproef vallen - bijvoorbeeld de bedrijven in een bepaalde gemeente - zijn meestal geen betrouwbare steekproef-schattingen te geven.

Een onderzoekgebied dat zich speciaal richt op schattingen voor kleine deelpopulaties is dat van de "small domain estimation" (Platek et al, 1987; Purcell and Kish, 1979). Op dit terrein worden alternatieve (synthetische) schatters ontwikkeld die niet alleen gebruik maken van het kleine aantal steekproefelementen in de deelpopulatie, maar ook van de elementen in andere deelpopulaties.

Een voorbeeld van het gebruik van synthetische schatters is te vinden in Wijnands et al. (1983). In deze studie worden op basis van gegevens van steekproefbedrijven regressievergelijkingen geschat die het verband aangeven tussen meetingsgegevens en kunstmestgiften per hectare. Door in deze regressie-vergelijkingen bedrijfskenmerken uit de meting in te vullen voor alle bedrijven in de populatie, was het mogelijk om schattingen voor de kunstmestgiften per hectare op gemeente-niveau te maken; ook voor gemeenten waarvoor de steekproef geen enkel bedrijf bevatte.

Een dergelijke benadering staat of valt met de aanname dat het gebruikte model - de regressievergelijkingen - dat geschat is op basis van een grotere deelpopulatie ook geldig is voor kleine deelpopulaties. De schattingen zijn "model-dependent" in tegenstelling tot "design-consistent" schattingen die geen model gebruiken (Särndal, 1984).

Een andere toepassing op het LEI is het schatten van arbeidsopbrengsten voor gebieden waar plannen met betrekking tot landinrichting bestaan. Deze gebieden zijn meestal relatief klein waardoor er maar enkele steekproefbedrijven in het betreffende gebied voorkomen. Daarom wordt de kleine deelpopulatie verdeeld in bedrijfstype-sbe-klassen waarbij iedere klasse gerepresenteerd wordt door er in thuishorende steekproefbedrijven uit vergelijkbare gebieden.

Nader onderzoek op dit terrein verdient aanbeveling gezien de vraag die er is naar gegevens over kleinere deelpopulaties en de beschikbaarheid van veel gegevens over de populatie in de vorm van de jaarlijkse metingen.

15. Het boekhoudnet als roterend panel

15.1 Inleiding

De keuze van de bedrijven voor het boekhoudnet gebeurt niet ieder jaar opnieuw (zie hoofdstuk 8). Aan het begin van ieder boekjaar wordt 20 à 25% van de bedrijven die in administratie zijn, vervangen door nieuwe bedrijven. Op de consequenties van deze werkwijze wordt in dit hoofdstuk ingegaan.

15.2 Bedrijven niet langer dan zes jaar in administratie

Ieder jaar wordt aan de hand van de meest recente metelling die beschikbaar is, de gewenste samenstelling van de steekproef vastgesteld. Dat betekent niet dat er ieder jaar een volledig nieuwe samenstelling van de steekproef gekozen wordt. De bedrijfskeuze richt zich op het aanvullen van de steekproef in verband met de uitval sinds de vorige bedrijfskeuze, en het verversen van de steekproef om structuurveranderingen in de Nederlandse landbouw op te vangen (Lodder, 1987:9-10). Uitgangspunt daarbij is dat bedrijven in de regel niet langer dan zes jaar in administratie worden gehouden. Omdat jaarlijks ongeveer 10% van de bedrijven die in administratie zijn, wegens uitval niet voor een volgend boekjaar beschikbaar is, zal in de praktijk ieder jaar ongeveer 25% van de bedrijven vervangen worden door nieuwe bedrijven die steekproefsgewijs getrokken worden.

Door het langer dan één jaar in administratie houden van bedrijven, worden mogelijkheden tot het uitvoeren van tijdreeksanalyses geschapen. Bovendien wordt daarmee de invloed van het steekproeftoeval bij het opstellen van voorlopige uitkomsten en prognoses verminderd (zie paragraaf 15.4 en 15.5). Daar staat tegenover dat door de termijn van zes jaar niet geheel kan worden voorkomen dat er enige leereffecten bij de deelnemers aan het boekhoudnet op zouden kunnen treden. Een voordeel van de wat langere termijn is echter dat niet ieder jaar alle bedrijven opnieuw aangezocht hoeven te worden met alle kosten die daaraan verbonden zijn. Ook de kwaliteit van de gegevens van bedrijven die langer meedoen, zal in het algemeen groter zijn dan van bedrijven die net starten met deelname.

15.3 Jaarlijkse vervanging afhankelijk van doelstellingen

15.3.1 Inleiding

Het boekhoudnet van landbouwbedrijven is een verzameling van jaarlijkse steekproeven. Gegeven die reeks van steekproeven zijn

er drie soorten grootheden waarvoor een schatting gemaakt zou kunnen worden (Cochran, 1977:344-345):

- 1) de verandering in de gemiddelde waarde \bar{Y} van een variabele y tussen twee steekproeven;
- 2) de gemiddelde waarde van \bar{Y} over een aantal steekproeven samen;
- 3) de gemiddelde waarde van \bar{Y} in elk van de steekproeven afzonderlijk.

Tot nu toe is in dit onderzoeksverslag vooral de aandacht gelegd bij de derde grootheid. Afgezien van steekproefuitval en leereffecten, kan - vanuit het gezichtspunt van maximale betrouwbaarheid van schattingen - het volgende worden opgemerkt over de optimale vervangingsstrategie van bedrijven in de steekproef:

- 1) voor het schatten van veranderingen kan de gehele steekproef het beste gehandhaafd blijven;
- 2) voor het schatten van een gemiddelde over een aantal jaren samen, is het jaarlijks trekken van een nieuwe steekproef optimaal;
- 3) voor schattingen per jaar hangt het van een aantal factoren af wat de beste strategie is.

15.3.2 Het schatten van veranderingen

Bij het schatten van veranderingen geldt dat de geschatte variantie van een verandering

$$d = (\bar{x} - \bar{y}) \quad (15.3.1)$$

bij gelijkblijvende bedrijven gelijk is aan

$$S^2_x + S^2_y - 2R_{xy}S_xS_y \quad (15.3.2)$$

en bij verschillende bedrijven gelijk is aan

$$S^2_x + S^2_y \quad (15.3.3)$$

waarbij:

- \bar{x} = de waarde van de variabele in de ene steekproef
- \bar{y} = de waarde van de variabele in de andere steekproef
- R_{xy} = de correlatie-coëfficiënt tussen x en y
- S^2_x = de variantie van x

Aangezien R_{xy} bijna altijd positief is, is een zo gering mogelijke vervanging de beste strategie. Dan wordt immers een maximaal gebruik gemaakt van de positieve correlatie tussen waarnemingen.

Kish (1965) geeft aan wat de variantie van een geschat verschil is bij een gedeeltelijke overlap van twee steekproeven.

$$\text{var}(d) = \frac{S^2_x}{n_x} + \frac{S^2_y}{n_y} - \frac{2 * R_{xy} * S_x * S_y * P_x * P_y}{n_c} \quad (15.3.4)$$

waarbij:

n_x = aantal bedrijven in de ene steekproef
 n_y = aantal bedrijven in de andere steekproef
 n_c = aantal bedrijven dat in beide steekproeven voorkomt
 $P_x = (n_c/n_x)$
 $P_y = (n_c/n_y)$

In deze formule is afgezien van de eindigheidcorrectie $(1-(n/N))$. Duidelijk wordt dat de winst die voortvloeit uit de overlap van de twee steekproeven, afhankelijk is van de correlatiecoëfficiënt en het percentage overlap.

Bij gelijkblijvende steekproefgrootte $n=n_x=n_y$ - met daaruit voortvloeiend $P=P_x=P_y$ - en gelijkblijvende variantie in de populatie luidt vergelijking 15.3.4:

$$\text{var}(d) = \frac{2}{n} * S^2 * (1 - P * R_{xy}) \quad (15.3.5)$$

De winst ten opzichte van de situatie zonder een overlappend steekproefdeel is $(P * R_{xy}) * 100\%$. Bij een overlap van 80% - het percentage dat ongeveer geldt voor het LEI-boekhoudnet - en een correlatie-coëfficiënt van 60% tussen waarden in opeenvolgende boekjaren, blijkt de variantie met 50% gereduceerd te worden 1). Dat betekent een reductie van 30% in de standaardfout.

15.3.3 Het schatten van een meerjarig gemiddelde

Bij het schatten van een "overall" gemiddelde voor twee jaren, is de variantie op basis van gelijkblijvende steekproefbedrijven

$$(S^2_x + S^2_y + 2 * R_{xy} * S_x * S_y) / 4$$

- 1) Anders gezegd, de variantie van het verschil is niet de som van de varianties van de afzonderlijke schattingen, maar slechts de helft daarvan. Voor schattingen voor bijvoorbeeld de nationale boerderij kan de standaardfout voor het verschil voor een bepaalde variabele tussen twee boekjaren dus rechtstreeks afgelezen worden uit tabel 12.2.

terwijl deze

$$(S^2_x + S^2_y) / 4$$

is bij nieuw gekozen elementen. Het ligt dus voor de hand om voor deze schatting ieder jaar opnieuw een steekproef te kiezen omdat R_{xy} vrijwel altijd groter dan 0 zal zijn.

15.3.4 Het schatten van gemiddelden voor een bepaald jaar

Over geval drie valt iets meer te zeggen (Cochran, 1977: 346-355). Indien gebruik gemaakt wordt van een directe schatter die alleen de gegevens van de bedrijven in het betreffende boekjaar gebruikt, maakt de wijze van vervanging niet uit. Indien echter een "double sample regression estimate" 1) voor de blijvende bedrijven gebruikt wordt - waarbij de waarneming in het vorige boekjaar de hulpvariable is - en een gewone schatting voor de nieuwe bedrijven, en deze schattingen op een bepaalde manier worden gewogen tot een schatting \bar{x}_2 , dan is de variantie van de gemiddelde waarde \bar{x}_2 gelijk aan

$$\text{Var}(\bar{x}_2) = S^2_2 * \left(\frac{n - u * r^2}{n^2 - u^2 * r^2} \right) \quad (15.3.3)$$

waarbij

u = het aantal bedrijven in de steekproef dat vervangen wordt

$\text{Var}(\bar{x}_2)$ bereikt een minimale waarde voor:

$$u/n = \frac{1}{1 + \sqrt{(1-r^2)}}$$

waarbij:

r = de correlatie-coëfficiënt tussen de vorige en de nieuwe waarneming bij blijvende bedrijven

De winst ten opzichte van complete vervanging (u=n) of complete handhaving (u=0) is dan gelijk aan

1) Een regressie-schatting waarbij de hulpinformatie afkomstig is uit een andere steekproef in plaats van uit de populatie.

$$\frac{1 + \sqrt{(1-r^2)}}{2}$$

Voor $r=0,5$ komt dat neer op 7%, voor $r=0,8$ op 25% en voor $r=0,95$ op 52%. De winst is maximaal 100% voor $r=1$.

De schatting \bar{x}_2 wordt als volgt berekend:

$$\bar{x}_2 = \left(\frac{W_{2v}}{W_{2b} + W_{2v}} \right) * \bar{x}_{2v} + \left(1 - \frac{W_{2v}}{W_{2b} + W_{2v}} \right) * \bar{x}_{2b} \quad (15.3.7)$$

waarbij W de inverse variantie ($1/S$) aanduidt en de indices b en v betrekking hebben op blijvende respectievelijk vervangen bedrijven.

Voor het schatten van een gemiddelde voor een bepaald boekjaar kan ook een splitsing gemaakt worden in bedrijven die 1, 2, 3, 4 of meer jaren meedoen in de steekproef. Ook kunnen schattingen van een voorgaand jaar herzien worden op basis van informatie uit het lopende boekjaar. Cochran geeft aan dat de reductie in de varianties van schattingen bescheiden zullen zijn indien r kleiner is dan 0,8. Als r groter of gelijk aan 0,95 is, zijn reducties van 50% mogelijk. Bij een optimale vervangingsstrategie wordt in het algemeen meer dan de helft van de bedrijven vervangen.

De reducties in varianties bij het schatten van *verschillen* (veranderingen tussen jaren) kunnen bij $r > 0,7$ fors zijn, met name indien weinig bedrijven vervangen worden (zie paragraaf 15.3.1). Cochran (1977:354) doet daarom de suggestie dat een vervanging van 1:3, 1:4 of 1:5 tussen twee opeenvolgende jaren een goede praktische regel kan zijn indien schattingen voor afzonderlijke steekproeven ("current estimators") en schattingen van verschillen tussen steekproeven beide belangrijk zijn. De LEI-steekproef is in dit opzicht dus redelijk optimaal. Tot nu toe worden de in deze subparagraaf genoemde schatters voor afzonderlijke steekproeven echter nauwelijks gebruikt.

15.4 Voorlopige uitkomsten op basis van blijvende bedrijven

15.4.1 Inleiding

De definitieve uitkomsten van een bepaald boekjaar zijn in het algemeen pas een jaar na afloop van het boekjaar bekend, omdat er op dat moment genoeg bedrijven zijn uitgewerkt om over te kunnen gaan tot publikatie. Om al in een eerder stadium een indicatie te kunnen geven van de bedrijfsuitkomsten, worden een half

jaar na afloop van een boekjaar voorlopige uitkomsten gegeven. Deze voorlopige uitkomsten worden ook gebruikt bij het opstellen van prognoses (paragraaf 15.5). Ze hebben betrekking op de arbeidsopbrengst per ondernemer en het netto bedrijfsresultaat (Prognoses, 1986).

15.4.2 Oude wegingsfactoren

De voorlopige uitkomsten worden berekend op basis van uitkomsten van ongeveer 60% van het totaal aantal bedrijven waarop de definitieve uitkomsten zullen zijn gebaseerd. Op het moment van het berekenen van de voorlopige uitkomsten staan de definitieve wegingsfactoren voor het betreffende boekjaar nog niet vast. De meetellingsaantallen per stratum aan het begin van het boekjaar zijn weliswaar bekend, maar de aantallen steekproefbedrijven per stratum met een uitgewerkte boekhouding zijn nog onbekend. De uitkomsten van deze ongeveer 600 bedrijven worden daarom gewogen met de wegingsfactor die ze in het vorige boekjaar hadden. Dat heeft tot gevolg dat bij de berekening van de voorlopige uitkomsten alleen die bedrijven opgenomen kunnen worden die langer dan 1 jaar in administratie zijn.

De uitkomsten van deze bedrijven worden per groep gecorrigeerd met een term die aangeeft in hoeverre ze in het vorige boekjaar qua uitkomsten afwaken van alle bedrijven met een uitgewerkte boekhouding. Als de arbeidsopbrengst op alle weidebedrijven in de zandgebieden vorig jaar 40.000 gulden bedroeg en op de bedrijven die meedoen in de voorlopige uitkomsten 42.000 gulden, dan worden de voorlopige uitkomsten van deze laatste groep met 2.000 gulden naar beneden bijgesteld 1).

Hoewel berekening op basis van "oude" wegingsfactoren in de praktijk redelijk werkt, blijkt bij analyse achteraf vaak dat verschillen tussen voorlopige en definitieve uitkomsten voor een aanzienlijk deel veroorzaakt worden door verschillen tussen voorlopige en definitieve wegingsfactoren. De wegingsfactoren die een bedrijf in twee opeenvolgende jaren toegekend krijgt, blijken in de praktijk nogal te kunnen verschillen, ook al zijn er geen grote verschuivingen in de populatie en ook al is het betreffende bedrijf niet of nauwelijks veranderd van karakter. De oorzaak daarvan ligt voornamelijk in de ver doorgevoerde stratificatie. Indien een bedrijf voor slechts een van de vijf stratificatiecriteria een andere waarde krijgt - bijvoorbeeld door een verandering van de bedrijfsgrootte van 19 naar 21 hectare - is er een grote kans dat het bedrijf van stratum veranderd. Indien er grote

1) In feite hebben we hier te maken met een gecombineerde verschilschatter met coëfficiënt $b=1$ (paragraaf 15.4). Voor toepassing van een separate verschilschatter zijn de steekproefaantallen per stratum in het algemeen te klein.

aantallen steekproefbedrijven per stratum zouden zijn, zou een dergelijke verandering de wegingsfactoren nauwelijks beïnvloeden. Maar vanwege de kleine aantallen per stratum, zijn de veranderingen vaak fors. Natuurlijk wordt bij de keuze van nieuwe bedrijven zoveel mogelijk geprobeerd om een evenwichtige stratumvulling te realiseren, maar door allerlei factoren is het aantal steekproefbedrijven dat uiteindelijk in een bepaald stratum terecht komt, enigszins aan het toeval onderhevig.

15.4.3 Voorlopige wegingsfactoren

Een alternatief zou zijn om de 600 bedrijven nieuwe wegingsfactoren toe te kennen op basis van de stratum-indeling die hoort bij het betreffende boekjaar. De verhouding tussen het meetlingsaantal en het aantal al uitgewerkte boekhoudingen per stratum is dan de voorlopige wegingsfactor voor dat stratum. Indien niet alleen op populatie-niveau 60% van de bedrijven zou zijn uitgewerkt, maar ook in ieder afzonderlijk stratum, zouden de voorlopige wegingsfactoren een goede relatieve schatting zijn voor de definitieve wegingsfactoren. In de praktijk verschillen de uitwerkingspercentages nogal van stratum tot stratum, waardoor met name bedrijven in strata met nog weinig uitgewerkte bedrijven, relatief veel sterker meewegen in de voorlopige uitkomsten dan in de definitieve uitkomsten.

In de LEI-praktijk wordt het alternatief van voorlopige ("variabele") wegingsfactoren alleen toegepast indien circa 80% van de steekproefboekhoudingen is uitgewerkt.

Een andere mogelijkheid zou zijn om op het moment dat de voorlopige uitkomsten berekend worden, een schatting te maken - per stratum - van het aantal bedrijven dat in de definitieve uitkomsten meegenomen zal worden. Op grond daarvan kunnen schattingen van de definitieve wegingsfactoren worden gemaakt. Deze mogelijkheid biedt waarschijnlijk meer perspectief dan de hiervoor genoemde.

15.4.4 Verschillen met definitieve uitkomsten

Voorlopige uitkomsten zullen door de rol van het toeval enigszins kunnen afwijken van de definitieve uitkomsten. De standaardfout van het verschil tussen voorlopige en definitieve uitkomsten is te berekenen met een soortgelijke formule als in paragraaf 15.3.1 voor elkaar gedeeltelijk overlappende steekproeven.

Voor een enkelvoudige steekproef geldt (Kish, 1965:460):

$$\text{var}(\bar{y}_{\text{voorl}} - \bar{y}_{\text{def}}) = \frac{S^2}{n_c} * (1 - P_x) = \frac{S^2}{n_x} * \left(\frac{1 - P_x}{P_x} \right) \quad (15.4.1)$$

waarbij

\bar{y}_{voorl} = schatting van \bar{Y} op basis van een deel van de steekproefbedrijven
 \bar{y}_{def} = schatting van \bar{Y} op basis van alle steekproefbedrijven

Dat wil zeggen dat de variantie van het betreffende verschil P_x procent lager is dan de standaardfout van het geschatte voorlopige gemiddelde, of $(1-P_x)/P_x$ procent van de standaardfout van het geschatte definitieve gemiddelde. Voor de praktijk van het LEI-boekhoudnet betekent dit, dat de standaardfout van het verschil ongeveer 60% ($100 * \sqrt{(1-0,6)}$ %) van de standaardfout van de voorlopige uitkomsten bedraagt. In termen van de standaardfout van de definitieve uitkomsten is dit ongeveer 80% (de wortel uit $100 * (1-0,6)/0,6$). Voor een gestratificeerde steekproef gelden overeenkomstige percentages.

Nog grotere winst zou behaald kunnen worden indien voor de definitieve uitkomsten ook gebruik gemaakt zou worden van een

Tabel 15.1 Gemiddelde arbeidsopbrengst bij voorlopige en definitieve uitkomsten naar BUL-groep (gemiddeld over de boekjaren 1982-1985)

BUL-groep	Voorlopige uitkomsten	Definitieve uitkomsten	Absoluut verschil
Grotere bedrijven			
Akkerbouw:			
- noordelijk klei	41.300	39.200	5.700
- centraal klei	61.800	71.400 *)	11.300
- zuidwestelijk klei	71.000	68.900	4.300
- veenkolonien	37.300	34.900	3.600
Weide:			
- nrd. klei/veenweide	41.300	43.900	3.500
- westelijk weide	29.800	31.100	4.400
- zandgebieden	38.000	38.400	4.000
Gemengd	26.300	24.200	5.300
Kleinere bedrijven			
Weide	10.800	9.000	3.100
Gemengd	7.100	3.000	5.600

*) Indien het extreme jaar 1983 buiten beschouwing wordt gelaten, is het verschil 4.800. In 1983 waren de voorlopige en de definitieve uitkomst respectievelijk 140.000 en 173.800.

Bron: LEI, Prognoses, diverse jaren.

verschilschatter, of van een "double sample regression estimate" (paragraaf 15.3.3).

In tabel 15.1 wordt inzicht gegeven in de verschillen tussen voorlopige en definitieve uitkomsten voor de onderscheiden BUL-groepen.

De verschillen tussen voorlopige en definitieve uitkomsten zijn niet groot. Desondanks zijn ze in het algemeen groter dan verwacht zou mogen worden op basis van de standaardfouten van de definitieve uitkomsten. De verwachting van het absolute verschil tussen een waarde uit een normale verdeling en de verwachte waarde uit die verdeling is ongeveer 40% van de standaarddeviatie 1). Vergelijking van 40% * 80% van de standaardfouten uit tabel 15.4 met de absolute fouten uit tabel 15.1 leert dat laatstgenoemde fouten hoger liggen. Het verschil moet verklaard worden uit het gebruik van oude wegingsfactoren op het moment dat de voorlopige uitkomsten berekend worden, en uit het gebruik van een verschilschatter bij het berekenen van de voorlopige uitkomsten. Zou bij de berekening van de definitieve uitkomsten ook een dergelijke verschilschatter gebruikt worden (zie paragraaf 15.3.3) dan zou het verschil tussen voorlopige en definitieve uitkomsten waarschijnlijk kleiner zijn. Bij de correctie van de blijvende bedrijven ten opzichte van alle bedrijven op het moment van het berekenen van de voorlopige uitkomsten wordt namelijk ook de toevalscomponent in het verschil weg-gecorrigeerd.

Worden alle voorlopige uitkomsten in tabel 15.1 gemiddeld, dan blijkt dat ze gemiddeld vrijwel gelijk zijn aan de gemiddelde definitieve uitkomsten. Er is dus in algemene zin geen sprake van systematische over- of onderschatting ten opzichte van de definitieve uitkomsten.

15.5 Prognoses voor een lopend boekjaar

Op het moment dat de voorlopige uitkomsten voor een al afgesloten boekjaar berekend zijn, is ook de basis gelegd voor het maken van een prognose voor een lopend boekjaar, waarvan dan inmiddels de helft verstreken is. Vertrekpunt bij de prognoses van de bedrijfsuitkomsten voor een aantal groepen bedrijven, zijn de voorlopige uitkomsten. Deze uitkomsten worden gesplitst in kosten- en opbrengsten-componenten, die elk weer gesplitst worden in prijs- en hoeveelheids-componenten. Op basis van externe gegevens worden aanpassingen doorgevoerd in de prijzen en hoeveelheden van

1) Met behulp van de dichtheidsfunctie van de normale verdeling is af te leiden dat deze verwachting gelijk is aan $(1/\sqrt{2*\pi})$ maal de standaardisatie (Mood, Grayhill and Boes, 1974:177).

produkten en produktiemiddelen en worden de voorlopige uitkomsten bijgesteld om zo te komen tot een prognose. De externe gegevens betreffen oogstramingen, prijs-prognoses, en dergelijke.

16. Conclusies periode 1975-1987

16.1 Inleiding

In de hoofdstukken 4 tot en met 9 is een beschrijving gegeven van de steekproefopzet in de boekjaren vanaf 1975. De berekening van steekproefuitkomsten is daarbij ook aan de orde gekomen.

De hoofdstukken 10 tot en met 15 gaven een evaluatie van de opzet en de daaruit resulterende steekproefuitkomsten. Op grond van de inhoud van deze hoofdstukken wordt in de volgende paragrafen een aantal conclusies getrokken. Bij de bespreking van de nieuwe opzet van de steekproef in deel 3 wordt op een aantal van deze conclusies teruggegrepen.

16.2 Representativiteit

In hoofdstuk 11 is nagegaan in hoeverre er op het niveau van de nationale boerderij verschillen zijn tussen steekproef- en populatiebedrijven. Vergelijking van het gemiddelde en het totaalcijfer in de twee groepen bedrijven voor een aantal variabelen in de metelling, leert dat er op de belangrijkste punten nauwelijks afwijkingen zijn te constateren. De grootte van de bedrijven en de sbe-verdeling over gewassen en dieren in de steekproef sluiten goed aan bij de verhoudingen in de populatie. Enigszins oververtegenwoordigde gewassen zijn peulvruchten, landbouwzaden, pootaardappelen en fabrieksaardappelen. Ondervertegenwoordigd zijn de voedergewassen en de tuinakker- en tuinbouwgewassen. Diersoorten die in de steekproef relatief te veel voorkomen zijn leghennen, moederdieren van slachtpluimvee en kalkoenen. Sterk ondervertegenwoordigd zijn mestkalveren, overig mestvee, schapen, geiten, paarden, eenden, en in iets mindere mate slachtkuikens. De overen ondervertegenwoordiging betreft dus vooral gewassen en diersoorten die in de populatie relatief weinig voorkomen.

Ten aanzien van de overige metellingsgegevens springen vooral de ondervertegenwoordiging van vreemde en vrouwelijke arbeidskrachten en van bedrijven met rechtspersoonlijkheid in het oog. Een iets te sterke vertegenwoordiging is er van de gepachte oppervlakte.

Metellingsgegevens die iets zeggen over de wijze van bedrijfsvoering zijn uiteraard interessant voor een vergelijking. Steekproefbedrijven hebben relatief vaker een ligboxenstal, hebben een grotere huiskavel - maar ook een groter aantal kavels - en beregenen een groter deel van hun oppervlakte cultuurgrond. De afwijkingen zijn echter maar klein.

De conclusie kan getrokken worden dat de steekproef ten aanzien van de meetingsvariabelen een grote mate van representativiteit bezit. Dat is echter nog geen garantie dat de steekproef ook ten aanzien van de doelvariabelen (toegevoegde waarde, arbeidsopbrengst en dergelijke) representatief is. Het is mogelijk dat de steekproef relatief veel bedrijven zou bevatten die qua meetingsvariabelen vergelijkbaar zijn met de groep bedrijven in de meting die ze moeten representeren, maar die gemiddeld hogere of lagere waarden van de doelvariabelen realiseren. Om dit na te gaan is een aantal cijfers in de steekproef vergeleken met cijfers uit externe bronnen. Omdat er nauwelijks extern vergelijkingsmateriaal is voor de doelvariabelen, zijn variabelen die daar enigszins mee samenhangen onderzocht.

Het blijkt dat de melkvetproductie per melkkoe door het boekhoudnet met enkele procenten wordt overschat. Het percentage bedrijven dat deelneemt aan de melkcontrole ligt op de steekproefbedrijven aanzienlijk hoger dan op het totaal van populatiebedrijven.

De kilogramopbrengsten per hectare voor de belangrijkste akkerbouwgewassen lopen in de meeste gevallen niet ver uiteen tussen steekproef en externe bronnen. De vergelijking is in sommige gevallen niet eenduidig te maken vanwege definitieverschillen. Voor de intensieve veehouderij is een vergelijking met externe bronnen niet gemakkelijk uit te voeren.

Een vergelijking van de meetingsvariabelen voor de nationale FIP-boerderij, leidt tot de conclusie dat cijfers over de financiële positie van de ondernemer betrekking hebben op bedrijven die gemiddeld iets kleiner zijn dan de bedrijven in de FIP-populatie. Verder zijn de meeste afwijkingen die voorkomen bij de nationale BUL-boerderij ook hier aanwezig. Een aantal afwijkingen in de FIP-steekproef wordt veroorzaakt door het doelbewust uitsluiten van NV's, BV's en dergelijke; dat zijn meestal grotere bedrijven. Daarnaast zijn er afwijkingen die sterker zijn dan op grond van deze uitsluiting verwacht zouden mogen worden. Zo is het areaal granen en handelsgewassen ondervertegenwoordigd in de steekproef, terwijl het aantal vreemde arbeidskrachten nog sterker onderschat wordt dan in de BUL-steekproef.

De vrij goede representatie van de populatie door de steekproef is voor een groot deel toe te schrijven aan de gevolgde steekproefopzet en de wijze van berekening van de uitkomsten. In hoofdstuk 10 is immers aangegeven dat er grote verschillen in respons zijn tussen verschillende groepen bedrijven in de populatie. Blijkbaar hebben de stratificatie naar bedrijfstype, grootte, oppervlakte, leeftijd en regio en het gebruiken van deze criteria bij de vervangende keuzes effect gesorteerd. Vanuit het oogpunt van representativiteit van de nationale boerderij zijn er weinig redenen om een andere opzet van de steekproef voor te staan. Het systeem kan wel op een aantal punten worden aange-

scherpt. Zo zou een splitsing van de groep "rundveebedrijven" in "zuivere melkveebedrijven" en "overige rundveebedrijven" er aan kunnen bijdragen dat de oververtegenwoordiging van de melkkoeien in de steekproef wordt teruggedrongen en parallel daaraan de ondervertegenwoordiging van mestvee, schapen, geiten, paarden en dergelijke.

Een verdere aanbeveling die gedaan kan worden is de splitsing van de groep "pluimveebedrijven" in "leghennenbedrijven" en "overige pluimveebedrijven". In de akkerbouwsector zou bij de (vervangende) keuze eventueel gekeken kunnen worden naar de gewassoorten op de bedrijven.

Een verdere splitsing van de bestaande acht bedrijfstypen bij de stratificatie zal ten koste gaan van een verdergaande stratificatie naar oppervlakte, leeftijd en regio. In paragraaf 11.5 is echter aangetoond dat deze laatste drie criteria maar een beperkte invloed op de representativiteit hebben.

16.3 Betrouwbaarheid

De centrale doelstelling bij de steekproef is het betrouwbaar schatten van een aantal doelvariabelen in de sfeer van de bedrijfsuitkomsten en van de financiële positie van de ondernemer voor de gemiddelde en totale nationale boerderij. De genormaliseerde toegevoegde waarde (sbe) en de arbeidsopbrengst van de ondernemer zijn daarbij gekozen als richtvariabelen.

Hoofdstuk 12 laat zien dat de steekproefopzet heeft geresulteerd in een standaardfout van 1,5 sbe bij een gemiddeld aantal sbe van ongeveer 210 sbe. De standaardfout van de arbeidsopbrengst van de ondernemer was in de afgelopen jaren ongeveer 1.600 gulden bij een gemiddelde van ruim 25.000 gulden. Door de stratificatie kon een gemiddelde reductie in de standaardfout van het aantal sbe worden gerealiseerd van 50 tot 60%. Voor de arbeidsopbrengst was die reductie ongeveer 10%.

Veel grootheden in de opbrengsten-, kosten- en inkomenssfeer hebben door de stratificatie een standaardfout die 40 tot 50% lager ligt dan bij een enkelvoudige steekproef het geval zou zijn geweest. Daarbij speelt vooral de stratificatie-variabele "sbe" een grote rol. Opvallend is dat het hanteren van de stratificatie-variabelen oppervlakte, leeftijd en regio nauwelijks van invloed is op de betrouwbaarheid van de meeste nationale gemiddelden. Een vergaande stratificatie is daarom voor de betrouwbaarheids-doelstelling minder relevant.

Door het grote aantal van ruim 1.000 steekproefbedrijven zijn de relatieve standaardfouten van veel schattingen erg klein; meestal in de orde van grootte van 1 tot 2%. Voor schattingen van grootheden in de financiële sfeer zijn de relatieve standaardfouten in het algemeen iets groter. De grotere heterogeniteit in de populatie, het kleinere aantal steekproefbedrijven en de geringere reductie door de stratificatie spelen daarbij een rol. De resultaten blijven echter alleszins acceptabel.

Ook voor quotiënten levert de steekproef schattingen met kleine standaardfouten. De gemiddelde opbrengst per f 100,- kosten bijvoorbeeld heeft een standaardfout die kleiner is dan 40 cent.

Bij de stratificatie van de populatie en de allocatie van steekproefbedrijven over de strata, wordt gebruik gemaakt van de variabele $(sbe)^k$. Voor het betrouwbaar schatten van de doelvariabelen blijkt dit een zeer geschikte hulpvariabele te zijn. De geringere toewijzing van steekproefbedrijven aan strata die betrekking hebben op de rundveehouderij dan op grond van de standaarddeviatie in $(sbe)^k$ "optimaal" zou zijn, blijkt een juiste strategie te zijn in verband met het betrouwbaar schatten van de gemiddelden van de doelvariabelen. Handhaving van het gebruik van $(sbe)^k$ bij de stratificatie en allocatie ligt voor de hand. Kleine verbeteringen zijn wel mogelijk. Aan de groep gemengde rundveebedrijven met intensieve veehouderij zouden in navolging van de groep zuivere rundveebedrijven, circa 20% minder bedrijven toegewezen kunnen worden. De groep pluimveebedrijven zou een hoger steekproefpercentage kunnen krijgen dan op grond van de standaarddeviatie in $(sbe)^k$ "optimaal" is. Dit zou eventueel gecombineerd kunnen worden met een splitsing van deze groep bedrijven in homogeneren deelgroepen. Ook voor de akkerbouwbedrijven ligt een iets hoger steekproefpercentage dan het huidige voor de hand.

De tweede doelstelling van de steekproef is het betrouwbaar schatten van gemiddelden voor BUL- en FIP-groepen. Hoofdstuk 13 laat zien dat de standaardfouten op dit niveau aanzienlijk groter kunnen zijn dan op nationaal niveau. De kleinere aantallen steekproefbedrijven per groep zijn daarvan de hoofdoorzaak. De samenstelling van BUL- en FIP-groepen op grond van andere criteria dan die bij de stratavorming gebruikt worden, speelt daarbij ook een rol. Doordat de groepen dwars door de strata heenlopen, wordt een extra "fout" geïntroduceerd, namelijk die in het aantal bedrijven. De som van de wegingsfactoren is immers een schatting van het aantal bedrijven dat daadwerkelijk tot de betreffende groep behoort. In sommige gevallen leidt dit ertoe dat geschatte gemiddelden een standaardfout hebben die enkele tientallen procenten hoger ligt dan het geval zou zijn bij vaste wegingsfactoren. Voor totaalschattingen wordt de standaardfout in veel gevallen zelfs met een factor 2 tot 3 vergroot.

Voor een deel zijn de schattingsfouten in de wegingsfactoren niet te voorkomen. De BUL- en FIP-groepen worden namelijk samengesteld op basis van kenmerken van bedrijven zoals die tijdens het boekjaar worden waargenomen. Er is echter wel een verkleining van de standaardfouten te realiseren door de stratificatie meer te laten aansluiten bij de groepsvorming voor BUL en FIP. Zowel de indeling in bedrijfstypen als die in grootte-klassen en regio's verschilt tussen stratificatie en publikatie. Er zijn weinig redenen om deze twee niet te laten aansluiten, temeer omdat

een enigszins gewijzigde indeling bij de stratificatie nauwelijks gevolgen zal hebben voor de representativiteit en betrouwbaarheid van uitkomsten op nationaal niveau.

De standaardfouten van verschillen tussen BUL-groepen of tussen FIP-groepen zijn in het algemeen aanzienlijk, waardoor op het oog aanwezige verschillen vaak niet statistisch significant kunnen worden vastgesteld.

Omdat ieder jaar slechts een deel van de steekproefbedrijven vervangen wordt door nieuwe bedrijven, zijn de standaardfouten van verschillen tussen twee opeenvolgende boekjaren voor dezelfde BUL-groep, enkele tientallen procenten kleiner dan op grond van de afzonderlijke standaardfouten verwacht zou mogen worden. Om dezelfde reden is de standaardfout van het verschil tussen voorlopige en definitieve uitkomsten relatief klein.

Deel 3: de steekproef vanaf boekjaar 1988

"Wat wil je weten over wie en met welke betrouwbaarheid?"

(K. Lodder)

17. Overgang op de NEG-typering

17.1 Inleiding

Met ingang van 1986 is in Nederland de VAT-typering vervangen door de NEG-typering. Het LEI heeft besloten om met ingang van boekjaar 1988 deze typering te gaan gebruiken bij de presentatie van steekproefuitkomsten in de BUL en in de FIP. De afbakening van de populatie waar de steekproefgegevens betrekking op hebben zal ook in termen van de NEG-typering gaan plaatsvinden. In dit hoofdstuk zal op deze veranderingen worden ingegaan.

De gevolgen die de nieuwe vorm van groepsindeling in LEI-publicaties heeft voor de steekproefopzet worden in de paragrafen 5 en 6 van dit hoofdstuk kort aangeduid. In de hoofdstukken 18 tot en met 20 wordt - aansluitend bij de conclusies van deel 2 - een enigszins gewijzigde steekproefopzet ontwikkeld.

17.2 De NEG-typering

In de Europese Gemeenschap is in de zeventiger jaren een EG-typering voor land- en tuinbouwbedrijven ontwikkeld. Deze typering, die is gebaseerd op het bruto standaard saldo (bss) van dieren en gewassen, wordt gebruikt om de bedrijven in het EG-boekhoudnet te kunnen groeperen. Zowel bij de stratificatie van de EG-steekproef als bij de presentatie van gegevens wordt deze typering gehanteerd. Voor een uitvoerige beschrijving van de wijze waarop de EG bedrijven typeert en hoe deze typering een rol speelt in het EG-boekhoudnet, zij verwezen naar Poppe (1984, 1986).

Binnen Nederland werden de laatste decennia verschillende bedrijfstyperingssystemen gebruikt. In Landbouwcijfers en in andere publicaties over de Landbouwtelling werd de VAT-typering gebruikt. Het LEI publiceerde resultaten over groepen landbouwbedrijven aan de hand van de BUL-typering (zie tabel 3.2). De EG-typering werd weinig gebruikt.

In 1986 is in Nederland de NEG-typering in gebruik genomen (Landbouwcijfers, 1987:III). Deze typering sluit nauw aan bij de EG-typering en bevordert daardoor de internationale vergelijkbaarheid van statistische gegevens. De NEG-typering zal niet alleen de VAT-typering gaan vervangen, maar zal voor het LEI ook de typering zijn die gebruikt gaat worden voor het presenteren van bedrijfsuitkomsten. De oude BUL-typering zal daarmee verdwijnen. De sbe als grootte-maatstaf zal gehandhaafd blijven voor onderzoeks- en voorlichtingsdoeleinden, maar zal niet meer gebruikt worden bij typering en dimensionering van bedrijven. Het bruto standaard saldo (bss) zal deze rol over gaan nemen.

Een beschrijving van het NEG-typeringssysteem en van de motivering voor de ontwikkeling ervan is te vinden in CBS (1986), CBS (1988) en voor den Dag (1988).

17.3 BUL-groepen op NEG-basis

In hoofdlijnen zijn de oude en de nieuwe BUL-typering niet essentieel verschillend. Er blijven groepen akkerbouwbedrijven, weide- of melkveebedrijven, gemengde bedrijven en intensieve veehouderijbedrijven. De samenstelling van die groepen is soms wat ruimer en soms wat krapper dan in de oude typering. De groepen intensieve veehouderijbedrijven komen wat duidelijker in beeld. De gemengde groepen worden tot een vrij grote groep samengevoegd. In hoofdstuk 21 wordt de aansluiting van de ene typering op de andere nagegaan. In figuur 17.1 worden de nieuwe BUL-groepen op een rij gezet. Deze groepsindeling zal ook voor de FIP gebruikt gaan worden.

17.4 Afbakening steekproefpopulatie

Evenals in het verleden vindt er bij de bepaling van de populatie een uitsluiting van kleine en extreem grote bedrijven plaats. De grootte van bedrijven wordt uitgedrukt in Nederlandse grootte-eenheden (nge), die zijn afgeleid van de bss. Bedrijven beneden een bepaalde nge-ondergrens en boven een bepaalde nge-bovengrens, worden om soortgelijke redenen als in het verleden (zie paragraaf 4.2) uitgesloten van de populatie. De afbakening ten opzichte van de tuinbouwsector wordt niet meer direct uitgevoerd op basis van het procentuele aandeel van land- en tuinbouw op de bedrijven, maar volgt uit de NEG-typologie. Het boekhoudnet landbouw heeft in de toekomst betrekking op de volgende NEG-hoofdtypen:

- A) Akkerbouwbedrijven
- B) Graasdierbedrijven
- C) Hokdierbedrijven
- D) Gewassencombinaties (exclusief tuinbouw/blijvende teeltbedrijven)
- E) Veeteeltcombinaties
- F) Gewassen/veeteeltcombinaties

Het boekhoudnet van tuinbouwbedrijven zal zich richten op de overige bedrijfstypen: tuinbouw- en blijvende teelt-bedrijven en een deel van de gewassencombinaties. Daarmee is afbakening van de boekhoudnetten van land- en tuinbouw duidelijker geworden dan in het verleden.

1) Akkerbouwbedrijven (NEG-type 1):

Grotere bedrijven:

- Noordelijk kleigebied (regio 1)
- Centraal kleigebied (regio 3-4)
- Zuidwest. kleigebied (regio 5)
- Veenkolonien (regio 6-8)

Kleinere bedrijven:

- Kleigebieden (regio 1,3-5)
- Veenkolonien (regio 6-8)

2) Sterk gespecialiseerde melkveebedrijven (NEG-type 411):

Grotere bedrijven:

- Noordelijk klei/Veenw.gebied (regio 1-3)
- Westelijk weide gebied (regio 4,5,11)
- Zand gebieden (regio 6-10,12,13)

Kleinere bedrijven:

- Klei/veenweidegebieden (regio 1-5,11)
- Zand gebieden (regio 6-10,12,13)

3) Minder sterk gespecialiseerde melkveebedrijven
(NEG-type 412-437):

Grotere bedrijven
Kleinere bedrijven

4) Hokdierenbedrijven (NEG-type 5):

Fokvarkensbedrijven (NEG-type 5011):
Andere varkensbedrijven (NEG-type 5012-5013):
Legkippenbedrijven (NEG-type 5021):

5) Combinatie-bedrijven en overige bedrijven
(NEG-typen 438-4449,6-8):

Grotere bedrijven
Kleinere bedrijven

Figuur 17.1 Indeling van BUL-groepen vanaf boekjaar 1988

Bron: Poppe en Dijk (1986).

17.5 Naar een uniforme typering

In de periode 1975-1987 werden door het LEI verschillende typeringen gebruikt voor verschillende doelen. Daar is op zich-

zelf niets vreemds aan, omdat de ene typering voor een bepaald doel tot een groter inzicht leidt dan een andere typering. Zo was er een typering voor de publikatie van bedrijfsuitkomsten - de BUL-typering -, een typering voor de stratificatie van de steekproefpopulatie, een EG-typering, en daarnaast de meer algemeen gebruikte VAT-typering. Hoewel deze typeringen in hoofdlijnen wel iets op elkaar leken, was het in de praktijk toch moeilijk om aan te geven hoe ze op elkaar aansloten. In deel 2 is aangegeven dat er goede redenen zijn om de typering bij de stratificatie af te stemmen op de typering bij de publikatie. Nu de typering voor de BUL gebaseerd gaat worden op de NEG-typering en deze bovendien sterk overeenkomt met de EG-typering, die gebruikt wordt voor publikatie van resultaten van het Europese boekhoudnet, is er alle reden om bij de stratificatie ook gebruik te maken van de NEG-typering.

Door bij de stratificatie en bij publikatie de NEG-typering te hanteren, wordt de vergelijkbaarheid met cijfers uit andere bronnen groter. Immers, als gepubliceerd wordt over "akkerbouwbedrijven" of over "legkippenbedrijven", dan gaat het in LEI-publicaties om dezelfde groepen als bij bijvoorbeeld het CBS of de EG. Ook wordt duidelijker op welke groep bedrijven de totale LEI-steekproef zich richt, omdat de populatie-afbakening in algemeen bekende termen plaats vindt.

In het volgende hoofdstuk zal nader worden ingaan op deze afstemming van stratificatie- op publikatie-indelingen. Een probleem dat zich voordoet bij de indeling van de BUL-bedrijven, verdient echter eerst nog enige aandacht. Het betreft de vraag op welke gegevensbasis de indeling moet plaatsvinden. De volgende paragraaf gaat daar op in.

17.6 BUL-typering: homogeniteit of herkenbaarheid?

Tot nu toe is er in dit hoofdstuk van uitgegaan dat de indeling in BUL-groepen gebeurt op basis van de bedrijfskenmerken in de meitelling. Immers, alleen dan is er aansluiting van de stratificatie op de BUL-indeling, en is er de mogelijkheid om BUL-uitkomsten te vergelijken met gegevens uit andere statistieken. Anders gezegd, alleen dan is de som van de wegingsfactoren voor een bepaalde BUL-groep gelijk aan de aantallen in de meitelling in die BUL-groep.

De huidige werkwijze in de BUL is anders. Bedrijven worden ingedeeld op basis van de bedrijfskenmerken gemiddeld gedurende het boekjaar. Zowel het type als de grootte van het bedrijf worden daarmee bepaald 1). Daardoor kunnen er verschillen ontstaan

1) De berekening van wegingsfactoren en de daaraan ten grondslag liggende stratificatie hebben uiteraard betrekking op meitellingsgegevens van bedrijven.

tussen de som van de wegingsfactoren in de betreffende BUL-groep en het aantal bedrijven in de metelling dat op grond van metellingsgegevens tot die BUL-groep behoort.

Uit het oogpunt van het vormen van homogene groepen bedrijven verdient het de voorkeur om te typeren op basis van gemiddelden gedurende het boekjaar. Een bedrijf dat per 1 mei getypeerd wordt als klein akkerbouwbedrijf kan in werkelijkheid een groot pluimveebedrijf zijn. Opname van dit bedrijf in de BUL-groep "klein akkerbouw" zou de cijfers van "klein akkerbouw" iets moeilijker interpreteerbaar maken.

Anderzijds is de herkenbaarheid en aanwijsbaarheid van een groep bedrijven belangrijk. Dat pleit voor typering volgens de metelling. Als de BUL-metellingsgroep "klein akkerbouw" in de metelling bijvoorbeeld 2.000 bedrijven bevat, dan kan uit het boekhoudnet gehaald worden dat daarvan bijvoorbeeld 1.800 bedrijven ook gemiddeld gedurende het boekjaar die typering hebben gehad, terwijl de overige 200 bedrijven een andere typering hadden. De BUL-uitkomsten voor de totale metellingsgroep van 2.000 bedrijven hebben dan betrekking op een enigszins heterogene groep, maar het is in elk geval wel duidelijk welke bedrijven dat zijn, namelijk de bedrijven die door hun opgave in de metelling aangeven "klein akkerbouwbedrijf" te zijn.

Wie geïnteresseerd is in een koppeling van structuurgegevens (metelling) en boekhoudnetgegevens zal kiezen voor de laatste benadering. De EG vormt bijvoorbeeld bedrijfsgroepen op basis van bedrijfskenmerken in de metelling. Wie wil weten hoe een gemiddeld akkerbouwbedrijf er uitziet en minder geïnteresseerd is in het aantal bedrijven waarop dat betrekking heeft, zal kiezen voor de tweede benadering.

In de afweging tussen homogeniteit en herkenbaarheid moet overigens ook worden meegenomen dat de typering die aansluit bij de stratificatie - de typering op basis van herkenbaarheid - in het algemeen lagere standaardfouten zal opleveren dan de andere typering, ondanks de grotere variantie (heterogeniteit) in de eerstgenoemde typering (zie paragraaf 13.4).

Voor de publikatie van steekproefuitkomsten in de BUL en de FIP zal het LEI ook na 1987 uitgaan van indelingen op basis van bedrijfskenmerken tijdens het boekjaar. Om deze groeps cijfers te kunnen vertalen naar groepen bedrijven in de metelling zal in het vervolg wel aangegeven worden uit welke metellingsgroepen de bedrijven in een bepaalde BUL-groep afkomstig zijn. Dat kan bijvoorbeeld op de manier die is aangegeven in bijlage 6.

18. Van doelstellingen naar steekproef

18.1 Inleiding

Aan het begin van de zeventiger jaren is de landbouwsteekproef opgezet zoals die tot en met boekjaar 1987 heeft gefunctioneerd. Vanaf die tijd zijn de structurele ontwikkelingen in de landbouw in hoog tempo doorgegaan. Dat kan een reden zijn om na te gaan of deze opzet nog goed aansluit bij de informatie-behoefte over landbouwbedrijven, zeker nu een nieuwe bedrijfstypering en populatie-afbakening gebruikt zal gaan worden. Bovendien is er inmiddels meer dan tien jaar ervaring met de steekproefopzet. Op grond daarvan is in deel 2 bekeken of er verbeteringen zijn te realiseren, in die zin dat met dezelfde middelen een betere kwaliteit van de informatie is te bewerkstelligen.

In dit en het volgende hoofdstuk worden de doelstellingen van de steekproef opnieuw geformuleerd - wat niet betekent dat de doelstellingen op zichzelf veel veranderen - en wordt aangegeven op welke wijze die doelstellingen gerealiseerd kunnen worden.

18.2 Uitwerking hoofd- en nevendoelelstellingen 1975-1987

In hoofdstuk 4 van dit onderzoekverslag is aangegeven dat de hoofddoelelstelling van het boekhoudnet bestond uit het weergeven van een zo representatief en betrouwbaar mogelijk beeld van de Nederlandse landbouw. In de praktische uitwerking betekende dit, dat gestreefd werd naar maximale betrouwbaarheid van schattingen voor een aantal bedrijfseconomische en financiële grootheden voor de gemiddelde en totale nationale boerderij, gegeven een aantal maatregelen die de representativiteit moesten waarborgen.

De uitwerking van deze hoofddoelelstelling zorgde ervoor dat aan nevendoelelstellingen niet zonder meer voldaan kon worden. Die nevendoelelstellingen bestaan uit het publiceren van gegevens over afzonderlijke groepen bedrijven binnen de populatie in de BUL en in de FIP. Dat laatste is pas verantwoord indien er genoeg steekproefbedrijven voor een afzonderlijke groep beschikbaar zijn. De toewijzing (allocatie) van bedrijven over de verschillende delen van de populatie die voortvloeide uit de hoofddoelelstelling, bleek in de praktijk niet zodanig te zijn dat automatisch voldoende bedrijven beschikbaar waren voor de afzonderlijke BUL- en FIP-groepen.

Daarom werden bij deze allocatie ongeveer 60 bedrijven buiten de verdeling gelaten en achteraf toegevoegd aan de verschillende delen van de populatie waar dat uit hoofde van de nevendoelelstellingen noodzakelijk was. Er werd naar gestreefd om zo min mogelijk bedrijven achteraf toe te wijzen om zo min mogelijk afbreuk te doen aan de hoofddoelelstelling.

In de praktijk bleek het niet altijd eenvoudig te zijn om te bepalen in welke strata de 60 BUL-aanvullingsbedrijven gekozen moesten worden. De keuze werd bepaald door de definitie van de BUL-groep waar de bedrijven uiteindelijk in moesten terechtkomen. Maar deze definitie was niet altijd even gemakkelijk te vertalen in de definities van de strata. In de eerste plaats verschilden de bedrijfstype-definities en sloten de sbe-klassen 1 en 2 bij de stratificatie niet direct aan op de indeling in groot en klein in de BUL. In de tweede plaats bestond een stratum veelal uit bedrijven uit een aantal regio's tegelijk, terwijl voor de BUL-aanvulling bedrijven uit specifieke regio's nodig waren. Bij handhaving van het principe van aselechte trekking binnen een stratum, was de kans aanwezig dat bedrijven gekozen worden die uit het oogpunt van BUL-aanvulling niet nodig waren. De aanvullende keuze moest dan noodzakelijkerwijs beperkt worden tot strata waarin alleen bedrijven uit de betreffende regio voorkwamen. Indien het stratificatie-proces in andere jaren tot andere strata leidde, kon dit tot vertekeningen leiden vanwege uiteenlopende historische trekkingskansen binnen een stratum.

Een betere afstemming van de vorming van enerzijds strata en anderzijds BUL-groepen is daarom zowel vanuit theoretische als praktische overwegingen aantrekkelijk. Er zijn echter nog zwaarwegender redenen om de typering die gebruikt wordt bij de stratificatie aan te laten sluiten bij de typering die bij publikatie van gegevens wordt gehanteerd. Daarop zal in de volgende paragraaf worden ingegaan.

18.3 Afstemming stratificatie op groepsindeling bij publikatie

Bij de stratificatie van de populatie werden de bedrijven in de populatie eerst ingedeeld in acht bedrijfstypen en vervolgens in vier sbe-klassen, resulterend in 32 basisstrata. Het principe achter de indeling in bedrijfstypen was, dat er binnen de typen homogeniteit en tussen de typen heterogeniteit moest ontstaan. Vanuit het homogeniteits-criterium verdient de ene indeling in bedrijfstypen de voorkeur boven de andere indeling. Een indeling die de bedrijven alleen indeelt op basis van het aantal paarden dat er voorkomt, is duidelijk niet de meest geschikte indeling voor de steekproef. Een indeling op basis van de verdeling van het totale aantal sbe over de verschillende dieren en gewassen ligt meer voor de hand. In het verleden werden wel typeringen gemaakt op basis van de oppervlakte-verdeling over de gewassen, of op basis van bewerkingseenheden. De VAT-typering die van 1968 tot 1986 onder andere door het CBS werd gehanteerd, was gebaseerd op de sbe. Ook de stratificatie-typering en de BUL-typering waren erop gebaseerd. De sbe blijkt een goede maat om aan te geven waar de economische accenten op een bedrijf liggen. Hoe deze typering vervolgens een meer precieze vorm wordt gegeven, hangt af van het doel van de typering.

Het doel bij de steekproef is het realiseren van representatieve en betrouwbare uitkomsten. In paragraaf 12.10 bleek dat er op het punt van de betrouwbaarheid van de uitkomsten van de gemiddelde nationale boerderij, weinig verschillen zijn te constateren tussen een aantal mogelijke typeringen. Of bij de stratificatie nu gebruik gemaakt wordt van de VAT-typering, de BUL-typering, de huidige stratificatie-typering of een andere er op gelijkende typering, maakt weinig verschil voor de standaardfouten. De representativiteit is op sommige punten wel gevoelig voor de indeling in bedrijfstypen (paragraaf 11.2). Zo werd tot 1987 een bedrijfstype pluimveebedrijven onderscheiden, met daarbinnen zowel legkippen-, slachtkuiken-, als kalkoenenbedrijven. Door een lage effectieve respons van slachtkuikenbedrijven en een hoge respons van kalkoenenbedrijven en bedrijven met slachtkuikenmoederdieren, is de eerste groep ondervertegenwoordigd en de tweede oververtegenwoordigd in het boekhoudnet. Op een soortgelijke manier zijn de bedrijven met mestvee en schapen ondervertegenwoordigd bij de rundveebedrijven.

De bedrijfstypering bij de stratificatie moet dus bedrijven die zowel qua bedrijfskenmerken als qua respons uiteenlopen, zoveel mogelijk van elkaar scheiden.

Zoals gezegd zijn er vanuit het gezichtspunt van de gemiddelde nationale boerderij verschillende typeringen mogelijk die elkaar weinig ontlopen. Indien echter gelet wordt op de neven-doelstelling van het boekhoudnet, dan is de typering van veel groter belang. In hoofdstuk 13 kwam naar voren dat indien de indeling bij stratificatie en publikatie uiteenlopen, dit kan leiden tot aanmerkelijke vergrotingen van de standaardfouten. Het is daarom gewenst om bij de indeling in strata een bedrijfstypering te kiezen die aansluit bij de BUL-bedrijfstypering. Een bijkomend voordeel is dan dat de regionale aanvullingen veel gericht plaats kunnen vinden (zie paragraaf 18.2).

Ook uit het oogpunt van de hoofddoelstelling is er reden om bij de stratificatie rekening te houden met indelingen die bij de publikatie een rol spelen. De EG publiceert niet alleen cijfers over de gehele Nederlandse landbouw, maar ook over bepaalde EG-bedrijfstypen en EGE-groottesklassen.

18.4 Explicitering van doelstellingen

Een verandering in de typering bij de stratificatie brengt niet direct de noodzaak met zich mee om ook verder iets te veranderen in de steekproefopzet. Toch is deze verandering aangegrepen om op een aantal punten de opzet wat bij te sturen, met name om de afweging van hoofd- en neven-doelstellingen en van de neven-doelstellingen onderling, explicieter te maken.

De "oude" steekproefopzet ging ervan uit dat de hoofd- en neven-doelstellingen vertaald konden worden in een eerste alloca-

tie van 1.100 bedrijven op een vanuit de hoofddoelstelling optimale wijze en een verdere allocatie van 60 bedrijven om aan de nevendoelestellingen te voldoen. Impliciet zit achter deze allocatie een afweging tussen het belang van de hoofddoelstelling en dat van de nevendoelestellingen. Voor een goede afweging is het achter gewenst om de allocatie meer expliciet te laten aansluiten bij de geformuleerde doelstellingen.

Uitgangspunt blijft dat de hoofddoelstelling van het boekhoudnet ligt in het representatief en betrouwbaar weergeven van de gemiddelde nationale boerderij. De vraag is dan in hoeverre afbreuk gedaan wordt aan deze doelstelling indien een aantal bedrijven los van deze doelstelling wordt gealloceerd aan bepaalde strata die corresponderen met bepaalde BUL-groepen. In paragraaf 7.2 bleek dat een vermindering van het aantal steekproefbedrijven dat in eerste instantie optimaal gealloceerd wordt, en dientengevolge een vergroting van het aantal aanvullingsbedrijven, tot erg kleine vergrotingen van de standaardfouten van de schattingen voor de nationale boerderij leidt.

Voor de BUL zijn een aantal nieuwe BUL-groepen gevormd (zie paragraaf 17.3) en voor deze nieuwe groepen is bepaald hoeveel steekproefbedrijven nodig zijn om te kunnen komen tot betrouwbare resultaten. Deze aantallen steekproefbedrijven zijn berekend op grond van informatie uit voorgaande boekjaren. Voor vier grootheden is vastgesteld welke norm gesteld moet worden aan de standaardfout van de schattingen voor de verschillende groepen. Op grond van deze normen en informatie uit voorgaande boekjaren over de variantie in de betreffende grootheden in de populatie, kon berekend worden hoeveel bedrijven voor elke groep nodig zijn. Deze aantallen zijn gelegd naast de aantallen die voortvloeien uit de optimale allocatie en vervolgens is een tussen-allocatie gezocht, waarbij de optimale allocatie in grote lijnen gevolgd wordt, maar waarbij op bepaalde punten wordt tegemoet gekomen aan de eisen die voor de BUL gesteld worden (zie hoofdstuk 19). Een en ander heeft overigens niet tot drastische wijzigingen in de allocatie geleid. Er vindt alleen een directere afweging tussen doelstellingen plaats.

19. Concretisering van de doelstellingen

19.1 Inleiding

Hoofdstuk 17 gaf aan op welke populatie en op welke NEG-groepen daarbinnen de hoofd- en nevendoelstellingen betrekking hebben. In hoofdstuk 18 is ingegaan op de noodzaak tot het expliciet maken van de doelstellingen van het boekhoudnet en de wenselijkheid van het laten aansluiten van de stratificatie op indelingen die bij publikatie van gegevens gehanteerd worden. In dit hoofdstuk zullen de doelstellingen voor de populatie (de nationale boerderij) en voor de BUL-groepen nader geconcretiseerd worden.

19.2 De nationale boerderij

Op basis van gegevens over een aantal boekjaren uit het recente verleden is nagegaan welke eisen impliciet aan de betrouwbaarheid van de uitkomsten van de gemiddelde nationale boerderij gesteld worden. Voor vier grootheden is nagegaan welke standaardfout wordt gerealiseerd bij een aantal steekproefbedrijven van ruim 1.000 voor de BUL en ongeveer 850 voor de FIP. In tabel 19.1 staan de resultaten.

Tabel 19.1 *Gemiddelden en standaardfouten voor vier doelvariabelen voor de gemiddelde nationale boerderij (gemiddeld over de boekjaren 1982-1986)*

Variabele	Schatting van het gemiddelde	Standaardfout van het geschatte gemiddelde	
		absoluut	in procenten
BUL:		*)	
opbr. per f 100,- kosten	90	0,53	0,6
faktoropbrengst	95.226	1.295	1,4
FIP:			
gezinsinkomen	80.792	1.574	1,9
besparingen	14.331	1.397	9,7

*) Berekend met vergelijking (6.4.5).

Deze standaardfouten zijn gerealiseerd op basis van een optimale stratificatie en een optimale allocatie. Indien de standaardfouten worden berekend op basis van enigszins afwijkende strata, waarbij de gekozen steekproefbedrijven achteraf in deze strata worden ingedeeld en de allocatie daardoor ook minder optimaal kan zijn, worden de standaardfouten in het algemeen weinig vergroot. Indeling van de strata op basis van de NEG-typering met handhaving van de al gevormde sbe-klassen, levert standaardfouten op die vergelijkbaar zijn met die van de huidige stratificatie. Overgang op een enigszins andere stratificatie om meer aansluiting te zoeken bij de indeling van de BUL-groepen en de indelingen op EG-niveau, hoeft dus vanuit de betrouwbaarheid van de uitkomsten van de gemiddelde nationale boerderij geen belemmering te vormen.

Verder is de steekproefomvang op nationaal niveau van een dermate grote omvang dat enige tegemoetkoming aan BUL-doelstellingen - dat wil zeggen een minder dan optimale allocatie - nauwelijks invloed heeft op de betrouwbaarheid op nationaal niveau. De 60 bedrijven die in het verleden gekozen werden als regionale aanvullingen, hebben de realisering van de hoofddoelstelling van het boekhoudnet dan ook nauwelijks geschaad. Ook voor de toekomst zijn er daarom mogelijkheden om te voldoen aan eisen die gesteld worden om voor afzonderlijke groepen binnen de Nederlandse landbouw, gegevens te kunnen publiceren.

Ervan uitgaande dat de betrouwbaarheid die in het verleden is gerealiseerd voor de nationale boerderij, norm is voor wat "redelijk betrouwbaar" is, kan geconcludeerd worden dat een nieuwe opzet (NEG-stratificatie) leidt tot redelijk betrouwbare uitkomsten.

19.3 BUL-groepen

Omdat de afzonderlijke groepen in de BUL en de FIP enigszins anders samengesteld gaan worden, zal daarvoor opnieuw geformuleerd moeten worden wat "redelijk betrouwbaar" is. Daarbij kan voor groepen die vrijwel overeenkomen met "oude" BUL-groepen, de betrouwbaarheid in het verleden norm zijn voor de betrouwbaarheid in de toekomst. Op basis van inpassing van de steekproefbedrijven in een nieuwe NEG-stratificatie, is berekend welke standaardfouten voor elk van de groepen horen bij de huidige aantallen bedrijven per nieuwe BUL-groep. Deze cijfers zijn gezet naast de gewenste standaardfouten voor de verschillende BUL-groepen. Op grond daarvan is berekend welke aantallen per BUL-groep nodig zijn om te voldoen aan de eisen.

De betreffende gewenste standaardfouten en de bijbehorende steekproefaantallen zijn te vinden in tabel 19.2 en tabel 19.3. Bij het vaststellen van de gewenste standaardfouten is zoveel mogelijk aangesloten bij standaardfouten die haalbaar zijn vanuit

de optimale allocatie op nationaal niveau. Op sommige punten is daarvan afgeweken om tegemoet te komen aan BUL-wensen. Evenals in het verleden heeft dat betrekking op grotere akkerbouwbedrijven. Voor de groepen hokdierenbedrijven werden in het verleden geen afzonderlijke cijfers gepubliceerd. De wensen met betrekking tot de standaardfouten zijn daar zodanig gekozen dat daarvoor iets meer bedrijven nodig zijn dan uit de optimale allocatie voortvloeit. Daardoor kunnen erg hoge standaardfouten voor deze groepen worden voorkomen.

De gewenste standaardfouten in tabel 19.2 kunnen gerealiseerd worden indien per BUL-groep bepaalde minimale aantallen bedrijven in de steekproef vallen. Tabel 19.3 geeft deze aantallen.

Tabel 19.3 is in 1986 opgesteld op basis van informatie over het boekjaar 1984 en voorgaande jaren. Omdat de standaardfouten in een steekproef schattingen zijn, zijn de gewenste aantallen in de tabel ook schattingen. Voor de boekjaren vòòr 1984 was bovendien niet de NEG-typering van de boekhoudnetbedrijven beschikbaar, zodat voor die jaren gekeken is naar de standaardfouten van vergelijkbare BUL-groepen. Verder moet bedacht worden dat de typering van specifieke bedrijven ook voortdurend aan verandering onderhevig is, door veranderingen in de normen (bss) die bij de typering een rol spelen. De gewenste aantallen die horen bij de gewenste standaardfouten zijn voor een aantal groepen dan ook niet meer dan globale indicaties.

De aantallen in tabel 19.3 hebben expliciet een rol gespeeld bij de bedrijfskeuze voor boekjaar 1988 (zie hoofdstuk 20).

In 1988 kwam informatie beschikbaar voor de boekjaren 1984-1986. Voor bepaalde groepen bleken de gestelde aantallen-eisen naar beneden bijgesteld te kunnen worden - of indien dat vanuit nationaal oogpunt geen problemen opleverde - de gestelde standaardfouten-eisen strakker aangetrokken te kunnen worden. Voor andere groepen bleken meer bedrijven nodig te zijn om aan de gestelde eisen te kunnen voldoen. In sommige gevallen is daarvoor ruimte aanwezig. In andere gevallen moet genoeg worden genomen met grotere standaardfouten.

Tabel 19.2 *Explicitering van de BUL-doelstellingen; gewenste standaardfouten naar BUL- en FIP-groep en variabele*

BUL/FIP-groep	Variabele				
	opbrengst/ f100,-	factor- kosten	factor- opbr.	gezins- inkomen	bespa- ringen
Akkerbouwbedrijven	1,25	5.500	5.500	5.500	4.500
Grotere bedrijven	1,00	7.500	7.500	7.500	6.000
- noordelijk klei	2,50	12.500	12.500	12.500	10.000
- veenkolonien	2,00	6.000	6.000	6.000	6.000
- centraal klei	2,00	6.000	8.000	8.000	10.000
- zuidwestelijk klei	2,00	7.500	8.000	8.000	7.500
Kleinere bedrijven	3,00	4.000	4.500	4.500	5.000
- kleigebieden	3,50	6.000	6.000	6.000	6.000
- veenkolonien	2,50	4.000	5.000	5.000	4.000
St. gesp. melkveebedrijven	0,75	2.000	2.000	2.000	1.750
Grotere bedrijven	0,75	2.500	2.500	2.500	2.250
- nrd. klei/veenweide	1,25	5.000	4.000	4.000	4.000
- westelijk weide	1,50	6.000	5.000	5.000	4.000
- zandgebieden	1,00	3.500	3.500	3.500	3.500
Kleinere bedrijven	1,25	2.250	2.500	2.500	2.250
- klei/veenweide	1,50	2.500	3.000	3.000	3.000
- zandgebieden	1,50	2.500	3.000	3.000	2.250
Minder gespec. melkveebedr.	1,25	3.500	3.250	3.250	3.000
Grotere bedrijven	1,25	6.000	6.000	6.000	6.000
Kleinere bedrijven	2,00	3.000	3.000	3.000	3.000
Hokdieren-bedrijven	0,75	6.000	5.000	5.000	4.500
Fokvarkensbedrijven	1,25	7.500	7.500	7.500	7.000
Andere varkensbedrijven	1,25	7.500	7.500	7.500	6.000
Legkippenbedrijven	1,25	10.000	10.000	10.000	10.000
Overige bedrijven	1,50	5.000	6.000	6.000	6.000
Grotere bedrijven	1,75	7.500	7.500	7.500	8.000
Kleinere bedrijven	2,50	6.000	7.500	7.500	8.000

Tabel 19.3 Gewenste aantallen steekproefbedrijven naar BUL- en FIP-groep en variabele

BUL/FIP-groep	Variabele			
	opbrengst/ f100,- kosten	factor- opbr.	gezins- inkomen	bespa- ringen
Akkerbouwbedrijven	235	245	175	165
Grotere bedrijven	210	205	130	130
- noordelijk klei	55	60	35	35
- veenkolonien	40	35	30	30
- centraal klei	55	55	45	50
- zuidwestelijk klei	55	55	40	35
Kleinere bedrijven	55	55	45	45
- kleigebieden	30	30	25	30
- veenkolonien	25	25	25	25
St. gesp. melkveebedrijven	285	280	250	255
Grotere bedrijven	280	270	230	225
- nrd. klei/veenweide	60	65	50	45
- westelijk weide	70	60	45	50
- zandgebieden	150	140	130	115
Kleinere bedrijven	85	80	75	70
- klei/veenweide	40	45	45	45
- zandgebieden	80	80	55	60
Minder gesp. melkveebedr.	70	70	60	65
Grotere bedrijven	40	35	30	30
Kleinere bedrijven	30	30	30	30
Hokdieren-bedrijven	115	90	105	90
Fokvarkensbedrijven	65	55	40	40
Andere varkensbedrijven	35	45	40	35
Legkippenbedrijven	30	45	45	40
Overige bedrijven	80	85	65	85
Grotere bedrijven	50	60	60	65
Kleinere bedrijven	35	35	25	35

20. Steekproefopzet vanaf boekjaar 1988

20.1 Inleiding

Voor boekjaar 1988 is eind 1987 een steekproefopzet gemaakt waarbij een expliciete afweging heeft plaatsgevonden tussen de EG-hoofddoelstelling en de BUL-nevendoelelstellingen. De keuze die daaruit voortvloeide voor de verschillende strata kon nog niet geheel gerealiseerd worden, omdat ieder jaar slechts een deel van de steekproefbedrijven vervangen kan worden.

20.2 Steekproefopzet op basis van hoofddoelstelling

In eerste instantie wordt evenals bij de oude steekproefopzet een optimale stratificatie en optimale allocatie uitgevoerd om daarmee te bepalen welke strata gevormd zouden moeten worden en welke aantallen bedrijven per stratum gekozen zouden moeten worden om zo betrouwbaar mogelijke schattingen voor de gemiddelde nationale boerderij te realiseren. Bij dit vormen van strata worden vooraf al basisstrata gevormd die vrij nauw aansluiten bij de groepsindeling in de BUL, omdat dit de betrouwbaarheid van de BUL-uitkomsten aanmerkelijk kan vergroten. In plaats van het onderscheiden van 4 sbe-klassen binnen ieder type bedrijven, wordt gekozen voor een indeling in kleinere en grotere bedrijven, waarbij de grotere bedrijven nog worden onderverdeeld in 3 sbe-klassen - via optimale stratificatie aan de hand van de variabele (sbe)¹⁾ (zie paragraaf 6.6). De regionale BUL-indelingen worden al doorgevoerd in de stratificatie vòòr de optimale stratificatie plaatsvindt.

Er worden eerst 65 basisstrata gevormd. Per basisstratum wordt een steekproefpercentage bepaald, waarna een verdergaande stratificatie wordt doorgevoerd. Tabel 20.1 geeft de 65 basisstrata weer.

Optimale allocatie op basis van de variantie in (sbe)¹⁾ leidde tot het allocatie-schema in tabel 20.2. Daarbij is uitgegaan van een aantal van 1.100 te verdelen bedrijven, evenals bij de optimale allocatie in het verleden 1).

1) Eerst worden 1.200 bedrijven gealloceerd, waarna er 100 in mindering gebracht worden op de groep sterk gespecialiseerde melkveebedrijven (zie paragraaf 6.6.3). In volgende boekjaren zal de reductie van 100 bedrijven evenredig verdeeld worden over alle melkveebedrijven (zie paragraaf 12.9.3).

Tabel 20.1 Sbe-stratumgrenzen voor de basisstrata bij de bedrijfskeuze in boekjaar 1988

Stratificatie- groep *)	Sbe-klasse			
	1	2	3	4
Akkerbouwbedrijven:				
- regio 1		158 - 295	295 - 519	>= 519
- regio 3,4	} 79 - 158	158 - 252	252 - 397	>= 397
- regio 5		158 - 252	252 - 417	>= 417
- regio 6-8	79 - 158	158 - 241	241 - 404	>= 404
- regio 2, 9-13	79 - 158	158 - 248	248 - 467	>= 467
Sterk gespec. melkveebedrijven:				
- regio 1-3	} 79 - 158	158 - 241	241 - 365	>= 365
- regio 4,5,11		158 - 220	220 - 320	>= 320
- regio 6-10,12,13	79 - 158	158 - 232	232 - 340	>= 340
Minder sterk gespec. melkveebedrijven	79 - 158	158 - 237	237 - 362	>= 362
Fokvarkensbedrijven	79 - 169	169 - 259	259 - 465	>= 465
Andere varkens- bedrijven	79 - 164	164 - 263	263 - 451	>= 451
Legkippenbedrijven	79 - 187	187 - 330	330 - 641	>= 641
Andere hokdierbedr.	79 - 158	158 - 247	247 - 442	>= 442
Andere graasdier- bedrijven	79 - 158	158 - 230	230 - 365	>= 365
Gewassencombinaties	79 - 158	158 - 270	270 - 475	>= 475
Veeteeltcombinaties	79 - 158	158 - 241	241 - 382	>= 382
Gewassen/veeteelt- combinaties	79 - 158	158 - 248	248 - 427	>= 427

*) Voor een omschrijving van de regio's zie Bedrijfsuitkomsten, 1988.

Tabel 20.2 Optimale allocatie van 1.100 steekproefbedrijven over de basisstrata

Stratificatie- groep	Sbe-klasse			
	1	2	3	4
Akkerbouwbedrijven:				
- regio 1		11	11	11
- regio 3,4	26	15	15	15
- regio 5		16	16	16
- regio 6-8	13	10	10	10
- regio 2,9-13	6	4	4	4
Sterk gespec. melkveebedrijven:				
- regio 1-3	46	24	24	24
- regio 4,5,11		17	17	17
- regio 6-10,12,13	66	58	58	58
Minder sterk gespec. melkveebedrijven	25	20	20	20
Fokvarkensbedrijven	16	16	16	16
Andere varkensbedr.	17	17	17	17
Legkippenbedrijven	8	8	8	8
Andere hokdierbedr.	4	4	4	4
Andere graasdierbedr.	14	10	10	10
Gewassencombinaties	8	8	8	8
Veeteeltcombinaties	21	21	21	21
Gewassen/veeteelt- comb.	14	13	13	13

Deze tabel kan vergeleken worden met tabel 19.3 om te bekijken in hoeverre er vanuit hoofd- en nevendoelestellingen verschillende allocaties gewenst zijn. In de volgende paragraaf gebeurt dat.

20.3 Confrontatie hoofd- en nevendoelestellingen

Om te kunnen vergelijken of de hoofd- en de nevendoelestellingen met elkaar sporen op het punt van de aantallen bedrijven die in de verschillende strata gekozen moeten worden, moeten de BUL-aantallen uit tabel 19.3 nog worden opgehoogd om rekening te houden met de uitval tussen keuze en uitwerking. De volgende tabel houdt daar rekening mee. De aantallen uit de tabellen 19.3 en 20.2 zijn te vinden in de kolommen "gewenst aantal BUL" en "aantal volgens optimale allocatie". De kolom "gewenst startaantal BUL" geeft de startaantallen die nodig zijn om uiteindelijk de aantallen bedrijven uit de kolom "gewenst aantal BUL" over te houden.

Tabel 20.3 Confrontatie BUL-wensen en EG-wensen

Stratificatie- groep	Gewenst aantal BUL	Gewenst start aantal BUL	Aantal volgens optimale allocatie	Keuze- aantal (compro- mis)
Akkerbouwbedrijven:	210	250	213	306
grotere bedrijven:	210	250	168	237
- regio 1	55	60	33	60
- regio 3,4	55	60	45	60
- regio 5	55	60	48	60
- regio 6-8	40	45	30	45
- regio 2,9-13	--	--	12	12
kleinere bedrijven:	55	65	45	69
- regio 1,3-5	30	35	26	35
- regio 6-8	25	30	13	30
- regio 2,9-13	--	--	6	4
Sterk gespec. melkveebedrijven:	405	455	409	456

Tabel 20.3 (vervolg)

Stratificatie- groep	Gewenst aantal BUL	Gewenst start aantal BUL	Aantal volgens optimale allocatie	Keuze- aantal (compro- mis)
grotere bedrijven:	280	320	297	321
- regio 1-3	65	75	72	75
- regio 4,5,11	65	75	51	75
- regio 6-10,12,13	150	170	174	171
kleinere bedrijven: (*)	125	135 (100)	112	135 (100)
- regio 1-5,11	45	50	46	50
- regio 6-10,12,13	80	85	66	85
Minder sterk gespec. melkveebedrijven	70	80	85	95
- grotere	40	45	60	60
- kleinere	30	35	25	35
Hokdierenbedrijven:	115	130	164	172
Fokvarkensbedrijven	60	65	64	64
Andere varkensbedr.	40	45	68	68
Legkippenbedrijven	35	40	32	40
Overige bedrijven: (**)	35+60	40+70	61+168	40+87
Andere hokdierbedr.			4+ 12	3+ 9
Andere graasdierbedr.			14+ 30	10+18
Gewassencombinaties			8+ 24	4+12
Veeteeltcombinaties			21+ 63	15+30
Gewassen/veeteeltcomb.			14+ 39	8+18
Alle bedrijven (*)			1.100	1.156 (1.121)

*) Na inkrimping boekhoudnet met 35 bedrijven (paragraaf 20.4).

***) Kleinere bedrijven + grotere bedrijven.

In bijlage 7 is aangegeven in hoeverre de keuze-aantallen uit de laatste kolom afwijken van de huidige aantallen bedrijven in administratie en hoe groot de steekproefpercentages zijn die horen bij deze keuze-aantallen.

De grootste verschillen tussen BUL-wensen en EG-wensen betreffen de groepen grotere en kleinere akkerbouwbedrijven. Daar moet een fors aantal bedrijven bijgekozen worden om de BUL-aantallen te halen. Uit bijlage 8 blijkt dat dit geen nieuw element in de steekproefopzet is. De aantallen te kiezen bedrijven per BUL-groep liggen in de orde van grootte van de nu al in administratie zijnde aantallen. De regionale aanvullingen in het verleden (zie hoofdstuk 7) zorgen daarvoor. Ook in andere groepen bedrijven zijn grotere keuze-aantallen genomen dan op grond van de optimale allocatie gerechtvaardigd was. Naast de "reserve" van ongeveer zestig "aanvullingsbedrijven" is daarvoor gebruik gemaakt van het aantal bedrijven dat volgens de optimale allocatie zou moeten worden toegewezen aan de overige en de gecombineerde bedrijven. Vanwege de heterogeniteit van deze groep zeggen gemiddelde cijfers over deze groep niet zo veel en zegt een iets hogere of lagere standaardfout ook weinig. Daarom is het aantal grotere bedrijven in deze groep gereduceerd van 168 tot 87, en het aantal kleinere bedrijven van 61 tot 40 (laatste twee kolommen in tabel 20.3). Door deze 127 bedrijven wordt dit deel van de populatie toch meegenomen voor de gemiddelde nationale boerderij. Ten opzichte van de oude steekproefopzet zijn er extra bedrijven gekozen in de groepen "grotere sterk gespecialiseerde melkveebedrijven in regio 4, 5, 11", "kleinere minder sterk gespecialiseerde melkveebedrijven" en "legkippenbedrijven". Deze laatste groep verdiende ook uit het oogpunt van betrouwbaarheid van de uitkomsten van de nationale boerderij een sterkere vertegenwoordiging in de steekproef (paragraaf 12.9.3).

20.4 Kleine inkrimping boekhoudnet

Eind 1988 is besloten het boekhoudnet landbouw met 35 bedrijven in te krimpen. De ruimte daarvoor is gevonden in het laten vallen van de BUL-eis dat er voor de groep "kleinere sterk gespecialiseerde melkveebedrijven" een indeling in twee regionale groepen gepubliceerd moet worden en dat daarvoor 125 bedrijven nodig zijn. In plaats van het oorspronkelijk geplande keuzeaantal van 135 bedrijven in de totale groep, is uitgegaan van 100 bedrijven. Dat betekent niet dat daarmee de mogelijkheid om over de afzonderlijke gebieden te publiceren van de baan is. De resultaten zullen dan wel een lagere betrouwbaarheid hebben.

20.5 Te verwachten betrouwbaarheid in boekjaar 1988

In hoofdstuk 19 is aangegeven dat na de keuze voor boekjaar 1988 extra informatie beschikbaar is gekomen over het verband tussen start-aantallen steekproefbedrijven en te verwachten betrouwbaarheid van de resultaten. Deze informatie kan gebruikt worden bij het eventueel bijsturen van de keuze in volgende boek-

jaren, indien de te verwachten betrouwbaarheid onvoldoende is voor bepaalde BUL- of FIP-groepen.

De tabellen 20.4 en 20.5 geven aan welke resultaten verwacht mogen worden bij de start-aantallen die in 1988 zijn gehanteerd. De varianties zijn geschat op basis van de boekjaren 1984-1986, de non-responspercentages in de uitwerkingsfase zijn ontleend aan de boekjaren 1982-1986 (zie hoofdstuk 10), terwijl de geschatte FIP-respons (zie paragraaf 10.5) betrekking heeft op de boekjaren 1982-1986 (zie ook Aukema en Overgaw, 1988:68).

De groepen in tabel 20.4 en 20.5 zijn gevormd op basis van het bedrijfstype in de metelling en de sbe-grootte gemiddeld gedurende het boekjaar. Bedrijfstypering op basis van gegevens gemiddeld gedurende het boekjaar was nog niet realiseerbaar met de beschikbare programmatuur. De standaardfouten zijn berekend met vergelijking 6.4.5, waarbij de groepen in tabel 20.5 als basisstrata zijn aangemerkt en deze basisstrata onderverdeeld worden in de gebruikelijke vier sbe-klassen uit de oude opzet.

De gekozen opzet zorgt ervoor dat de standaardfout van "opbrengsten per f 100,- kosten" voor de meeste groepen kleiner is dan f 2,-. Voor de hoofdgroepen "akkerbouw totaal", enz. zijn ze vanzelfsprekend beduidend lager.

De standaardfouten voor de "factoropbrengst" zijn - uitgezonderd enkele akkerbouw-groepen en de groep legkippen - niet of nauwelijks groter dan f 6.000,-. Dat zelfde geldt voor de standaardfouten van de grootheden "gezinsinkomen uit bedrijf" en "besparingen".

Voor de nationale boerderij worden vrij kleine standaardfouten gerealiseerd. Deze liggen in dezelfde orde van grootte als bij de oude opzet.

Tabel 20.4 Te verwachten aantallen bedrijven voor BUL- en FIP-groepen in boekjaar 1988 op basis van startaantallen per 1 mei (tussen haakjes de "respons"-percentages)

Stratificatie-groep	Startaantal boekhoudnet	Verwacht aantal BUL-bedrijven	Verwacht aantal FIP-bedrijven
Akkerbouwbedrijven :	306	290 (96%)	215(74%)
grotere bedrijven:	237	225 (96%)	160(72%)
- regio 1	60	55 (94%)	30(55%)
- regio 3,4	45	45(100%)	35(83%)
- regio 5	60	55 (94%)	45(81%)
- regio 6-8	60	60(100%)	45(74%)
- regio 2,9-13	12		
kleinere bedrijven:	69	65 (96%)	55(86%)
- regio 1,3-5	35	30 (83%)	25(82%)
- regio 6-8	30	30(100%)	25(94%)
- regio 2,9-13	4		
Sterk gespecialiseerde melkveebedrijven	421	365 (86%)	310(86%)
grotere bedrijven:	321	265 (84%)	220(83%)
- regio 1-3	75	65 (88%)	50(79%)
- regio 4,5,11	75	60 (82%)	50(80%)
- regio 6-10,12,13	171	140 (83%)	120(85%)
kleinere bedrijven:	100	95 (95%)	90(96%)
- regio 1-5,11	40	35 (90%)	35(98%)
- regio 6-10,12,13	60	60 (98%)	55(96%)
Minder sterk gespecialiseerde melkveebedrijven	95	85 (90%)	80(94%)
- grotere	60	50 (85%)	45(93%)
- kleinere	35	35 (95%)	35(98%)
Hokdierenbedrijven:	172	160 (95%)	135(86%)
Fokvarkensbedrijven	64	60 (97%)	45(79%)
Andere varkensbedrijven	68	65 (97%)	60(94%)
Legkippenbedrijven	40	35 (92%)	30(86%)
Overige bedrijven:	127	125 (95%)	110(89%)
Grotere bedrijven	87	85 (95%)	70(85%)
Kleinere bedrijven	40	40 (95%)	40(99%)
Alle bedrijven	1.121	1.025 (91%)	850(86%)

Tabel 20.5 Te verwachten standaardfout van het gemiddelde in boekjaar 1988 voor BUL- en FIP-groepen op basis van startaantallen per 1 mei 1988, naar variabele (in gulden)

Stratificatie- groep	Opbrengst/ f 100,- kosten	Factor- opbrengst	Gezinsin- komen uit bedrijf	Besparin- gen
Akkerbouwbedrijven:	0,60	3.500	3.500	3.500
grotere bedrijven:	1,00	5.000	5.000	5.000
- regio 1	2,70	13.500	11.500	10.000
- regio 3,4	1,90	6.000	7.000	7.000
- regio 5	1,80	6.500	8.000	9.500
- regio 6-8	1,90	7.000	7.500	7.500
- regio 2,9-13				
kleinere bedrijven:	2,30	3.500	3.500	4.000
- regio 1,3-5	2,90	5.000	5.000	5.500
- regio 6-8	2,10	3.500	4.500	4.500
- regio 2,9-13				
Sterk gespecialiseerde melkveebedrijven	0,50	2.000	2.000	1.500
grotere bedrijven:	0,70	2.500	2.500	2.500
- regio 1-3	1,30	5.500	4.500	4.500
- regio 4,5,11	1,50	5.000	4.000	4.500
- regio 6-10,12,13	0,90	3.500	3.500	3.500
kleinere bedrijven:	1,10	2.000	2.500	2.000
- regio 1-5,11	1,50	3.000	3.500	3.500
- regio 6-10,12,13	1,60	3.000	3.000	3.000
Minder sterk gespeciali- seerde melkveebedrijven:	1,10	3.000	3.000	3.000
- grotere	1,20	4.500	5.000	5.000
- kleinere	1,60	3.500	3.500	3.500
Hokdierenbedrijven:	0,90	4.500	4.500	4.000
Fokvarkensbedrijven	1,40	6.000	5.000	5.000
Andere varkensbedr.	1,00	6.500	7.000	5.500
Legkippenbedrijven	2,20	12.000	12.500	11.000
Overige bedrijven:	1,40	4.500	4.500	4.500
Grotere bedrijven	1,50	6.000	6.500	7.000
Kleinere bedrijven	2,50	6.000	6.000	6.000
Alle bedrijven	0,40	1.400	1.400	1.300

21. Vergelijkbaarheid oude en nieuwe opzet

21.1 Inleiding

Om een indruk te geven van de verschillen tussen de oude steekproefopzet en de opzet vanaf boekjaar 1988, worden in dit hoofdstuk enkele tabellen gepresenteerd. Paragraaf 2 geeft weer hoe de bedrijfstypen die bij de stratificatie gehanteerd worden, op elkaar aansluiten. Vervolgens wordt in paragraaf 3 nagegaan hoe de nieuwe stratificatie de resultaten voor de nationale boerderij beïnvloedt. Paragraaf 4 gaat in op de veranderingen in de BUL.

21.2 Populatie en bedrijfstypen bij de stratificatie

Tabel 21.1 geeft weer hoe de oude en nieuwe bedrijfstypen bij de stratificatie met elkaar corresponderen. Het gaat bij beide bedrijfstyperingen om bedrijven met een minimale omvang van 79 sbe en een maximale omvang van 2.000 sbe.

Tabel 21.1 Aantallen per bedrijfstype; vergelijking oude en nieuwe stratificatie-typen (mettelling 1986)

NEG- type	LEI-type								
	akker- bouw	rund vee	var- kens	pluim vee	gem. akk	gem. rund	gem. rund	gem. int	tuin- bouw
akkerbouw	10624	5	0	0	99	0	0	0	86
st.gesp.melkv.	0	33131	0	0	0	26	64	0	0
m.gesp. melkv.	0	3771	0	0	0	176	866	0	0
fokvarkens	0	0	3230	0	0	0	0	8	0
and.varkens	0	0	3429	0	0	0	0	56	0
legkippen	0	0	0	1219	0	0	0	3	0
ov.hokdieren	0	0	25	445	0	0	0	390	0
mestk/ov.graasd.	0	388	0	0	34	154	252	1085	0
sch./geiten/gr.	0	722	0	0	2	75	93	10	0
gewassencomb.	645	0	0	0	150	0	0	1	829
veeteeltcomb.	6	23	263	55	127	197	2137	2101	6
gew/veet.comb.	165	46	94	21	1042	713	0	569	232
tuinbouw	91	345	0	0	0	23	49	0	18240

Duidelijk is dat groepen die qua terminologie op elkaar lijken, in grote lijnen dezelfde bedrijven bevatten. De groep akkerbouwbedrijven wordt in het nieuwe systeem iets scherper afgebakend. Een deel van de "oude" akkerbouwbedrijven verdwijnt naar de bedrijfstypen "gewassencombinaties" en "gewassen/veeteelt-combinaties" en "tuinbouw".

De groepen "sterk gespecialiseerde" en "minder sterk gespecialiseerde melkveebedrijven" bevatten samen een groot deel van de "rundveebedrijven". Laatstgenoemde groep bevat bovendien een deel van de "gemengde rundveehouderij/intensieve veehouderij". Een deel van de "(gemengde) rundveebedrijven" komt terug in de nieuwe typen "mestkalveren en overige graasdieren", "schapen, geiten en overige graasdieren" en "tuinbouw". Van de bedrijven met "gemengde rundveehouderij/intensieve veehouderij" komt bijna de helft terecht in de groep "veeteeltcombinaties".

Het oude type "varkenshouderij" wordt gesplitst in "fokvarkens" en "andere varkens". De resterende bedrijven verdwijnen naar de combinatie-typen.

Het oude type "pluimveehouderij" is terug te vinden in de NEG-groepen "legkippen" en "andere hokdieren". Deze laatste groep wordt aangevuld met bedrijven afkomstig uit de categorie "gemengd intensief".

De vier groepen gemengde bedrijven in de oude typering worden herverdeeld over een aantal nieuwe groepen combinatie-bedrijven, terwijl ook een deel terecht komt in het niet-sterk op melkveehouderij gespecialiseerde deel van de rundveebedrijven.

Wat de totale populatie betreft, valt op dat daaruit bedrijven verdwijnen, maar dat er ook bedrijven toetreden. Totaal verdwijnen er 508 bedrijven, met name uit de oude groep "rundveebedrijven". Door een andere afbakening tussen land- en tuinbouw, worden deze nu getypeerd als "tuinbouwbedrijven". Anderzijds wordt de populatie vergroot met 1.153 bedrijven, voornamelijk bedrijven die nu thuishoren in de NEG-groepen "gewassencombinaties" en "gewassen/veeteeltcombinaties". Per saldo wordt de populatie 645 bedrijven groter.

21.3 De oude en de nieuwe nationale boerderij

Door de lichte veranderingen die de populatie heeft ondergaan is de nieuwe nationale boerderij niet geheel vergelijkbaar met de oude nationale boerderij. De procentuele verdeling van het totaal aantal sbe over de verschillende gewassen en dieren zal licht veranderen. De plantaardige sector zal iets sterker naar voren komen en de rundveehouderij iets minder sterk (zie vorige paragraaf).

Tabel 21.2 geeft aan hoe de populatie er in meitellingstermen uitziet voor boekjaar 1986. Daarbij is er van uitgegaan dat

de sbe-ondergrens gehandhaafd blijft op 79 sbe l). Het totaal aantal bedrijven in de populatie neemt met 1,5% toe. Het totaal aantal sbe neemt met 1,7% toe, waardoor de bedrijven gemiddeld een fractie groter worden. Het aantal koeien neemt met minder dan 0,1% toe, de oppervlakte bouwland daarentegen met 1,3% en de oppervlakte tuinbouw open grond neemt met 32% toe tot ruim 30.000 hectare. Dit laatste is het gevolg van het de wijze waarop de bedrijven met akkerbouwmatige groenten in de NEG-typering worden opgenomen.

Berekening van totalen of gemiddelden voor de nationale boerderij, zal leiden tot licht gewijzigde bedrijfsresultaten. Tabel 21.2 geeft aan waar die wijzigingen door veroorzaakt worden.

Er is echter nog een andere reden waarom de totale of de gemiddelde nationale boerderij niet geheel vergelijkbaar is met de oude nationale boerderij. Vanaf 1988 wordt immers een nieuw stratificatie-ontwerp gebruikt, waardoor bepaalde groepen bedrijven die in het verleden oververtegenwoordigd waren, nu minder sterk naar voren zullen komen in de nationale boerderij. Tot nu toe ondervertegenwoordigde bedrijven zullen hopelijk sterker naar voren gaan komen.

Berekening van de uitkomsten voor de nationale boerderij met wegingsfactoren die berekend worden op basis van de nieuwe stratificatie, wijzen al enigszins in die richting. Een goede vergelijking is voor boekjaar 1986 nog moeilijk te maken, omdat bedrijven die tot nu toe ondervertegenwoordigd zijn in de steekproef, ook bij een hernieuwde weging niet in de steekproef voorkomen, waardoor sommige strata nog bijna of geheel leeg zijn voor wat betreft de steekproefbedrijven. De volgende uitkomsten geven dan ook niet meer dan een aanwijzing.

Het gemiddeld aantal melkkoeien per bedrijf - gemiddeld gedurende het boekjaar - neemt bij de nieuwe weging af met ongeveer 3%. Daarbij dient echter bedacht te worden dat het aantal bedrijven in de nieuwe populatie 1,5% groter is, terwijl het totaal aantal melkkoeien in de populatie ongeveer gelijk is gebleven (tabel 21.2). Wordt gecorrigeerd voor de veranderingen in de populatie, dan is het gemiddeld aantal melkkoeien per bedrijf 1,5% lager.

De oppervlakte cultuurgrond per bedrijf neemt licht af (0,7%), terwijl de oppervlakte marktbaar gewassen licht toeneemt. Daarbij spelen zowel de veranderde populatie als de veranderde wegingsfactoren een rol. Het sbe-aantal per bedrijf neemt 1,0% toe; dat is meer dan op grond van de populatie-verandering verwacht mocht worden. Het aandeel van de intensieve veehouderij in de nationale boerderij wordt iets vergroot.

1) In 1989 is deze grens gelegd bij 20 nge, waardoor de samenstelling van de populatie opnieuw iets gewijzigd is.

Tabel 21.2 Overzicht van totalen en gemiddelden van de nieuwe nationale boerderij (mettelling 1986)

Variabele	Totaal	Verandering ten opzichte van de oude nationale boerderij		
		totaal	aantal bedrijven	gem. waarde
Aantal bedrijven	70.306	+1.5	+1.5	----
Aantal rechtspersonen	843	+3.7	+3.7	----
Sbe-totaal	15.177.258	+1.7	+1.5	+0.2
Nge-totaal	4.413.388	+1.6	+1.5	+0.1
Arbeidskrachten:				
- mannelijk totaal	96.679	+1.9	+1.5	+0.4
- mannelijk vreemd	8.228	+4.9	+3.1	+1.7
- vrouwelijk totaal	27.603	+1.3	+1.3	+0.0
- vrouwelijk vreemd	404	+3.3	+4.2	-0.8
Oppervlakten:				
- bouwland	652.015	+1.3	+1.9	-0.6
- blijvend grasland	933.376	+0.3	+0.7	-0.4
- tijdelijk grasland	29.674	+1.4	+1.9	-0.5
- tuinbouw open grond	30.446	+32.1	+15.9	+13.9
- tuinbouw onder glas	52	+164.8	+97.0	+34.4
- braakland	2.011	+6.7	+12.8	-5.3
- cultuurgrond totaal	1.647.404	+1.2	+1.6	-0.4
- aantal melkkoeien	2.201.563	+0.1	+0.2	-0.1
Sbe:				
- rundvee fokkerij	6.875.149	+0.1	+0.3	-0.2
- mestkalveren	190.484	+0.6	+0.8	-0.2
- rundvee mesterij	203.116	+0.6	+0.9	-0.3
- schapen	71.296	+1.6	+1.0	+0.6
- paarden	110.364	+0.7	+0.8	-0.1
- mestvarkens	896.736	+0.8	+1.0	-0.2
- fokvarkens	1.486.873	+0.5	+0.7	-0.2
- slachtkuikens	102.987	+0.9	+1.2	-0.3
- leghennen (incl. md)	458.510	+0.3	+0.9	-0.6
- granen	378.140	+1.3	+2.7	-1.4
- suikerbieten	876.739	+2.5	+3.7	-1.0
- pootaardappelen	409.881	+3.9	+6.2	-2.2
- consumptieaardappelen	493.685	+2.4	+3.8	-1.3
- fabrieksaardappelen	232.048	+0.4	+0.6	-0.1
- peulvruchten	69.564	+1.2	+1.9	-0.7
- handelsgewassen	21.908	+1.0	+1.5	-0.4
- landbouwzaden	56.575	+1.7	+2.3	-0.6
- overige akkerbouw	826.501	+1.1	+1.4	-0.3
- voedergewassen	549.790	+0.6	+0.8	-0.2
- tuinbouwgewassen	331.188	+112.7	+20.6	+76.4
- tuinakkergewassen	166.674	- 2.1	+3.3	-5.2

Zowel de totale kosten als de totale opbrengsten nemen met enkele procenten toe. De arbeidsopbrengst en de factoropbrengst van het gemiddelde bedrijf stijgen ongeveer 1%, evenveel als de stijging in het aantal sbe. Het netto overschot stijgt iets meer. Dat hangt samen met een kleiner aantal arbeidsuren van de ondernemer. De arbeidsopbrengst per ondernemer daalt licht. Het deel van de arbeidskosten dat uitbetaald wordt, ligt hoger dan in de oude berekeningen. De gemiddelde leeftijd van de ondernemer neemt iets toe.

21.4 Oude en nieuwe BUL-groepen

Van de bedrijven die in boekjaar 1986 meededen in de berekeningen voor de BUL is bekend welk BUL-bedrijfstype ze hadden op basis van boekhoudnetgegevens. Ook is bekend welk NEG-stratificatie-type ze hadden in de metelling van 1986. Berekening van het NEG-bedrijfstype op basis van boekhoudnetgegevens is echter nog niet realiseerbaar 1). Daardoor is een precieze vergelijking van oude en nieuwe BUL-typen niet helemaal mogelijk. Een vergelijking van de NEG-metellingstypes met de oude BUL-typen zal echter wel een goede indicatie geven van de aansluiting van de twee typering.

Poppe (1984) heeft een dergelijke vergelijking uitgevoerd. Hij geeft aan dat er geen grote verschillen zijn tussen de twee indelingen. De meeste weidebedrijven komen terecht in de categorie gespecialiseerde melkveebedrijven. Van de gemengde bedrijven met overwegend rundveehouderij wordt een klein deel van de bedrijven eveneens ingedeeld bij de gespecialiseerde melkveebedrijven. De overige bedrijven in deze groep worden grotendeels ingedeeld bij de minder op melkvee gespecialiseerde graasdierenbedrijfstypen.

De gemengde bedrijven met intensieve veehouderij - een zeer heterogene groep in de BUL - worden over diverse NEG-typen verdeeld.

Om vergelijking van bedrijfseconomische en financiële gegevens tussen verschillende boekjaren mogelijk te maken, zullen de bedrijfsresultaten van boekjaar 1987 en enkele voorgaande jaren zowel voor de oude als de nieuwe BUL-typering door het LEI worden gepubliceerd.

1) De computer-programmatuur die op basis van metellingsgegevens het NEG-type van een bedrijf berekent, is nog niet toepasbaar op boekhoudnetgegevens.

22. Conclusies

Het LEI-boekhoudnet is vanaf de oprichting van het LEI in beweging geweest. Zowel de mogelijkheden om gegevens te verzamelen als de eisen die aan te verzamelen gegevens gesteld werden, waren daarbij aan verandering onderhevig.

De gerichte bedrijfskeuze tot het midden van de zestiger jaren was er de oorzaak van dat de bedrijfsuitkomsten op LEI-bedrijven minder variatie vertoonden dan in de populatie het geval was. Er zijn echter weinig aanwijzingen dat het niveau van de bedrijfsuitkomsten systematisch te hoog of te laag was. De overgang op een steekproef in 1966 heeft een sterkere bewaking van de representativiteit met zich meegebracht. Waar systematische verschillen tussen steekproef en populatie werden gesignaleerd, werd de steekproef bijgesteld. Daarnaast ging de betrouwbaarheid van steekproefuitkomsten een steeds grotere rol spelen. Onder invloed van EG-verplichtingen werd de hoofddoelstelling van het boekhoudnet verlegd naar het betrouwbaar schatten van grootheden voor de totale nationale boerderij. Het betrouwbaar schatten voor bepaalde groepen binnen de populatie werd een nevendoelstelling.

Die lijn werd voortgezet in 1975 toen een vernieuwde steekproefopzet werd geïntroduceerd. De doelstelling werd ook explicieter geformuleerd, namelijk in termen van specifieke doelvariabelen. Geconcludeerd kan worden dat de opzet van deze steekproef vrij optimaal is gezien de doelstellingen. De indeling van de populatie in strata en de verdeling van de 1.150 steekproefbedrijven over die strata leiden in het algemeen tot betrouwbare schattingen die ten aanzien van veel kenmerken representatief zijn voor de bedrijven in de populatie.

De belangrijkste diersoorten en gewassen zijn op nationaal niveau maximaal enkele procenten over- of ondervertegenwoordigd ten opzichte van de populatie. In de akkerbouw zijn de relatief weinig voorkomende gewassen in de meeste gevallen oververtegenwoordigd. In de rundveehouderij zijn de weinig voorkomende diersoorten daarentegen ondervertegenwoordigd. Verschillen in geschiktheid en bereidheid om deel te nemen aan de steekproef tussen bedrijven met verschillende gewassen en diersoorten liggen daaraan ten grondslag. Vergelijking met externe gegevensbronnen laat zien dat de fysieke opbrengsten per hectare gewas en per diersoort in de steekproef niet aan de hoge of de lage kant liggen. Aan de kostenzijde van de produktie is het moeilijker om te bepalen of de steekproefuitkomsten representatief zijn.

De uitgevoerde analyses met het oog op de representativiteit leveren als totaalbeeld dat de bedrijfsuitkomsten die op grond van de steekproef worden berekend, niet systematisch naar boven

of naar beneden afwijken ten opzichte van de bedrijfsuitkomsten in de achterliggende populatie.

De betrouwbaarheid van schattingen voor de nationale boerderij is bovendien zodanig dat toevallige afwijkingen tussen steekproef en populatie in de meeste gevallen niet meer dan enkele procenten bedragen. Toevallige fouten voor bepaalde groepen binnen de populatie - bijvoorbeeld de groepen waarover in de BUL en de FIP wordt gepubliceerd - zijn in het algemeen groter, maar de standaardfouten zijn niet onaanvaardbaar hoog. De toewijzing van steekproefbedrijven aan specifieke strata ten behoeve van bepaalde BUL-groepen waarover afzonderlijk gepubliceerd moet worden, brengt een zeker evenwicht tussen de hoofd- en de nevendoelelstellingen. Deze toewijzing van zestig bedrijven los van de allocatie voor de hoofddoelelstelling heeft er toe geleid dat voor verschillende groepen bedrijven afzonderlijke uitkomsten gegeven konden worden waar dit anders niet het geval zou zijn geweest, terwijl daardoor aan de hoofddoelelstelling nauwelijks afbreuk is gedaan.

Doordat de groepen waarover gepubliceerd wordt in het algemeen niet samenvallen met de groepen die bij de stratificatie gevormd worden, vindt in sommige gevallen een onnodige vergroting van de standaardfouten plaats. Met het oog daarop is bij de nieuwe steekproefopzet met ingang van boekjaar 1988 voor een nauwere aansluiting van de strata bij de BUL-groepen gekozen. Ook is veel explicieter aangegeven van welke doelelstellingen wordt uitgegaan ten aanzien van de verschillende BUL- en FIP-groepen. Deze nevendoelelstellingen krijgen ook een iets zwaarder gewicht ten opzichte van de hoofddoelelstelling dan in het verleden.

Veel van de conclusies die getrokken zijn op basis van ervaring met de steekproef sinds 1975 zijn meegenomen bij de verandering van de steekproef in 1988. Ook in de nieuwe opzet zal er echter sprake zijn van hoge non-respons-percentages waarbij de selectiviteit van de non-respons een bedreiging is voor de representativiteit van de steekproefuitkomsten. Evenals in het verleden moet de manier van trekken van de steekproefbedrijven en de manier van berekenen van steekproefuitkomsten garanderen dat er ondanks de selectieve non-respons geen vertekeningen ontstaan in de bedrijfsuitkomsten. Nieuwe methoden om wegingsfactoren aan bedrijven toe te kennen, kunnen daarbij mogelijk zorgen voor een grotere representativiteit. Nader onderzoek naar de toepassing van bijvoorbeeld een algemene regressieschatter is daarvoor nodig.

Daarnaast zal geprobeerd moeten worden om de respons van de potentiële deelnemers te vergroten. Met een snellere terugkoppeling van de gegevens naar de deelnemers door een verdergaande automatisering en het uitvoeren van aanvullende analyses worden stappen gezet op weg naar het aantrekkelijker maken van deelname. Maar het is duidelijk dat in een tijd waarin de boer(in) tal van mogelijkheden krijgt aangeboden om via bedrijfsoverzichten en -analyses meer inzicht te krijgen in de gang van zaken op het be-

drijf, het LEI veel aandacht zal moeten geven aan de werving van deelnemers voor de steekproef. Ondanks alle mogelijkheden om achteraf te corrigeren voor non-respons blijft voorkomen ervan altijd nog beter.

Literatuur

STATISTIEKEN VAN BEDRIJFSUITKOMSTEN

Aukema, S.

De financiële positie van landbouwbedrijven in het boekjaar 19.../..., (jaarlijks, boekjaren 1970/71-1972/73)
Den Haag, LEI, 1972-1974

Aukema, S. en G. van Donkelaar

De financiële positie van landbouwbedrijven; boekjaar 19.../...,
(boekjaren 1964/65/'65-1969/70)
Den Haag, LEI, 1969-1972)

Aukema, S. en J.G.A. Overgaw

De Financiële positie van de landbouw: Boekjaar 19.../.. (en vergelijkingen met voorgaande jaren),
(jaarlijks, boekjaren 1973/74 t/m 1986/87)
Den Haag, LEI, 1976-1988; Periodieke Rapportage 12

Bedrijfsuitkomsten in de landbouw: 19.../.. t/m 19.../..),
(jaarlijks, boekjaren 1965/66 t/m 1968/69 - 1983/84 t/m 1986/87)
Den Haag, LEI, 1970-1988; Periodieke Rapportage 11

Centraal Bureau voor de Statistiek

Statistiek van bedrijfsuitkomsten in de landbouw, 19.../...,
(jaarlijks, boekjaren 1950/51-1963/64)
Hilversum, de Haan, 1954-1966

Centraal Bureau voor de Statistiek

Statistiek van de bedrijfsuitkomsten in de landbouw 19.../..., Bedrijfseconomische exploitatie-rekeningen,
(jaarlijks, boekjaren 1963/64-1966/67)
Den Haag, Staatsuitgeverij, 1966-1968

Centraal Bureau voor de Statistiek

Statistiek van de bedrijfsuitkomsten in de landbouw 19.../...,
(jaarlijks, boekjaren 1971/72-1973/74)
in: Maandstatistiek van de Landbouw (juni)
Den Haag, CBS, 1974-1976

Centraal Bureau voor de Statistiek

Produktierekening van de land- en tuinbouw 19.../...,
(jaarlijks, boekjaren 1975/76-1985/86)
Den Haag, Staatsuitgeverij, 1978-1988

LITERATUUR (1e vervolg)

Directie van den Landbouw
Bedrijfsuitkomsten van den Landbouw in het boekjaar 19../19..,
(boekjaren 1923-1939)
in: De Economische Toestand van den Landbouw
Den Haag, Directie van den Landbouw, 1924-1940

Directie van den Landbouw
Bedrijfsuitkomsten van den Landbouw 1942-1943 (voorlopige cijfers)
Den Haag, Directie van den Landbouw, 1944

Horring, J.
Rapport inzake het onderzoek naar de kostprijzen van de voornaamste akkerbouwprodukten voor den oogst 1940
Den Haag, LEI, 1940; LEI-rapport no.1

Koppejan, A.W.G.
Bedrijfseconomisch Onderzoek ten dienste van de landbouw
Den Haag, Ministerie van Landbouw, 1951; Verslagen van Landbouwkundige Onderzoekingen No. 54.10

LEI
Statistiek van de bedrijfsuitkomsten in de landbouw,
(jaarlijks, boekjaren 1946/47-1949/50)
Den Haag, LEI, 1949, 1950a, 1950b, 1951.

LEI
Gemiddelde uitkomsten van een aantal groepen bedrijven over
1947/48 t/m 1950/51 en over 1950/51 t/m 1951/52
Den Haag, LEI, 1952/1954; BEM no.9, BEM no.13

LEI
Uitkomsten van weidebedrijven van 1947/48 tot en met 1950/51
Den Haag, LEI, 1952a; BEM no. 5

LEI
Uitkomsten van gemengde bedrijven op zandgrond over 1948/49 tot en met 1950/51
Den Haag, LEI, 1952b; BEM no. 6

LEI
Statistisch Overzicht van de uitkomsten van landbouwbedrijven in 19../.., (boekjaren 1952/53 t/m 1962/63)
Den Haag, LEI, 1954-1964; BEM no. 14, 18, 20, 23, 25, 30, 32, 40, 41;
Verslagen No. 75,

LITERATUUR (2e vervolg)

Ministerie van Landbouw

Verslag over de landbouw in Nederland; bedrijfsuitkomsten over de jaren 1946, 1947 en 1948

Den Haag, Ministerie van Landbouw, 1951

Poppe, K.J. (red.)

Van bedrijfsuitkomsten tot financiële positie (BEF); Samenvattend overzicht van landbouwbedrijven tot en met 19../.., (jaarlijks, boekjaren 1975/76-1986/87)

Den Haag, LEI, 1977-1988; Periodieke Rapportage 13

Prognose van bedrijfsuitkomsten op akkerbouw- en veehouderijbedrijven in 19../.., (jaarlijks, boekjaren 1975/76-1987/88)

Den Haag, LEI, 1975-1988; Periodieke Rapportage 14

LEI-BOEKHOUDNET

Dijk van, J.

De bedrijfskeuze voor het boekhoudnet, (jaarlijks, boekjaren 1975-1987)

Den Haag, LEI, 1975-1987; interne notitie, niet gepubliceerd

LEI

Tien jaren Landbouw-Economisch Instituut; opbouw van het economische onderzoek in landbouw, tuinbouw en visserij

Assen, van Gorcum, 1950

Lodder, K.

Het boekhoudnet landbouwbedrijven (periode: boekjaar 1975 t/m 1983/84)

Den Haag, LEI, 1983; interne notitie, niet gepubliceerd

Lodder, K.

Het boekhoudnet landbouwbedrijven, een statistische verantwoording

Den Haag, LEI, 1987; Mededeling 358

Plas, M. van der

Trekking en tussentijdse aanpassing van de steekproef voor het LEI-boekhoudnet tuinbouw, theoretische achtergronden

Den Haag, LEI, 1985; Onderzoekverslag no. 16

LITERATUUR (3e vervolg)

Foppe, K.J., en J. Dijk
Het boekhoudnet landbouw en de introductie van de NEG-typing
Den Haag, 1986; niet-gepubliceerde discussienota

Snoek, Th.J.
Representativiteit van LEI-bedrijven
Den Haag, LEI, 1963; LEI-studies no.9

STEEKPROEF-THEORIE

Bethlehem, J.G.
Representativiteit: Wat, waarom en hoe?
Voorburg, CBS, 1985; Intern rapport

Bethlehem, J.G., and W.J.Keller
"Linear weighting of sample survey data"
Journal of Official Statistics 3(1987), pp.141-153

Bethlehem, J.G., en H.M.P. Kersten
Werken met non-respons
Den Haag, CBS, 1986; CBS-Statistische onderzoeken M30

Brackstone, G.J., and J.N.K. Rao
"An investigation of raking ratio estimators"
Sankhya : The Indian Journal of Statistics 41(1979)
Series C, Part 2, pp.97-114

Cochran, W.G.
Sampling techniques
New York, Wiley, 1977

Kish, L., and M.R. Frankel
"Inference from Complex Samples"
Journal of the Royal Statistical Society 36(1974)
Ser. B, pp. 1-37

Kish, L.
Survey Sampling
New York, Wiley, 1965

Mood, A.M., F.A. Graybill, and D.C. Boes
Introduction to the theory of statistics
Tokyo, Mc Graw-Hill, 1974

LITERATUUR (4e vervolg)

Moors J.J.A., en J. Muilwijk
Steekproeven; een inleiding tot de praktijk
Amsterdam, Elsevier, 1975

Muilwijk, J.
Statistiek en samenleving, drie voordrachten voor de cursus We-
tenschap en Samenleving
Groningen, Econometrisch Instituut, 1981

Platek, R. et al (eds.)
Small area statistics, an international symposium
New York, Wiley, 1987

Purcell, N.J., and L. Kish
"Estimation for Small domains"
Biometrics 35(1979), pp.365-384

Sarndal, C.E.
"Design-consistent versus model-dependent estimation
for small domains"
Journal of the American Statistical Association 79(1984),
pp.624-631

OVERIGE

CBS
Typering van de Nederlands landbouwbedrijven volgens de NEG-type-
ring
Den Haag, Staatsuitgeverij, 1986

CBS
Het NEG-systeem van bedrijfstypering
In: Landbouwtelling 1985 en 1986
Den Haag, CBS, 1988

Dag, B. voor den
De Nederlandse landbouwbedrijven in 1986 getypeerd volgens de
NEG-typologie: spreiding-activiteiten-grootte-omvang
Den Haag, Ministerie van Landbouw, 1988

Dröge, H., K.J. Poppe en H. Prins
Toepassing van standaardbedrijfseenheden in onderzoek en voor-
lichting op landbouwbedrijven
Den Haag, LEI, 1983; Mededeling 273

LITERATUUR (5e vervolg)

Landbouwcijfers

Den Haag/Voorburg, LEI/CBS, diverse jaren

Poppe, K.J.

Het EG-boekhoudnet

In: Van bedrijfsuitkomsten tot financiële positie (BEF)
(boekjaar 1982/83)

Den Haag, LEI, 1984; Periodieke Rapportage 13-83/84

Poppe, K.J.

Nederlandse bruto standaard saldi voor bedrijfstypering in EG-
verband (bss 1982)

Den Haag, LEI, 1986; Mededeling 347

Wijnands J. et al.

Het kunstmestgebruik in de land- en tuinbouw in 1979/1980; het
gebruik van de landbouwtelling om de betrouwbaarheid van popula-
tiegegevens van steekproefanalyses te vergroten

Den Haag, LEI, 1983; LEI-publikatie 3.125

Bijlagen

Bijlage 1 Stratifikatie bij de steekproefopzet 1975-1987

Indeling van de populatie in strata

De landbouwbedrijven uit de metelling worden ingedeeld in groepen (strata) aan de hand van de volgende vijf criteria:

A De produktierichting:

Deze wordt berekend met behulp van een sbe-indeling aan de hand van de metellingsgegevens. Er wordt onderscheid gemaakt in:

- I - sbe akkerbouw en tuinakkergewassen
- II - sbe grasland en weidedieren
- III - sbe varkenshouderij
- IV - sbe pluimveehouderij
- V - sbe overige intensieve veehouderij
- VI - sbe tuinbouw
- VII - sbe totaal

Een bedrijf hoort niet bij de populatie van landbouwbedrijven indien minstens 50% van het totale aantal sbe tot "tuinbouw" behoort.

Er worden acht bedrijfstypen onderscheiden op basis van de verdeling van het totaal aantal sbe (VII) over de zes onderscheiden sbe-groepen:

A	: Akkerbouwbedrijven	: I + VI > 2/3 * VII
R	: Rundveebedrijven	: II > 2/3 * VII
Va	: Varkensbedrijven	: III > 2/3 * VII
Pl	: Pluimveebedrijven	: IV > 2/3 * VII
A+	: Akkerbouw/gemengd	: I+VI > II; I+VI > III+IV+V
RA	: Rundvee/Akkerbouw	: I+VI < II; I+VI > III+IV+V
RV	: Rundvee/Int. veehouderij	: II >= III+IV+V; I+VI <= III+IV+V
V+	: Int. veehouderij/gemengd	: II < III+IV+V; I+VI <= III+IV+V

Een bedrijf wordt dus bijvoorbeeld een rundveehouderijbedrijf genoemd indien meer dan 2/3 van het totaal aantal sbe betrekking heeft op de rundveehouderij.

B De bedrijfsgrootte:

Elk bedrijfstype wordt onderverdeeld in vier sbe-klassen op basis van statistische criteria (paragraaf 6.6). Voor boekjaar 1987 zijn de volgende klassen aangehouden, gebaseerd op metelling 1985:

Bijlage 1 (le vervolg)

Tabel Bl.1 Sbe-classes bij de stratum-indeling van de populatie in de metelling van 1985 naar bedrijfstype

Bedrijfstype	Sbe-klasse (totaal aantal sbe)			
	1	2	3	4
Akkerbouw	79 - 161	161 - 255	255 - 420	>= 420
Rundveehouderij	79 - 151	151 - 226	226 - 337	>= 337
Varkenshouderij	79 - 160	160 - 250	250 - 428	>= 428
Pluimveehouderij	79 - 169	169 - 298	298 - 553	>= 553
Akkerbouw/gemengd	79 - 148	148 - 232	232 - 401	>= 401
Rundvee/akkerbouw	79 - 145	145 - 227	227 - 362	>= 362
Rundvee/intensieve veeh.	79 - 156	156 - 232	232 - 355	>= 355
Gemengd/intensieve veeh.	79 - 150	150 - 232	232 - 387	>= 387

Het totaal aantal landbouwbedrijven in de populatie per 1 mei 1985 bedroeg 69.764 bedrijven. De volgende tabel geeft de aantallen voor elk van de 32 bedrijfstype-grootte-classes (basisstrata) weer.

Tabel Bl.2 Aantal landbouwbedrijven van 79 tot 2.000 sbe per bedrijfstype en per sbe-klasse in metelling 1985

Bedrijfstype	Sbe-klasse				
	1	2	3	4	totaal
Akkerbouw	4.174	3.855	2.631	794	11.454
Rundveehouderij	13.281	13.018	9.897	3.593	39.789
Varkenshouderij	2.471	2.328	1.423	407	6.629
Pluimveehouderij	711	573	301	114	1.699
Akkerbouw/gemengd	587	508	311	96	1.502
Rundvee/akkerbouw	517	449	311	121	1.398
Rundvee/intensieve veehouderij	1.101	1.064	791	313	3.269
Gemengd/intensieve veehouderij	1.514	1.384	881	245	4.024

C Oppervlakte:

De volgende vijf classes worden onderscheiden:

- 1) minder dan 5 hectare
- 2) 5 tot 10 hectare
- 3) 10 tot 20 hectare
- 4) 20 tot 50 hectare
- 5) 50 hectare en meer

Bijlage 1 (2e vervolg)

D Leeftijd van het bedrijfshoofd:

Er worden drie leeftijdsklassen onderscheiden; indien er geen leeftijd is opgegeven, zoals bijvoorbeeld het geval is bij BV's, dan wordt een bedrijf ingedeeld bij leeftijdsklasse 2:

- 1) jonger dan 35 jaar
- 2) van 35 tot 60 jaar
- 3) 60 jaar en ouder

E Landbouwgebied:

De gebruikelijke LEI-indeling wordt gevolgd (zie Bedrijfsuitkomsten, 1988: 173):

- 1) Noordelijk zeekleigebied
- 2) Noordelijk veenweidegebied
- 3) Centraal kleigebied
- 4) Westelijk weidegebied
- 5) Zuidwestelijk kleigebied
- 6) De Wouden
- 7) Veenkoloniën
- 8) Noordelijk zandgebied
- 9) Oostelijk zandgebied
- 10) Centraal zandgebied
- 11) Rivierkleigebied
- 12) Zuidelijk zandgebied
- 13) Zuid-Limburg

Bijlage 2 Steekproefpercentages per basisstratum (1987)

Op grond van statistische criteria wordt bepaald hoeveel bedrijven in elk basisstratum in de steekproef moeten vallen (zie paragraaf 6.5). Dat gebeurt op een zodanige wijze dat per bedrijfstype iedere sbe-klasse evenveel steekproefbedrijven bevat. In tabel B2.1 is een overzicht gegeven van de aantallen te kiezen bedrijven per basisstratum en de daarbij - uit tabel B1.2 te berekenen - steekproefpercentages voor boekjaar 1987.

Tabel B2.1 Steekproefaantallen en -percentages per basisstratum in metelling 1985 naar bedrijfstype

Bedrijfstype	Aantal bedrijven per basis stratum	Sbe-klasse			
		1	2	3	4
Akkerbouw	56	1,3%	1,5%	2,1%	7,1%
Rundveehouderij	132	1,0%	1,0%	1,3%	3,7%
Varkenshouderij	32	1,3%	1,4%	2,2%	7,9%
Pluimveehouderij	10	1,4%	1,7%	3,3%	8,8%
Akkerbouw/gemengd	7	1,2%	1,4%	2,3%	7,3%
Rundvee/akkerbouw	6	1,2%	1,3%	1,9%	5,0%
Rundvee/intens. veeh.	14	1,3%	1,3%	1,8%	4,5%
Gemengd/intens. veeh.	18	1,2%	1,3%	2,0%	7,3%

Bijlage 3 De opbouw van het boekhoudnet

Ieder jaar valt een deel van de bedrijven die deelnemen aan het boekhoudnet af, en worden nieuwe bedrijven toegevoegd. De verdeling van de bedrijven die in administratie zijn naar jaar van eerste deelname, is hieronder weergegeven.

Tabel B3.1 Aantal bedrijven in het boekhoudnet naar eerste jaar van deelname van de bedrijven en boekjaar

Boekjaar eerste deelname	In administratie in boekjaar					
	1982	1983	1984	1985	1986	1987
1974	1					
1975	1					
1976	110	5				
1977	169	123				
1978	182	159	112	11		
1979	192	180	161	139	6	2
1980	183	174	167	149	123	5
1981	137	124	114	111	94	75
1982	205	178	166	156	149	122
1983		244	205	188	176	160
1984			242	199	183	165
1985				233	202	179
1986					263	228
1987						279
Totaal	1.180	1.187	1.167	1.186	1.196	1.215

Bijlage 4 Overzicht van standaardfouten voor FIP-grootheden

Tabel B4.1 geeft een aanvulling op tabel 12.4.

Tabel B4.1 Steekproefschattingen voor het gemiddelde en de standaardfout voor een aantal FIP-variabelen voor de gemiddelde nationale FIP-boerderij (gemiddelde van de boekjaren 1982-1986)

Variabele	Gemiddelde waarde	Standaardfout van gemiddelde	Reductie st. fout door stratificatie
*)			
Arbeidsopbrengsten bedrijfshoofd	26.899	1.608 6,0%	10%
Incidentele opbrengsten	3.111	274 8,8%	- 2%
Berekende rente	17.209	197 1,2%	44%
Berekende pacht	24.980	534 2,1%	25%
Afschrijving gebouwen	9.382	227 2,4%	27%
Eigenaarslasten	2.602	63 2,4%	15%
Betaalde rente lang	20.261	581 2,9%	23%
Betaalde rente kort	775	63 8,2%	7%
Verg.b.-vermogen	9.169	551 6,0%	2%
Ondernemersinkomen	39.179	1.312 3,0%	13%
Berekend loon boerin	10.833	447 4,1%	- 7%
Berekend loon overige gezinsleden	17.261	917 5,3%	6%
Opbrengsten bezittingen bij bedrijven	4.189	213 5,1%	1%
Uitkering verzekering	6.178	285 4,6%	- 19%
Overige neveninkomsten	3.087	275 8,9%	- 13%
Totaal gezinsinkomen	80.792	1.574 2,0%	20%
Totaal eigen middelen	44.131	1.690 3,8%	13%
Nieuwe leningen lang	27.144	3.003 11,1%	4%
Mutaties kredieten lang	1.409	907 64,4%	2%
Mutaties vreemd vermogen kort	223	959 430,8%	- 9%
Totaal middelen	72.907	3.898 5,4%	7%
Bruto inventaris in grond	15.193	2.269 14,9%	5%
Bruto inventaris in gebouwen	15.581	1.892 12,1%	- 13%
Pachtersinvesteringen	1.480	410 27,7%	- 1%
Bruto inventaris in dode inventaris	15.778	721 4,6%	9%
Bruto inventaris in vee	- 471	985 209,1%	- 13%
Aflossingen leningen	19.705	1.204 6,1%	2%
Mutaties veldinventaris + voorraden	2.419	741 30,7%	17%
Pachtersinvesteringen	11.887	1.461 12,3%	8%
Dode inventaris	88.462	1.464 1,7%	33%
Veldinventaris/voorraden	24.907	856 3,4%	34%
Kas en giro	2.102	173 8,2%	- 6%
Lopende rekening rabobank	11.135	728 6,5%	- 6%
Lopende rekening overig	1.237	257 20,9%	9%
Voorraad eindproduct	211	58 27,7%	8%
Vorderingen kort	38.195	996 2,6%	21%
Spaarsaldi rabo	36.883	2.840 7,7%	- 4%
Spaarsaldi overig	3.632	746 20,6%	- 5%
Effekten	2.479	569 22,9%	13%
Totale liquiditeit	95.876	3.428 3,6%	5%

Bijlage 4 (vervolg)

Tabel B4.1 (vervolg)

Variabele	Gemiddelde waarde	Standaardfout van gemiddelde	Reductie st. fout door stratificatie
*)			
Ledenrekening coöperaties	10.988	456 4,1%	5%
Leningen u/g	4.370	866 19,8%	- 14%
Totaal vorderingen lang	15.358	1.001 6,5%	- 8%
Woning	79.160	2.263 2,9%	- 3%
Auto	8.528	267 3,1%	4%
Andere bezittingen	18.440	2.131 11,6%	7%
Balanstotaal	1.069.250	14.328 1,3%	32%
Lang vreemd vermogen rabo	189.076	6.600 3,5%	20%
Lang vreemd vermogen overig	19.021	3.242 17,1%	10%
Lang vreemd vermogen verz.m.	859	421 49,1%	- 11%
Lang vreemd vermogen familie	35.085	2.407 6,9%	8%
Lang vreemd vermogen overig	2.124	545 25,7%	- 2%
Lang vreemd vermogen totaal	252.687	6.986 2,8%	23%
Lopende rekening rabo	6.240	627 10,1%	2%
Lopende rekening overig	622	240 38,6%	16%
Leverancierskrediet	8.672	541 6,2%	- 2%
Overige kortlopende schuld	5.971	387 6,5%	7%
Totaal kort vreemd vermogen	21.504	1.018 4,7%	5%

*) Voor omschrijvingen van de variabelen, zie Aukema en Overgauw (1988:9-23).

Bijlage 5 (Sbe)^x als hulpvariabele

Tabel B5.1 Standaarddeviaties van enerzijds (sbe-mei)^x bij de bedrijfskeuze en anderzijds doelvariabelen in de boekhouding (gemiddeld over de boekjaren 1981-1986)

Type/sbe-klasse	Meitell. 2 jaar geleden ----- (sbe-mei) ^x	Boekhouding					
		(sbe-mei) ^x	(sbe-bkh) ^x	sbe-bkh	arb. *) opbr. per ond. *1000	arb. *) opbr. ondern. *1000	fakt. opbr. *1000
1- 1	10	11	15	36	31	32	34
2	11	11	18	48	50	52	57
3	17	17	25	74	65	74	81
4	59	51	57	220	100	120	137
2- 1	9	9	12	29	26	27	22
2	10	10	13	30	34	35	28
3	13	14	17	43	39	45	40
4	35	40	41	142	55	71	91
3- 1	10	10	13	31	28	28	27
2	11	11	14	35	40	41	45
3	17	19	21	58	69	69	71
4	64	37	42	144	106	114	123
4- 1	11	11	20	56	49	50	56
2	14	15	23	67	49	49	56
3	24	25	33	107	62	77	80
4	79	49	63	250	141	169	187
5- 1	9	9	15	37	30	31	28
2	11	9	14	33	43	44	40
3	16	19	26	79	53	56	57
4	47	38	38	128	86	104	129
6- 1	9	11	12	26	32	32	28
2	10	12	13	31	29	30	20
3	14	15	16	41	59	62	58
4	33	27	27	81	50	60	80
7- 1	9	10	11	24	25	26	19
2	10	10	16	41	34	37	34
3	13	11	14	34	42	44	37
4	33	32	35	113	62	72	77
8- 1	9	9	14	35	33	33	30
2	10	11	19	51	37	40	40
3	14	15	20	55	45	47	47
4	51	37	36	118	79	113	108

*) In de FIP staat de arbeidsopbrengst van de ondernemer(s) centraal, in de BUL de arbeidsopbrengst van de ondernemer per ondernemer.

***) 1= akkerbouw; 2=rundvee; 3=varkens; 4=pluimvee; 5= gem. akkerbouw; 6= rundvee/akkerbouw; 7= rundvee/int.; 8= gem.intensief.

Bijlage 6 Stratumherkomst van grotere akkerbouwbedrijven

De BUL-groepen worden gevormd uit delen van strata. Tabel B6.1 geeft bij wijze van illustratie aan hoe strata en BUL-groepen op elkaar aansluiten. De tabel geeft aan hoeveel bedrijven uit de strata 8 tot en met 56 (akkerbouwbedrijven, sbe-klassen 2, 3 en 4) terechtkomen in de vier BUL-groepen met grotere akkerbouwbedrijven en in de overige BUL-groepen. Duidelijk komt naar voren dat bedrijven uit een bepaald stratum in veel gevallen in meerdere BUL-groepen terechtkomen.

Tabel B6.1 Aantal bedrijven (som van de wegingsfactoren) naar stratum van herkomst en BUL-groep (boekjaar 1986)

Stratum nummer (*)	Aantal bedrij- ven	BUL-groep					Totaal
		Nrd. Klei	Veen- kol.	Centr. Klei	Zuidw. Klei	Overige groepen	
8	2	89				88	177
10	1			225			225
12	11		165			198	363
13	8		97			161	258
14	3			92		184	276
15	5				117	176	293
16	5		48			191	239
17	1			161			161
18	4				219		219
19	5			186			186
20	2	162					162
22	12			634		57	691
23	11				705		705
24	9		209			27	236
25	4		226				226
26	3	61	61			61	182
27	3			202			202
28	3				149	74	223
29	3	105	52				157
30	5		191	48			239
.	.						.
.	.						.
55	1			30			30
56	4				25	25	50
Totaal	**)	983	1.614	2.479	2.288	1.564	8.928

*) Voor de indeling in strata, zie van Dijk (1986).

***) Zie Bedrijfsuitkomsten (1988), staat A1 en A2.

Bijlage 7 Keuze-aantallen en steekproefpercentages bij de steekproefopzet voor boekjaar 1988

Tabel B7.1 Vergelijking aantallen in administratie boekjaar 1987 en keuze-aantallen boekjaar 1988

Stratificatie- groep	Sbe- klasse	Aantal bedr. in admini- stratie boekjaar 87	Keuzeaan- tal boek- jaar 1988	Steekproef- percentage
Grotere akkerbouwbedrijven				
- regio 1	2	17	20	3,38%
	3	19	20	4,71%
	4	22	20	14,71%
- regio 3,4	2	21	20	1,82%
	3	21	20	2,54%
	4	16	20	8,51%
- regio 5	2	18	20	1,72%
	3	16	20	2,64%
	4	19	20	9,43%
- regio 6-8	2	20	15	1,99%
	3	10	15	3,43%
	4	6	15	12,20%
- regio 2,9-13	2	1	4	1,34%
	3	2	4	2,41%
	4	5	4	8,51%
Kleinere akkerbouwbedrijven				
- regio 1,3-5	1	35	35	1,71%
- regio 3,4	1	30	30	2,97%
- regio 2,9-13	1	6	4	0,78%
Grotere sterk gespec. melkveebedrijven				
- regio 1-3	2	30	25	1,01%
	3	36	25	1,40%
	4	30	25	3,45%
- regio 4,5,11	2	22	25	1,23%
	3	22	25	1,69%
	4	20	25	4,55%
- regio 6-10,12,13	2	76	57	0,90%
	3	69	57	1,21%
	4	57	57	3,55%

Bijlage 7 (1e vervolg)

Tabel B7.1 (1e vervolg)

Stratificatie- groep	Sbe- klasse	Aantal bedr. in admini- stratie boekjaar 87	Keuzeaan- tal boek- jaar 1988	Steekproef- percentage

Kleinere sterk gespec. melkveebedr.		*)		
- regio 1-5,11	1	49	50 (41)	1,05% (0,88%)
- regio 6-10,12,13	1	76	85 (59)	1,26% (0,88%)

Minder sterk gespec. melkveebedr.				
- kleinere	1	19	35	1,73%
- grotere	2	15	20	1,21%
	3	14	20	1,67%
	4	15	20	5,60%

Fokvarkensbedrijven				
	1	13	16	1,34%
	2	22	16	1,33%
	3	17	16	2,46%
	4	15	16	8,65%

Andere varkens- bedrijven				
	1	11	17	1,33%
	2	15	17	1,41%
	3	21	17	2,21%
	4	9	17	7,36%

Legkippenbedrijven				
	1	5	10	2,01%
	2	10	10	2,42%
	3	15	10	4,29%
	4	4	10	12,66%

Overige hokdierenbedrijven				
	1	1	3	0,84%
	2	3	3	1,06%
	3	3	3	1,78%
	4	3	3	6,25%

Andere graasdierbedrijven				
	1	12	10	0,65%
	2	9	6	0,84%
	3	4	6	1,45%
	4	6	6	4,26%

Gewassen- combinaties				
	1	5	4	0,61%
	2	2	4	0,78%
	3	3	4	1,16%
	4	2	4	3,70%

Bijlage 7 (2e vervolg)

Tabel B7.1 (2e vervolg)

Stratificatie- groep	Sbe- klasse	Aantal bedr. in admini- stratie boekjaar 87	Keuzeaan- tal boek- jaar 1988	Steekproef- percentage
Veeteelt- combinaties	1	19	15	0,88%
	2	25	10	0,60%
	3	23	10	0,89%
	4	29	10	2,40%
Gewassen/veeteelt- combinaties	1	13	8	0,68%
	2	16	6	0,63%
	3	14	6	1,02%
	4	10	6	3,39%
Buiten de populatie		23		
Alle bedrijven		1.216	1.156	1,64%

*) Tussen haakjes de uiteindelijk gekozen aantallen in verband met de inkrimping van het boekhoudnet (zie paragraaf 20.4).

Bijlage 8 Vergelijking BUL-wensen voor en na boekjaar 1988

Tabel B8.1 geeft in kolom 2 de gewenste aantallen bedrijven in de verschillende BUL-en FIP-groepen (op NEG-basis) zoals die voor boekjaar 1988 zijn geformuleerd (zie tabel 20.3). De kolommen 3 en 4 bevatten de huidige uitgewerkte aantallen boekhoudingen in de betreffende groepen. Deze aantallen zijn berekend op basis van de kenmerken van de boekhoudnetbedrijven in de metellingen voorafgaande aan de boekjaren. Deze aantallen zullen niet veel afwijken van de aantallen die gebaseerd zijn op kenmerken gemiddeld gedurende het boekjaar. Temeer daar de indeling klein/groot wel gebaseerd is op boekhoudnetgegevens.

Tabel B8.1 Gewenste en huidige aantallen steekproefbedrijven in de nieuwe BUL-en FIP-groepen (ter vergelijking zijn in de laatste kolom de metelingsaantallen die horen bij de betreffende BUL-groep opgenomen)

BUL/FIP-groep *)	BUL-wensen 1988	BUL **) 1984-86	FIP 1984-86	metel- ling 86
Akkerbouwbedrijven	210	255	190	10.814
Grotere bedrijven	210	207	148	7.237
- noordelijk klei	55	48	24	1.155
- veenkolonien	40	39	33	1.316
- centraal klei	55	57	46	2.125
- zuidwestelijk klei	55	55	41	2.130
- overige gebieden		8	5	511
Kleinere bedrijven	55	48	41	3.577
- kleigebieden	30	24	19	2.050
- veenkolonien	25	21	20	1.011
- overige gebieden		3	2	516
Sterk gespecialiseerde melkveebedrijven	285	444	382	33.221
Grotere bedrijven	280	331	274	21.683
- nrd. klei/veenweide	65	91	71	5.001
- westelijk weide	65	60	48	4.064
- zandgebieden	150	181	154	12.618
Kleinere bedrijven	85	113	109	11.538
- klei/veenweide	45	42	41	4.776
- zandgebieden	80	70	67	6.762
Minder gespecialiseerde melkveebedrijven	70	57	53	5.230
Grotere bedrijven	40	39	37	3.207
Kleinere bedrijven	30	18	17	2.023

Bijlage 8 (vervolg)

Tabel B8.1 (vervolg)

BUL/FIP-groep *)	BUL-wensen 1986	BUL **) 1984-86	FIP 1984-86	meitel- ling 86
Hokdieren-bedrijven	115	105	90	7.944
Fokvarkensbedrijven	60	46	37	3.237
Andere varkensbedrijven	40	34	32	3.485
Legkippenbedrijven	35	25	21	1.222
Overige bedrijven	80	147	131	13.097
Grotere bedrijven	60	106	90	7.658
Kleinere bedrijven	35	41	41	5.439
		1.008	846	70.306

*) Bedrijfstype op basis van meetingsgegevens; bedrijfsgrootte op basis van gemiddeld aantal sbe gedurende het boekjaar.

***) Gemiddelde aantallen in de boekjaren 1984-1986.

Bijlage 9 Standaardfouten voor deelpopulaties

In deze bijlage wordt ingegaan op het berekenen van standaardfouten (varianties) van schattingen voor delen van de populatie die dwars door de strata heenlopen. Voorbeelden van dergelijke populaties zijn de BUL- en FIP-groepen.

De inhoud van deze bijlage is in hoofdlijnen gebaseerd op Moors en Muilwijk (1975), paragraaf 3.8 en 4.9.

De standaardfouten voor deelpopulaties in een *enkelvoudige* steekproef hangen af van het al dan niet bekend zijn van de omvang (A) van de deelpopulatie. Het subgemiddelde X_d en het subtotaal X_d van een deelpopulatie worden gedefinieerd als:

$$X_d = \sum_{i=1}^A x_i \quad (B9.1)$$

$$\bar{X}_d = \frac{1}{A} * \sum_{i=1}^A x_i \quad (B9.2)$$

waarbij de index d slaat op "deelpopulatie" en de elementen van de populatie geacht worden zo genummerd te zijn dat de eerste A elementen de deelpopulatie vormen.

De grootheden X_d en \bar{X}_d kunnen uit de steekproef zuiver geschat worden als:

$$\hat{X}_d = \frac{A}{a} * \sum_{j=1}^a x_j \quad (B9.3)$$

$$\hat{\bar{X}}_d = \frac{1}{a} * \sum_{j=1}^a x_j \quad (B9.4)$$

waarbij:

a = het aantal elementen in de steekproef dat tot de deelpopulatie behoort.

De variantie S^2_d in de populatie wordt gedefinieerd als:

$$S^2_d = \frac{1}{A-1} * \sum_{i=1}^A (x_i - \bar{X}_d)^2 \quad (B9.5)$$

Zuivere schatting hiervan kan plaatsvinden met de schatter:

$$s^2_d = \frac{1}{a-1} * \sum_{i=1}^a (x_i - \bar{X}_d)^2 \quad (B9.6)$$

De variantie van \bar{X}_d is:

$$\text{var}(\bar{X}_d) = \frac{s^2_d}{n} * \left(1 - \frac{a}{A}\right) \quad (B9.7)$$

waarbij (a/A) vervangen wordt door (n/N) indien A onbekend is.

Indien A bekend is kan ook het subtotaal X_d op de gebruikelijke wijze geschat worden:

Bijlage 9 (1e vervolg)

$$\hat{x}_d = \frac{A}{a} * \sum_{i=1}^a x_i \quad (B9.8)$$

en

$$\text{var}(\hat{x}_d) = A^2 * \text{var}(\hat{x}_d) \quad (B9.9)$$

Indien A onbekend is, kan echter geen gebruik gemaakt worden van de bovenstaande formule. A zal dan geschat moeten worden uit de steekproef. De onbekendheid van A zal zorgen voor een vergroting van de variantie. A kan geschat worden met:

$$\hat{A} = \left(\frac{N}{n} \right) * a \quad (B9.10)$$

De variantie van \hat{A} is gelijk aan: *

$$\text{var}(\hat{A}) = \frac{N^2}{n*(n-1)} * \left(1 - \frac{n}{N}\right) * \left(a - \frac{a^2}{n}\right) \quad (B9.11)$$

De variantie van \hat{x}_d is gelijk aan:

$$\text{var}(\hat{x}_d) = \frac{N^2 * s^2}{n} * \left(1 - \frac{n}{N}\right) \quad (B9.12)$$

waarbij:

$$s^2 = \frac{1}{n-1} * \sum_{i=1}^n (x_i' - \bar{x}')^2 \quad (B9.13)$$

waarbij:

$$\begin{aligned} x_i' &= x_i && \text{voor alle elementen in de deelpopulatie} \\ x_i' &= 0 && \text{voor alle elementen buiten de deelpopulatie} \end{aligned}$$

en

$$\bar{x}' = \frac{1}{n} * \sum_{i=1}^a x_i' \quad (B9.14)$$

Voor de standaardfout van het gemiddelde \bar{x}_d maakt het dus niet uit of A al dan niet bekend is, voor \hat{x}_d wel.

* Het geschatte percentage bedrijven ($\hat{\rho}$) dat tot een deelpopulatie behoort, is (\hat{A}/N). De variantie (standaardfout) van $\hat{\rho}$ is gelijk aan de variantie (standaardfout) van \hat{A} gedeeld door N^2 (N).

Bijlage 9 (2e vervolg)

Bij een *gestratificeerde* steekproef moet voor het bepalen van de standaardfouten onderscheid gemaakt worden tussen de volgende drie gevallen (Moors en Mulwijk, 1975:65):

- I - In ieder stratum is de omvang van de deelpopulatie bekend;
- II - Alleen de omvang van de gehele deelpopulatie is bekend;
- III - Over de omvang van de deelpopulatie is niets bekend.

I - In ieder stratum is de omvang van de deelpopulatie bekend.

Analoog aan het voorgaande en de definities in hoofdstuk 5 definiëren we:

- A_h = aantal elementen in de populatie dat tot de deelpopulatie behoort (in stratum h)
- a_h = aantal elementen in de steekproef dat tot de deelpopulatie behoort (in stratum h)
- \bar{x}_{dh} = het gemiddelde in de deelpopulatie (in stratum h)
- \hat{x}_{dh} = het gemiddelde in de deelpopulatie in de steekproef (in stratum h)
- S^2_{dh} = de variantie in de deelpopulatie (in stratum h)
- s^2_{dh} = de geschatte variantie in de deelpopulatie (in stratum h)

Dan zijn \bar{x}_{dh} en s^2_{dh} zuivere schatters voor \bar{x}_{dh} en S^2_{dh} als:

$$\bar{x}_{dh} = \frac{1}{a_h} * \sum_{i=1}^{a_h} x_{hi} \quad (B9.15)$$

$$s^2_{dh} = \frac{1}{a_h-1} * \sum_{i=1}^{a_h} (x_{hi} - \bar{x}_{dh})^2 \quad (B9.16)$$

Het geschatte totaal \hat{x}_d van de gehele deelpopulatie wordt dan:

$$\hat{x}_d = \sum_{h=1}^H \frac{A_h}{a_h} * \sum_{i=1}^{a_h} x_{hi} \quad (B9.17)$$

en

$$\bar{x}_d = \frac{\hat{x}_d}{A} \quad (B9.18)$$

De varianties die daarbij horen zijn:

$$\text{var}(\hat{x}_d) = \sum_{h=1}^H \frac{A^2_h * s^2_{dh}}{a_h} * \left(1 - \frac{a_h}{A_h}\right) \quad (B9.19)$$

en

Bijlage 9 (3e vervolg)

$$\text{var}(\bar{x}_d) = \frac{1}{A^2} * \sum_{h=1}^H \frac{A^2 h * s^2 dh}{a_h} * (1 - \frac{a_h}{A_h}) \quad (B9.20)$$

II *Alleen de omvang van de gehele deelpopulatie (A) is bekend*

Omdat A_h onbekend is, kunnen bovenstaande formules niet zondermeer gebruikt worden. In plaats van (A_h/a_h) zal de geschatte waarde (N_h/n_h) ingevuld moeten worden in (B9.17). Immers, $\hat{A}_h = N_h * a_h / n_h$. Analooq aan (B9.13) geldt voor de geschatte variantie binnen de deelpopulatie van een stratum:

$$s^2_{h'} = \frac{1}{n_h - 1} * \sum_{i=1}^{n_h} (x_{hi}' - \bar{x}_h')^2 \quad (B9.21)$$

waarbij:

$$\begin{aligned} x_{hi}' &= x_{hi} \text{ voor alle elementen in de deelpopulatie} \\ x_{hi}' &= 0 \text{ voor alle elementen buiten de deelpopulatie} \end{aligned}$$

en

$$\bar{x}_h' = \frac{1}{n_h} * \sum_{i=1}^{n_h} x_{hi}' \quad (B9.22)$$

Voor de variantie geldt:

$$\text{var}(\hat{x}_d) = \sum_{h=1}^H \frac{N^2 h * s^2_{h'}}{n_h} * (1 - \frac{n_h}{N_h}) \quad (B9.23)$$

en

$$\text{var}(\bar{x}_d) = \frac{1}{A^2} * \sum_{h=1}^H \frac{N^2 h * s^2_{h'}}{n_h} * (1 - \frac{n_h}{N_h}) \quad (B9.24)$$

III - *Over de omvang van de deelpopulatie is niets bekend.*

Voor het schatten van totalen (\hat{x}_d) van deelpopulaties kunnen de formules onder II gebruikt worden. Voor het schatten van gemiddelden ($\bar{x}_d = \hat{x}_d / \hat{A}$) moet A geschat worden.

$$\hat{A} = \sum_{h=1}^H \left(\frac{N_h * a_h}{n_h} \right) \quad (B9.25)$$

$$\bar{x}_d = \frac{1}{\hat{A}} * \hat{x}_d \quad (B9.26)$$

Bijlage 9 (4e vervolg)

De gekwadrateerde standaardfout is gelijk aan:

$$\text{var}(\bar{x}_d) = \frac{1}{\hat{A}^2} * \sum_{h=1}^H \frac{N^2 h * s^2_{qh}}{n_h} * \left(1 - \frac{n_h}{N_h}\right) \quad (\text{B9.27})$$

waarbij:

$$s^2_{qh} = \frac{1}{n_h - 1} * \sum_{i=1}^{a_h} (x_{hi} - \bar{x}_{dh})^2 + \frac{a_h}{n_h - 1} * \left(1 - \frac{a_h}{n_h}\right) * (\bar{x}_{dh} - \bar{x}_d)^2 \quad (\text{B9.28})$$

Deze formule komt niet overeen met de betreffende formule in Moors en Muilwijk (1975). De noemer in de eerste term is $(n_h - 1)$ in plaats van $(a_h - 1)$. Formule (B9.28) is te vinden in Cochran (1977:144) en in Kish and Frankel (1974).

Analoog aan (B9.11) kan de variantie van \hat{A} berekend worden uit:

$$\text{var}(\hat{A}) = \sum_{h=1}^H \frac{N^2 h}{n_h * (n_h - 1)} * \left(1 - \frac{n_h}{N_h}\right) * \left(a_h - \frac{a^2 h}{n_h}\right) \quad (\text{B9.29})$$

Herschrijving van de formules (B9.27) en (B9.23) vergemakkelijkt een vergelijking van de relatieve standaardfouten van het gemiddelde en het totaal bij onbekende A:

$$\frac{\text{var}(\bar{x}_d)}{(\bar{x}_d)^2} \approx \text{factor} * \sum_{i=1}^{a_h} (x_{hi}' - \bar{x}_h')^2 + a_h * \left(1 - \frac{a_h}{n_h}\right) * (\bar{x}_h' - \bar{x}')^2 \quad (\text{B9.30})$$

en

$$\frac{\text{var}(\hat{x}_d)}{(\hat{x}_d)^2} = \text{factor} * \sum_{i=1}^{a_h} (x_{hi}' - \bar{x}_h')^2 + a_h * \left(1 - \frac{a_h}{n_h}\right) * (x_{h1}')^2 \quad (\text{B9.31})$$

waarbij

$$\text{factor} = \sum_{h=1}^H \frac{N_h * (N_h - n_h)}{n_h * (n_h - 1)}$$

In het algemeen zal de relatieve standaardfout van het totaal groter zijn dan die van het gemiddelde. Maar uit bovenstaande formules wordt duidelijk dat het omgekeerde ook het geval kan zijn. Dat vloeit dan echter voort uit het feit dat (B9.30) een benaderingsformule is en (B9.31) een exacte formule.

Vergelijking van (B9.27) met (B9.20) leert dat berekening van de standaardfout met zekere wegingsfactoren (B9.20) kan leiden tot een grotere variantie dan berekening met onzekere factoren (B9.27). Ook daarbij speelt een rol dat (B9.27) slechts een benadering is.