



Tweede Review nationale studenten enquête 2019

Marja Exterkate

Lotte Schoot

Bram Rouw

CBS Den Haag
Henri Faasdreef 312
2492 JP Den Haag
Postbus 24500
2490 HA Den Haag
+31 70 337 38 00
www.cbs.nl

projectnummer 190233

15 juli 2019

Inhoudsopgave

1. Inleiding	4
1.1 Onderzoeksvragen	4
2. Effect van reparatie	5
2.1 Uitgevoerde reparaties	5
2.2 Effect van reparatie op responspercentages	6
2.3 Effect van reparatie op themascores	7
3. Respons na reparatie	9
3.1 Respons naar achtergrondkenmerken en instelling	9
3.2 Verdeling responspercentages	10
3.3 Representativiteit van de respons	13
4. Conclusies en aanbevelingen	15

1. Inleiding

Uit de review die het CBS in mei/juni 2019 heeft uitgevoerd voor Studiekeuze123 op de bestanden van de Nationale Studenten Enquête (NSE) 2019¹, blijkt dat de respons van de huidige NSE onvoldoende betrouwbaar is om tevredenheidsscores per opleiding te kunnen publiceren. Dit komt voornamelijk doordat niet met zekerheid kan worden vastgesteld dat studenten de opleiding volgen die zij hebben ingevuld in de enquête. Ook is van een deel van de respondenten op dit moment onbekend welke opleiding zij hebben beoordeeld omdat zij gebruik hebben gemaakt van de optie 'anders, namelijk...'.

Daarom is een reparatie uitgevoerd op het benchmarkbestand met behulp van de door onderwijsinstellingen aangeleverde populatiegegevens. Deze reparatie is op dit moment alleen uitgevoerd voor 17 van de 18 instellingen uit het niet-bekostigd onderwijs. Door de in de enquête ingevulde gegevens te verrijken met de na geleverde inschrijfgegevens, kan worden nagegaan of de beoordeelde opleiding overeenkomt met de opleiding waar de respondent staat ingeschreven. Vervolgens kan de respons worden toegeschreven aan de opleiding waar de student daadwerkelijk staat ingeschreven, zoals dat ook in voorgaande edities van de NSE is gedaan.

Studiekeuze123 heeft het CBS gevraagd om op het gerepareerde bestand een tweede review te doen. Het doel van de review is dat het CBS Studiekeuze123 antwoord geeft op de vraag of de uitkomsten van de NSE 2019 voor de 17 gerepareerde instellingen in het niet-bekostigd onderwijs voldoende betrouwbaar zijn om te publiceren.

1.1 Onderzoeksvragen

Na het verrijken van de responsdata met inschrijfgegevens, is Studiekeuze123 geïnteresseerd in de vraag in hoeverre de resultaten van de NSE2019 voor het niet-bekostigd onderwijs publicabel zijn.

Om een antwoord te kunnen geven op deze vraag, worden in de hoofdstukken 2 en 3 de volgende deelvragen behandeld:

- 1) In hoeverre verschilt de administratieve werkelijkheid van de door de student ingevulde gegevens, en wat is het effect van deze reparatie op responspercentages en themascores?
- 2) In hoeverre is de herstelde respons in 2019 vergelijkbaar met die uit eerdere jaren wat betreft kwantiteit (responspercentage) en kwaliteit (representativiteit)?

In de conclusie (hoofdstuk 4) wordt vervolgens de belangrijkste vraag beantwoord:

- 3) Zijn de herstelde NSE2019 data publicabel, en zo ja, op welk niveau kunnen de resultaten voor het niet-bekostigd onderwijs worden gepubliceerd?

¹ <https://www.cbs.nl/nl-nl/maatwerk/2019/25/review-nationale-studenten-enquete-2019>

2. Effect van reparatie

In dit hoofdstuk wordt onderzoeksvraag 1) beantwoord: *In hoeverre verschilt de administratieve werkelijkheid van de door de student ingevulde gegevens, en wat is het effect van deze reparatie op responspercentages en themascores?*

2.1 Uitgevoerde reparaties

Het CBS heeft van Studiekeuze123 een nieuw benchmarkbestand gekregen, waarin reparaties zijn uitgevoerd. Dat wil zeggen, de informatie over de instelling en opleiding per respondent komt nu overeen met de instelling en opleiding waarvoor de respondent staat ingeschreven volgens de onderwijsinstellingen. Hetzelfde geldt voor de opleidingslocatie, opleidingsvorm en type student. Dit gerepareerde benchmarkbestand is op emailadres van de respondenten gekoppeld aan het responsbestand voor herstel (de ruwe data). Door deze twee bestanden aan elkaar te koppelen kunnen we voor iedere respondent een vergelijking maken tussen de instelling en opleidingskenmerken die hij/zij zelf heeft ingevuld in de enquête, en de instelling en opleiding waarvoor hij/zij staat ingeschreven volgens de onderwijsinstelling.

In totaal wijken voor 1103 van de 3409 respondenten (32%) de ingevulde gegevens af van de inschrijfgegevens. Er is één instelling waarbij alle respondenten de juiste opleidingskenmerken hebben ingevuld. Voor de overige instellingen komt bij tenminste één respondent de instelling, opleiding, locatie, studievorm of het studietype niet overeen met de inschrijfgegevens.

2.1.1 Herstelpercentage en aantal instellingen per reparatietype

Kenmerk	Percentage hersteld	Aantal instellingen met ten minste één reparatie van dit type (van 17)
Instelling	3%	8
Opleiding	12,6%	14
Locatie	9,5%	10
Studievorm	10,2%	13
Type student	9,8%	12

Instelling

Alle respondenten waarvoor de instelling in het nieuwe bestand afwijkt van de ingevulde informatie hadden wel een instelling ingevuld. Het gaat hier dus om verschuivingen van respondenten tussen instellingen. Er zijn geen respondenten die niets of 'anders, namelijk...' hadden ingevuld en die nu wel ingedeeld kunnen worden bij de juiste instelling. Voor 14,4% van de respondenten waarvoor een reparatie is uitgevoerd op instelling, is wel de juiste opleiding, locatie, opleidingsvorm en type student ingevuld. Deze respondenten vallen allemaal binnen één 'joint degree'² opleiding. In alle andere gevallen dat de ingevulde instelling afwijkt van de administratieve gegevens, wijkt de respons ook af op één van de andere vlakken (opleiding, locatie, opleidingsvorm of type).

Opleiding

Op het niveau van opleiding bestaat bijna de helft (48,9%) van de reparaties uit respondenten die 'Anders, namelijk...' hebben ingevuld, waarvoor nu uit de inschrijvingsgegevens een opleiding bekend is. De overige 51,1% van de respondenten waarvoor een reparatie is uitgevoerd, heeft een

² Opleiding die wordt aangeboden door een samenwerking van instellingen

opleiding ingevuld die niet overeenkomt met de opleiding waarvoor zij ingeschreven staan volgens de onderwijsinstelling. Er zijn drie instellingen waarbij alle reparaties op opleidingsniveau van het type 'Anders, namelijk...' zijn. Voor de overige instellingen is gemiddeld 50% van de reparaties van het type 'Anders, namelijk...'.

Locatie

Op het niveau van locatie heeft de meerderheid van de reparaties (77,2%) plaatsgevonden voor respondenten die 'Anders, namelijk...' hebben ingevuld. Voor 30% van de respondenten waarvoor de locatie afwijkt van administratieve gegevens, wijkt ook de opleiding en/of instelling af. Voor drie instellingen komen locatieverschillen voor bij alle combinaties van instelling, opleiding, locatie, type student en opleidingsvorm.

Studievorm

De meerderheid van de respondenten waarvoor een reparatie is uitgevoerd op studievorm (73,8 %) heeft 'duaal' ingevuld als opleidingsvorm en is nu ingedeeld bij deeltijd of voltijd. In de overige gevallen gaat het om een verschuiving van voltijd naar deeltijd of andersom. Van de respondenten voor wie studievorm is gerepareerd heeft 70% wel de juiste instelling, opleiding en locatie ingevuld.

Type student

Het grootste deel van de reparaties op type student (76,6%) betreft een verschuiving van pre- of postmaster naar regulier. Een klein deel (9,6%) van de respondenten voor wie type student is gerepareerd had niets ingevuld. Overige reparaties betreffen verschuivingen van pre- naar postmaster of van regulier naar postmaster. Bijna driekwart (74,6%) van de respondenten voor wie het type student is gerepareerd, heeft wel de juiste instelling, opleiding en locatie ingevuld.

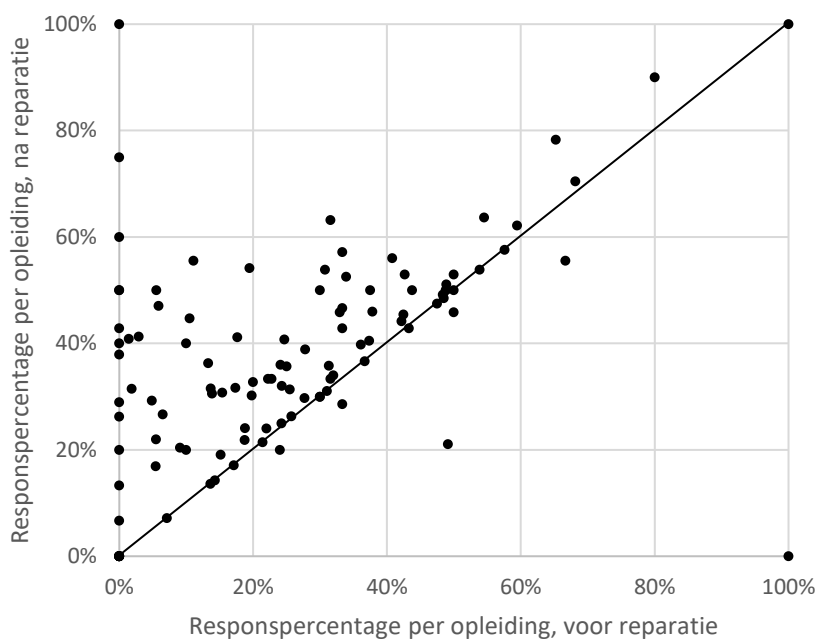
2.2 Effect van reparatie op responspercentages

Uit de vorige paragraaf komt naar voren dat reparaties twee effecten kunnen hebben. Enerzijds zijn voor respondenten die niets of 'Anders namelijk...' hadden ingevuld nu de juiste kenmerken bekend. Anderzijds zijn voor studenten die de instelling of opleidingsgegevens (inclusief locatie, type student, en opleidingsvorm) verkeerd hadden ingevuld, nu de juiste gegevens ingevuld.

Reparaties met het eerste effect zorgen voor een toename van het responspercentage voor de opleiding waar de student na reparatie is ondergebracht, terwijl het laatste type zorgt voor een verschuiving van respondenten tussen opleidingen. In figuur 2.2.1 hieronder is het verschil in responspercentage voor en na reparatie zichtbaar. Iedere waarneming is een combinatie van instelling, opleiding, locatie, opleidingsvorm en type student. Om een vergelijking mogelijk te maken zijn de populatieaantallen gebruikt uit het op persoonsniveau aangeleverde populatiebestand. Uit deze figuur is één opleiding weggelaten waarbij het aantal respondenten voor herstel hoger was dan het aantal ingeschreven studenten.

Gemiddeld is de respons per opleiding met 12,2 procentpunten toegenomen. Deze toename wordt waarschijnlijk voor een groot deel gedreven door de 'anders, namelijk...' respondenten die nu toegeschreven kunnen worden aan de juiste opleiding.

2.2.1 Responspercentage per opleiding, voor en na reparatie



2.3 Effect van reparatie op themascores

De studentenoordelen op de verschillende thema's zijn een belangrijke uitkomst van de NSE waarop instellingen en opleidingen vergeleken worden. Om te onderzoeken wat het effect is van herstel van de data op deze themascores, worden de door Studiekeuze123 berekende scores uit het benchmarkbestand na herstel vergeleken met de scores uit het benchmarkbestand voor herstel.

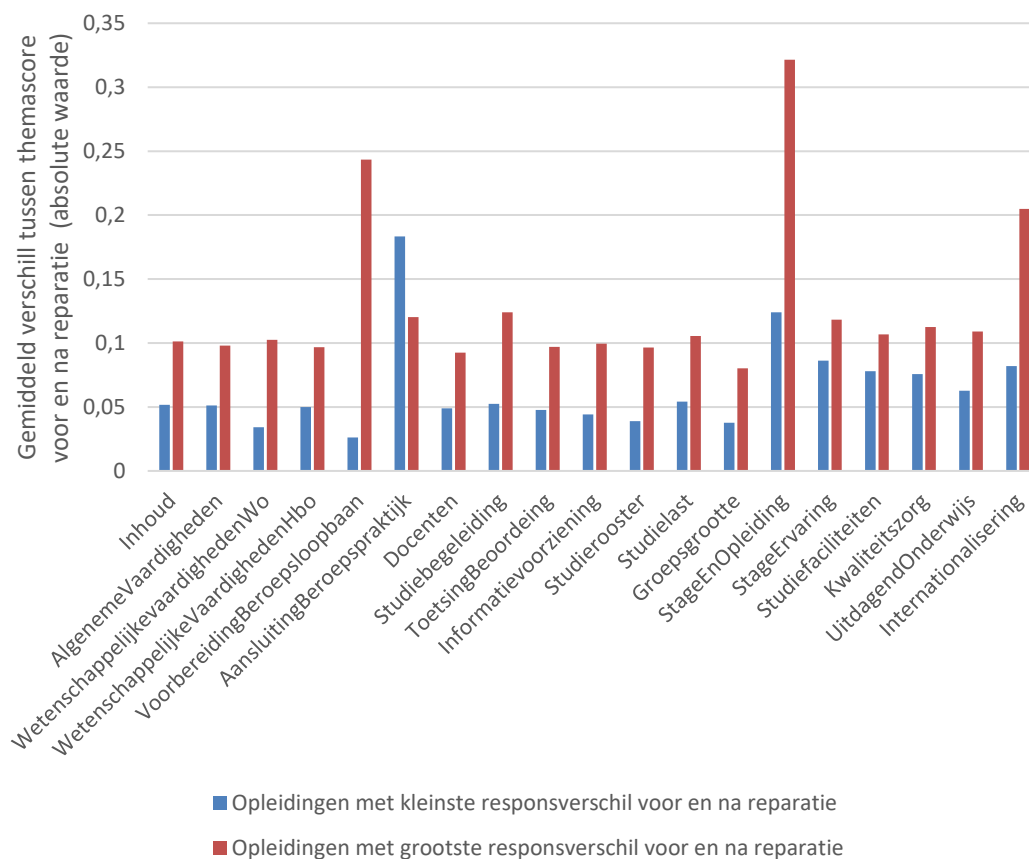
Deze analyse is uitgevoerd op het niveau van instelling, opleiding, locatie, type student en opleidingsvorm. Voor ieder thema is een gemiddelde score berekend per opleiding (ongewogen), met de restrictie dat het aantal respondenten per opleiding minimaal vijf is. Vervolgens zijn de (absolute) verschillen tussen de gemiddelden voor en na herstel berekend. Verschillen worden alleen gepresenteerd voor opleidingen die zowel voor als na reparatie voorkomen.

De verwachting is dat het verschil tussen themascores voor en na reparatie het grootst is voor opleidingen waarbij herstel veel effect heeft gehad. Dit wordt gemeten aan de hand van het verschil in responspercentage voor en na reparatie. Opleidingen zijn ingedeeld in twee groepen op basis van de mediaan in absolute responsverschillen: één groep waarbij het verschil in respons voor en na reparatie onder de mediaan ligt en één groep waarbij het verschil boven de mediaan ligt. Voor de groep met de kleinste verschillen in responspercentages is het gemiddelde verschil in respons voor en na reparatie 1,9% ($n = 34$, standaarddeviatie 1,6%) en voor de groep met de grootste verschillen in responspercentages is het gemiddelde verschil in respons 15,3% ($n = 35$, standaarddeviatie 7,8%).

Het gemiddelde verschil tussen themascores (absolute waarden) per groep wordt hieronder weergegeven in figuur 2.3.1. Hierbij moet worden opgemerkt dat niet alle vragen voor alle opleidingen zijn beantwoord. Hierdoor wijkt het aantal opleidingen per groep waarover het gemiddelde is berekend, af voor een aantal themascores. Dit geldt voor thema's "Wetenschappelijke vaardigheden Wo" (groep met kleinste verschillen in responspercentages: 11 opleidingen; groep met grootste verschillen in responspercentages: 20 opleidingen), "Wetenschappelijke vaardigheden Hbo"

(23/15), “Vorbereiding Beroepsloopbaan” (28/14), “Aansluiting beroepspraktijk” (5/21), “Stage en opleiding” (28/21), “Stage ervaring” (28/20) en “Internationalisering” (34/33). Het is zichtbaar dat opleidingen met een groot verschil in responspercentage voor en na reparatie een groter verschil laten zien in de themascores dan opleidingen met een klein verschil in responspercentage. Het gemiddelde verschil van alle themascores is respectievelijk 0,07 punten en 0,12 punten.

2.3.1 Gemiddeld verschil tussen themascores voor en na reparatie



3. Respons na reparatie

Na reparatie van de respons voor de niet-bekostigde instellingen met populatiegegevens, kan opnieuw worden beoordeeld hoe de kwaliteit van deze respons zich verhoudt tot eerdere jaren. De deelvraag die in dit hoofdstuk wordt beantwoord is: *In hoeverre is de herstelde respons in 2019 vergelijkbaar met die uit eerdere jaren wat betreft kwantiteit (responspercentage) en kwaliteit (representativiteit)?*

Populatiegegevens zijn aangeleverd op het niveau van instelling, opleiding, locatie, opleidingsvorm en type student. Omdat dit ook het niveau is waarop tevredenheidsscores in voorgaande jaren gepubliceerd zijn, zullen de volgende analyses zich richten op dit niveau. In vergelijking met de eerste review zal de totale respons niet zijn veranderd, maar is wel de samenstelling van de respons veranderd doordat ingevulde opleidingsgegevens zijn overschreven met inschrijfggegevens. Daarnaast is nu ook voor respondenten die handmatig hun opleiding hebben ingevuld via 'anders, namelijk...' bekend voor welke opleiding zij ingeschreven staan. Dit kan op het niveau van instelling, locatie, opleiding, type student en soort onderwijs wel leiden tot een hogere respons. Met de herstelde responsdata kan nu een zuiverdere vergelijking worden gemaakt met de respons van voorgaande jaren op dit niveau.

3.1 Respons naar achtergrondkenmerken en instelling

Tabel 3.1.1 toont de (netto) respons en populatieaantallen uitgesplitst naar onderwijsvorm en type student. Netto heeft betrekking op de correctie voor dubbeltellingen.

3.1.1 Netto respons naar inschrijfkenmerken, NSE 2018 en 2019

	2018			2019		
	Populatie	Respons	Respons %	Populatie	Respons	Respons %
Totaal	11 739	3 914	33,3	10 708	3 409	31,8
<i>Onderwijsvorm</i>						
Deeltijd	5 794	1 593	27,5	4 424	1 259	28,5
Voltijd	5 945	2 321	39,0	6 284	2 150	34,2
<i>Type student</i>						
Postmaster	362	95	26,2	313	78	24,9
Premaster	194	61	31,4	118	42	35,6
Regulier	11 183	3 758	33,6	10 277	3 289	32,0

Tabel 3.1.2 laat zien dat zowel in 2018 als in 2019 het grootste deel van de instellingen een responspercentage heeft tussen de 30% en 40%. In 2019 zijn er twee instellingen met een responspercentage van 10% tot 20%. In 2018 komt dit niet voor.

3.1.2 Responspercentages per instelling, NSE 2018 en 2019

Respons	2017	2018	2019	
10 tot 20%		1	-	2
20 tot 30%		3	3	3
30 tot 40%		5	8	8
40 tot 50%		6	3	3
> 50%		2	3	1

3.2 Verdeling responspercentages

Om inzicht te krijgen in hoe de respons zich na reparatie verhoudt tot eerdere jaren, wordt op het niveau van instelling, locatie, opleiding, type student en opleidingsvorm onderzocht wat de responsafwijking is ten opzichte van voorgaande jaren. Combinaties van deze uitsplitsing waarbij minder dan vijf personen hebben gerepsondeerd, zijn in deze analyse uitgesloten.

Tabel 3.2.1 laat de totale respons en populatie zien en daarnaast het percentage van de respons dat niet is meegenomen vanwege een totaal van minder dan vijf respondenten. De totale respons in 2019 is 31,8% (1,5 procentpunt lager dan in 2018). Hiervan wordt nog 0,9 procent uitgesloten omdat de opleiding minder dan vijf respondenten telt.

3.2.1 Overzicht respons en populatieaantallen niet-bekostigd onderwijs, NSE 2017-2019

Jaar	Respons	Populatie	Percentage	Vershil	Percentage n < 5
2017	3 923	10 996	35,7		0,6
2018	3 914	11 739	33,3	-2,3	0,5
2019	3 409	10 708	31,8	-1,5	0,9

Populatiebeschrijving

Respons en populatieaantallen

Om de respons te onderzoeken, zijn responspercentages berekend op basis van het aantal respondenten en aantal inschrijvingen per type student, opleidingsvorm, locatie, opleiding en instelling. Iemand geldt als respondent als minimaal vraag 15 (algemene tevredenheid) is ingevuld. Populatieaantallen zijn voor alle opleidingen overgenomen uit de door de instellingen aangeleverde inschrijfgegevens. De respons van de student is gekoppeld aan de inschrijving die deze student heeft bij de instelling (via het emailadres). De inschrijfgegevens krijgen dus altijd prioriteit boven de opleiding die de student zelf heeft ingevuld in de enquête. Deze methode is gelijk aan die van voorgaande jaren. Het belangrijkste verschil is dat studenten voorheen de opleiding waarvoor zij ingeschreven stonden ter controle voorgelegd kregen, terwijl dit jaar alle gegevens zijn uitgevraagd bij de student. Deze uitvraag is nu overschreven door de later aangeleverde administratieve werkelijkheid voor de variabelen instelling, locatie, opleiding, type student en opleidingsvorm.

Studenten met meerdere opleidingen

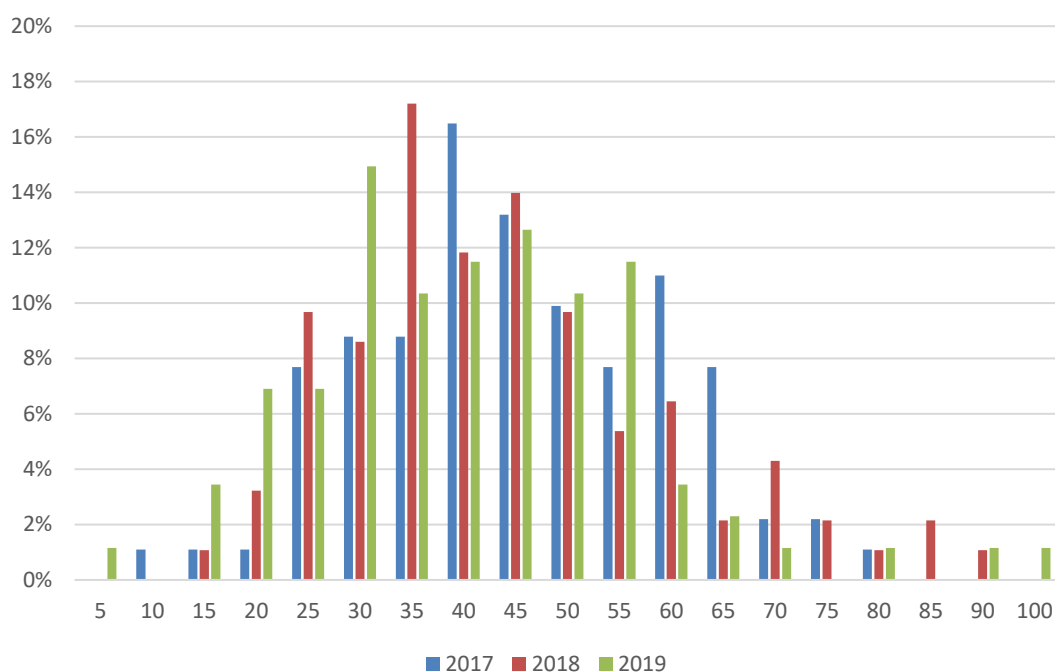
Wanneer studenten voor meerdere studies op dezelfde instelling zijn ingeschreven, wordt maar één inschrijving meegeteld (de hoofdschrijving). Wanneer studenten meerdere studies volgen aan verschillende instellingen, worden deze inschrijvingen beschouwd als aparte inschrijvingen. Dit komt overeen met de reguliere benaderingswijze van de NSE, waarbij studenten per instelling een uitnodiging ontvangen om hun studie te beoordelen.

Afbakening opleidingen en instellingen

Om de respons tussen opleidingen in deze analyse over de jaren heen te kunnen vergelijken, is een aantal keuzes gemaakt om de onderzoekspopulatie vast te stellen. Opleidingen die niet voorkomen in het voorgaande jaar, worden uitgesloten bij deze analyse. Daarnaast zijn combinaties van instelling, locatie, opleiding, type student en onderwijssoort met een responsaantal van minder dan vijf studenten niet meegenomen in de analyse.

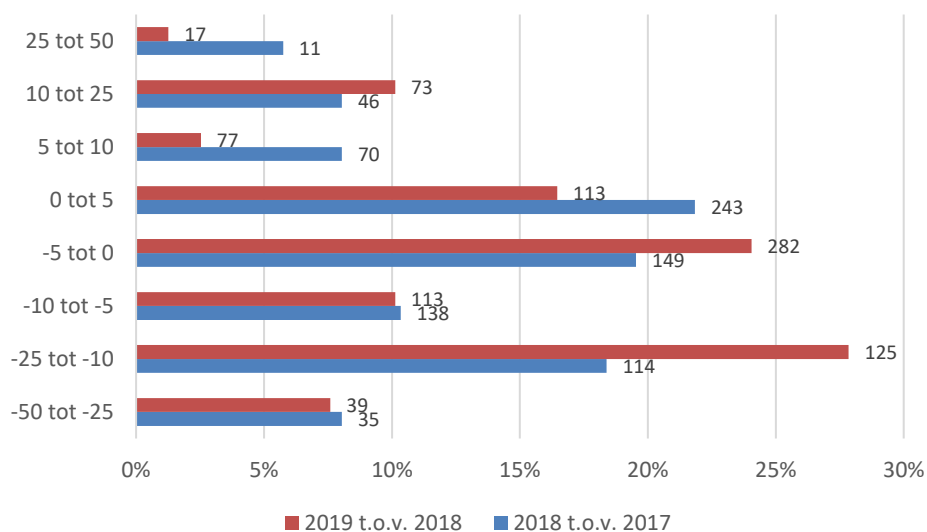
De volgende figuur laat de verdeling van responspercentages (horizontale as) zien per unieke combinatie (instelling, locatie, opleiding, type student en onderwijsvorm), afgerond op vijftallen. Hieruit komt naar voren dat lage responspercentages in 2019 in verhouding veel meer voorkomen dan in de jaren 2017 en 2018. Een derde (33%) van deze combinaties heeft in 2019 een responspercentage van maximaal 30%, terwijl dit in 2017 en 2018 bij respectievelijk 20% en 23% het geval was. Ook is de spreiding van de verdeling toegenomen ten opzichte van voorgaande jaren.

3.2.2 Verdeling afgeronde responspercentages, NSE 2017 – 2019



Om inzicht te geven in de jaar-op-jaar ontwikkeling van de hierboven weergegeven responspercentages, zijn in figuur 3.2.3 de jaar-op-jaar verschillen in responspercentages weergegeven. Deze figuur laat zien dat bij de meeste combinaties (28 procent) een afwijking van tussen de -25 en -10 procentpunten voorkomt. Gemiddeld zijn op deze combinaties 125 studenten ingeschreven in 2019 en 114 in 2018. In totaal is bij 70 procent van de combinaties de respons afgenomen in 2019, tegen 56 procent in 2018. Extreme verschillen (zowel positief als negatief) komen in 2019 niet vaker voor ten opzichte van 2018 dan in 2018 ten opzichte van 2017. Daarnaast toont de figuur dat de grootste responsverschillen voornamelijk voorkomen bij combinaties met een gemiddeld klein aantal ingeschreven studenten. Andersom komen relatief kleine verschillen het meeste voor bij combinaties met een gemiddeld groot aantal ingeschreven studenten. Bij combinaties waarbij de respons met 25 tot 50 procentpunten is toegenomen, zijn bijvoorbeeld gemiddeld 11 studenten ingeschreven in 2018 en 17 in 2019.

3.2.3 Jaar-op-jaar verschillen in responspercentages in procentpunten, inclusief gemiddeld aantal ingeschreven studenten



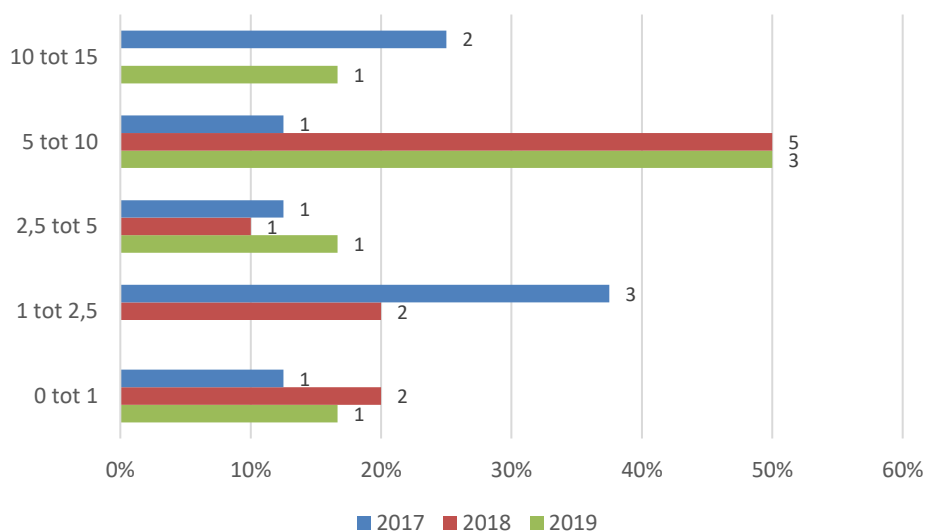
3.3 Representativiteit van de respons

Een afname van de respons betekent niet per definitie dat de respons daardoor minder betrouwbaar is of onbruikbaar wordt. Wel neemt met afnemende respons de kans toe op selectiviteit. Een gebruikelijke methode om de representativiteit van de (non) respons te beoordelen is door de verdeling van een aantal achtergrondkenmerken te vergelijken tussen de populatie (inclusief niet-respondenten) en de respons. Omdat de beschikbaarheid van achtergrondkenmerken van de populatie zich beperkt tot de kenmerken van de opleiding (onderwijsvorm en type student), worden in de volgende analyse alleen deze opleidingskenmerken vergeleken op het niveau van instelling, locatie en opleiding. Hiervoor zijn alleen de opleidingen geselecteerd die zowel voltijd als deeltijd worden aangeboden of verschillende type studenten hebben (premaster, regulier of postmaster). Dit betreft slechts een fractie van de opleidingen (respectievelijk 8, 10 en 6 opleidingen voor de jaren 2017, 2018 en 2019).

Voor de overige opleidingen geldt dat deze maar één opleidingsvorm of type student hebben, waardoor de representativiteit op het niveau van instelling, locatie en opleiding niet beoordeeld kan worden. Van deze opleidingen kan alleen worden beoordeeld of de verdeling van de respons binnen de instelling en locatie naar opleiding overeenkomt met die in de populatie. Omdat niet alle instellingen meerdere locaties hebben of meerdere opleidingen aanbieden, biedt ook deze analyse een beperkte mogelijkheid om de representativiteit te beoordelen.

Binnen de opleidingen die uitgesplitst kunnen worden naar onderwijsvorm en type student, kan overigens niet verder worden beoordeeld hoe representatief de respons is in vergelijking met de non-respons, aangezien er in 2019 geen overige achtergrondkenmerken (bijvoorbeeld geslacht, studiejaar) beschikbaar zijn om de representativiteit te onderzoeken. Deze analyse geeft dus alleen aan in hoeverre de uitsplitsing op het niveau van instellingen, locatie en opleiding representatief is naar onderwijsvorm en type student voor de opleidingen waarbij dit onderscheid mogelijk is.

3.3.1 Verschillen tussen de verdeling naar onderwijsvorm en type student in populatie en respons (in procentpunten)



Figuur 3.3.1 beschrijft de verschillen in de samenstelling onderwijsvorm en type student in de populatie en de groep respondenten. De labels met waarden tonen het aantal opleidingen waarbij de verschillen op de verticale as voorkomen. Een opleiding die voorkomt als voltijd en deeltijd variant of met reguliere en postmaster studenten wordt geteld als één opleiding. Per combinatie van opleiding en instelling is berekend welk aandeel deze studie bijvoorbeeld deeltijd volgde als reguliere student in de populatie en in de groep respondenten. Vervolgens is de absolute afwijking tussen beide verhoudingen berekend en ingedeeld in categorieën. Als een opleiding voorkomt in de categorie '5 tot 10' betekent dit bijvoorbeeld dat het percentage reguliere deeltijd studenten in de populatiegegevens 50% is, maar in de respons 42% bedraagt. Opleidingen met minder dan vijf respondenten op deze uitsplitsing zijn niet meegenomen in deze analyse.

De figuur toont dat de verschillen tussen de samenstelling van de respons en de populatie in 2019 zijn toegenomen ten opzichte van 2017. Ook ten opzichte van 2018 zijn de verschillen toegenomen, maar het gaat wel om een zeer klein aantal opleidingen dat dit verschil veroorzaakt. Slechts één van de zes opleidingen in 2019 heeft een verschil tussen de samenstelling van voltijdstudenten (en dus ook voor de deeltijdstudenten) in de respons en populatie van tussen de 10 en 15 procentpunten. De overige vijf opleidingen hebben afwijkingen die vergelijkbaar zijn met die in 2018. Hieruit kan geconcludeerd worden dat voor de meeste opleidingen in 2019 de representativiteit in termen van onderwijsvorm en type student vergelijkbaar is met voorgaande jaren. Voor één opleiding is een uitsplitsing naar deze opleidingskenmerken minder betrouwbaar. Hiervoor zou bijvoorbeeld gecorrigeerd kunnen worden door deze opleiding ook te wegen op het niveau van onderwijsvorm en type student.

4. Conclusies en aanbevelingen

In dit onderzoek staat de volgende vraag centraal: *zijn de herstelde NSE2019 data publicabel, en zo ja, op welk niveau kunnen de resultaten voor het niet-bekostigd onderwijs worden gepubliceerd?*

Een vergelijking tussen de door de instelling aangeleverde en door de student ingevulde gegevens, laat zien dat in totaal voor 32 procent van de respondenten ten minste één van de kenmerken (instelling, opleiding, locatie, studievorm of studietype) gerepareerd is. Deze reparaties komen de zowel de kwaliteit als de kwantiteit van het analysebestand ten goede. De kwaliteit is verbeterd omdat de respondentgegevens nu overeenkomen met de administratieve werkelijkheid. Omdat dit ook in vorige jaren dat de NSE is uitgevoerd zo gedaan is, komt dit ook de continuïteit van de NSE ten goede. Daarnaast is de kwantiteit verbeterd: doordat voor de respondenten die ‘anders, namelijk...’ of niets hadden ingevuld nu administratieve gegevens bekend zijn, is het aantal respondenten met een bruikbare opleiding toegenomen. Wanneer we alle unieke combinaties van instelling, opleiding, locatie, opleidingsvorm en type student bekijken gaat het gemiddeld om een toename in het responspercentage van 12,2 procentpunten.

Niettemin blijkt uit de responsontwikkeling dat ook na herstel nog steeds een afwijkende en dalende trend zichtbaar is. De meeste combinaties in 2019 (70 procent) hebben te maken met een daling in de respons ten opzichte van 2018. In 2018 was dit bij 56 procent van de combinaties het geval ten opzichte van het jaar daarvoor. Grote afwijkingen in de respons komen wel voornamelijk voor bij combinaties met een relatief klein aantal ingeschreven studenten. Hierdoor blijft het effect op de totale respons in 2019 beperkt tot een 1,5 procentpunt lagere respons ten opzichte van 2018.

Een lagere respons hoeft echter geen probleem te zijn, mits de respons representatief is voor de populatie. Het is echter niet mogelijk om de representativiteit van respondenten te beoordelen op het gewenste publicatieniveau (instelling, opleiding, locatie, type student en opleidingsvorm). Dit komt door toepassing van de Algemene verordening gegevensbescherming (AVG), waardoor de relevante achtergrondkenmerken (zoals studiejaar en geslacht) van de respondenten en non-respondenten niet beschikbaar zijn. Om een voorbeeld te geven: als het niet bekend is wat de verhouding eerstejaars en ouderejaars is in de respons en in de populatie, kan ook niet vastgesteld worden of die verhouding representatief is. Met andere woorden: omdat er geen andere kenmerken beschikbaar zijn op een lager niveau dan type student of opleidingsvorm, kan op dit niveau geen uitspraak worden gedaan over de representativiteit.

Het laagste niveau waarop representativiteit is onderzocht, is opleiding per combinatie van locatie en instelling, door de verdeling van onderwijsvorm en type student in de respons te vergelijken met die in de populatie. De meerderheid van de opleidingen in het niet-bekostigd onderwijs wordt echter maar in één onderwijsvorm aangeboden (voltijd óf deeltijd) en heeft maar één type student (regulier óf premaster óf postmaster). De representativiteit is daarom alleen onderzocht voor opleidingen die uitgesplitst kunnen worden naar onderwijsvorm of type student. Idealiter willen we voor deze opleidingen zien dat de verdeling in de respons identiek is aan de verdeling in de populatie, bijvoorbeeld tussen deeltijd- en voltijdstudenten binnen een opleiding.

Wij concluderen dat de resultaten op het niveau van instelling, opleiding en locatie gepubliceerd kunnen worden (als $n \geq 5$), maar raden daarbij wel sterk aan om te corrigeren voor verschillen in respons en populatie wat betreft onderwijsvorm en type student. Dat kan bijvoorbeeld door het toepassen van een weegfactor, waardoor de respons binnen een groep wordt opgehoogd naar de

verdeling in de populatie. Voorwaarde hierbij is dat beide groepen voorkomen in de respons. Daarnaast is het raadzaam de responspercentages te vermelden voor iedere uitsplitsing, zodat de lezer een indruk krijgt van het percentage respondenten waarop de beoordeling gebaseerd is.

Als de resultaten gepubliceerd worden op het niveau van onderwijsvorm of type student, bijvoorbeeld om de continuïteit tussen de rapportages over de jaren heen te kunnen waarborgen, moet -naast de restrictie van $n \geq 5$ per opleiding- wel vermeld worden dat de representativiteit van de respons niet onderzocht is. Om de lezer een indicatie te geven van de betrouwbaarheid moeten ook hier responspercentages worden vermeld voor iedere groep.