



Panteia

Research to Progress

Research voor Beleid | EIM | NEA | IOO | Stratus | IPM



Kans op Stage 2018

Onderzoeksverantwoording

Jan de Kok

Zoetermeer, 22 december 2017

De verantwoordelijkheid voor de inhoud berust bij Panteia. Het gebruik van cijfers en/of teksten als toelichting of ondersteuning in artikelen, scripties en boeken is toegestaan mits de bron duidelijk wordt vermeld. Vermenigvuldigen en/of openbaarmaking in welke vorm ook, alsmede opslag in een retrieval system, is uitsluitend toegestaan na schriftelijke toestemming van Panteia. Panteia aanvaardt geen aansprakelijkheid voor drukfouten en/of andere onvolkomenheden.

The responsibility for the contents of this report lies with Panteia. Quoting numbers or text in papers, essays and books is permitted only when the source is clearly mentioned. No part of this publication may be copied and/or published in any form or by any means, or stored in a retrieval system, without the prior written permission of Panteia. Panteia does not accept responsibility for printing errors and/or other imperfections.

Inhoudsopgave

1	Inleiding	5
1.1	Aanleiding	5
1.2	Opzet van deze onderzoeksverantwoording	5
2	Uitgangspunten: het onderliggende model	7
2.1	Het model	7
2.2	De kans op stage voor 2018	8
2.3	Kans op stage in zeven stappen	9
2.4	Kans op stage voor alle opleidingen bepaald	13
3	De gebruikte data	15
3.1	Registratiegegevens	15
3.2	Enquête onder erkende leerbedrijven	18
4	Complicaties en correctiefactoren	21
4.1	Correctiefactor ontbreken leerweg	21
4.2	Correctiefactoren inschrijvingen en ontbreken leerbedrijven	21
4.3	Correctiefactor omvang BPVOs	22
4.4	Correctiefactor dossierniveau	22
4.5	Aantal BPVOs per student: enkel voor publiek bekostigd onderwijs	22
4.6	Buitenlandse stages	23
4.7	Geen correctie voor toe- en uitstroom	23
5	De methodiek verder uitgewerkt	25
5.1	_1_hernoem.sps	25
5.2	_2_brongegevens vacatures corrigeren.sps	26
5.3	_3_weegfactoren leerbedrijvenenquête.sps	27
5.4	_4_leerbedrijvenenquête.sps	28
5.5	_5_correctiefactor dossier.sps	31
5.6	_6_aantal_inschrijvingen.sps	32
5.7	_7_omvang BPVOs.sps	32
5.8	_8_aanbod_jaar_Tplus1.sps	33
5.9	_9_vraag_jaar_T.sps	34
5.10	_10_vraag_jaar_Tplus1.sps	35
5.11	_11_regionale mobiliteit.sps	36
5.12	_12_indicatoren_prognoses_Panteia.sps	37



1 Inleiding

1.1 Aanleiding

SBB brengt jaarlijks de kans op stage in beeld voor de belangrijkste MBO opleidingen in Nederland. Door deze kans per kwalificatie (crebocode), type leerweg (BOL of BBL) en regio (35 arbeidsmarktregio's) te bepalen ontstaat een zeer gedetailleerd beeld van de kansen van MBO studenten om in het komende jaar een stageplaats te vinden.

In het verleden is de kans op stage bepaald op basis van de meningen en inzichten van de adviseurs en sectorspecialisten van SBB. Beschikbare informatie over aantallen studenten (per opleiding, type leerweg en/of regio) werd hierbij niet op een systematische en uniforme manier gebruikt.

In 2016 heeft SBB besloten om de kans op stage op een kwantitatieve manier te bepalen, met beschikbare registratiedata als uitgangspunt. Omdat niet alle benodigde informatie in deze registraties beschikbaar is, is een enquête onder erkende leerbedrijven gebruikt om additionele informatie te verzamelen. Deze werkwijze heeft een aantal voordelen ten opzichte van de manier waarop de kans op stage in het verleden is bepaald:

- De werkwijze is uniform: de kans op stage wordt voor alle opleidingen, alle typen leerwegen en alle arbeidsmarktregio's op dezelfde manier bepaald. De resultaten zijn hierdoor goed vergelijkbaar voor alle opleidingen.
- De werkwijze is transparant: door uitleg te geven over hoe de kans op stage berekend wordt (onder andere via deze onderzoeksverantwoording) kan iedereen nagaan hoe dit gedaan is. Dit vergemakkelijkt (interne en externe) discussies over de gevolgde werkwijze, waardoor in de toekomst verbeteringen eerder gerealiseerd kunnen worden.
- De werkwijze is geautomatiseerd: de jaarlijkse updates van de kans op stage kunnen hierdoor snel gerealiseerd worden.

In 2016 is een basisversie voor deze nieuwe werkwijze ontwikkeld. In 2017 is deze basisversie verder ontwikkeld.

1.2 Opzet van deze onderzoeksverantwoording

In het voorjaar van 2016 heeft SBB de hoofdlijnen bepaald van de manier waarop de kans op stage bepaald zou moeten worden: wat moet de kans op stage weergeven, wat zijn de onderliggende indicatoren die de kans op stage bepalen, en welke informatie is nodig om deze indicatoren te kunnen bepalen.

De kern van de methodiek is simpel: bepaal (per opleiding, type leerweg en arbeidsmarktregio) het verwachte aanbod en de verwachte vraag naar stages en leerbanen, en vergelijk die met elkaar. Er zijn echter een aantal complicerende factoren waardoor het lastig is om beschikbare data over vraag en aanbod direct met elkaar te vergelijken. In hoofdstuk 2 worden de hoofdlijnen van de methodiek besproken en worden de belangrijkste complicerende factoren vermeld.

Veel van deze complicaties hangen samen met de kenmerken van de beschikbare gegevens waarop de kans op stage berekend kan worden. Daarom worden in hoofdstuk drie de beschikbare databronnen besproken, voordat in hoofdstuk vier verder ingegaan wordt op de complicerende factoren (en hoe hiervoor wordt gecorrigeerd).



De methodiek bestaat uit een aantal SPSS syntaxen. Deze zijn zo opgezet dat ze met een minimum aan aanpassingen elk jaar opnieuw gebruikt kunnen worden. In hoofdstuk 5 bespreken we de opzet van elk van deze syntaxen. De syntaxen zelf zijn aan SBB overgedragen.

Leeswijzer

Om een overzicht te krijgen van de gebruikte methodiek is het voldoende om hoofdstuk 2 te lezen. Wie inzicht wil krijgen in de verschillende complicaties en de oplossingen die daarvoor gekozen zijn, kan daarna hoofdstukken 3 en 4 lezen. Hoofdstuk 5 is alleen relevant voor de lezers die zich willen verdiepen in hoe de methodiek (inclusief de correcties) is uitgewerkt en geprogrammeerd.



2 Uitgangspunten: het onderliggende model

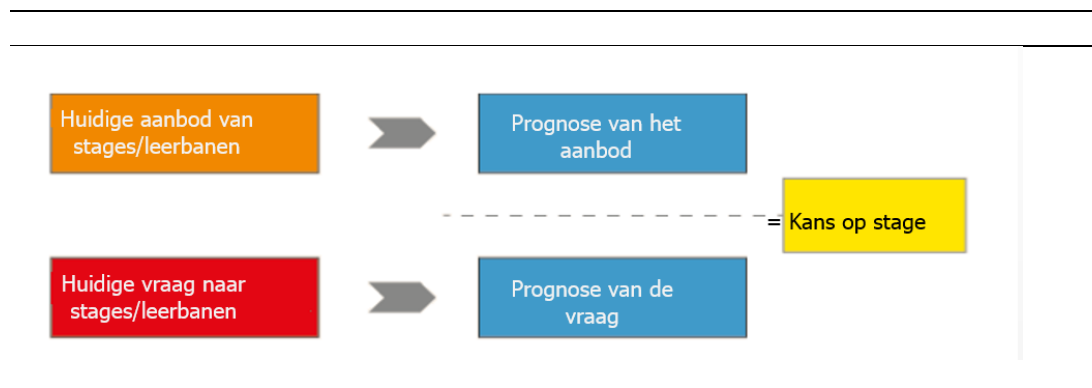
2.1 Het model

De kans op stage is door SBB als volgt gedefinieerd:

“Kans op stage is de graadmeter die een uitspraak doet over de verwachte kansen van mbo-studenten op het vinden van een stage of leerbaan. De graadmeter geeft de verhouding weer tussen het aantal stagezoekende studenten op school en het aanbod van stages en leerbanen bij bedrijven”
(Bron: SBB)

Om de kans op stage te bepalen, moet een prognose gemaakt worden van zowel het aanbod van stages en leerbanen als de vraag naar stages en leerbanen. Het aanbod komt van door SBB erkende leerbedrijven, de vraag komt van studenten die ingeschreven zijn bij MBO instellingen. Dit kunnen zowel publiek als particulier bekostigde instellingen zijn¹. De kern van het model wordt in figuur 1 weergegeven.

figuur 1 Kans op stage: de werkwijze op hoofdlijnen



Bron: SBB, 2016

2.1.1 Vier indicatoren voor de kans op stage

De uitkomsten van deze werkwijze worden sterk bepaald door de volledigheid en nauwkeurigheid van de beschikbare data. Deze is nooit 100%. Om hier rekening mee te houden is besloten om niet één, maar drie verschillende indicatoren af te leiden die allemaal inzicht kunnen geven in de (regionale spreiding van de) kans op stage (zie figuur 2).

De eerste twee indicatoren zijn gebaseerd op het model uit figuur 1. Het verschil tussen beide indicatoren is de manier waarop de prognose van het aanbod wordt bepaald (meer hierover in paragraaf 2.3.2).

Deze twee indicatoren kunnen per arbeidsmarktregio bepaald worden (en daarmee verschillen tussen arbeidsmarktregio's weergegeven), maar met één belangrijke maatstaf voor verschillen tussen regio's houden ze geen rekening: de regionale mobiliteit. Als derde indicator wordt daarom de regionale mobiliteit berekend, door per regio (en kwalificatie en type leerweg) te kijken naar de verhouding tussen het aantal

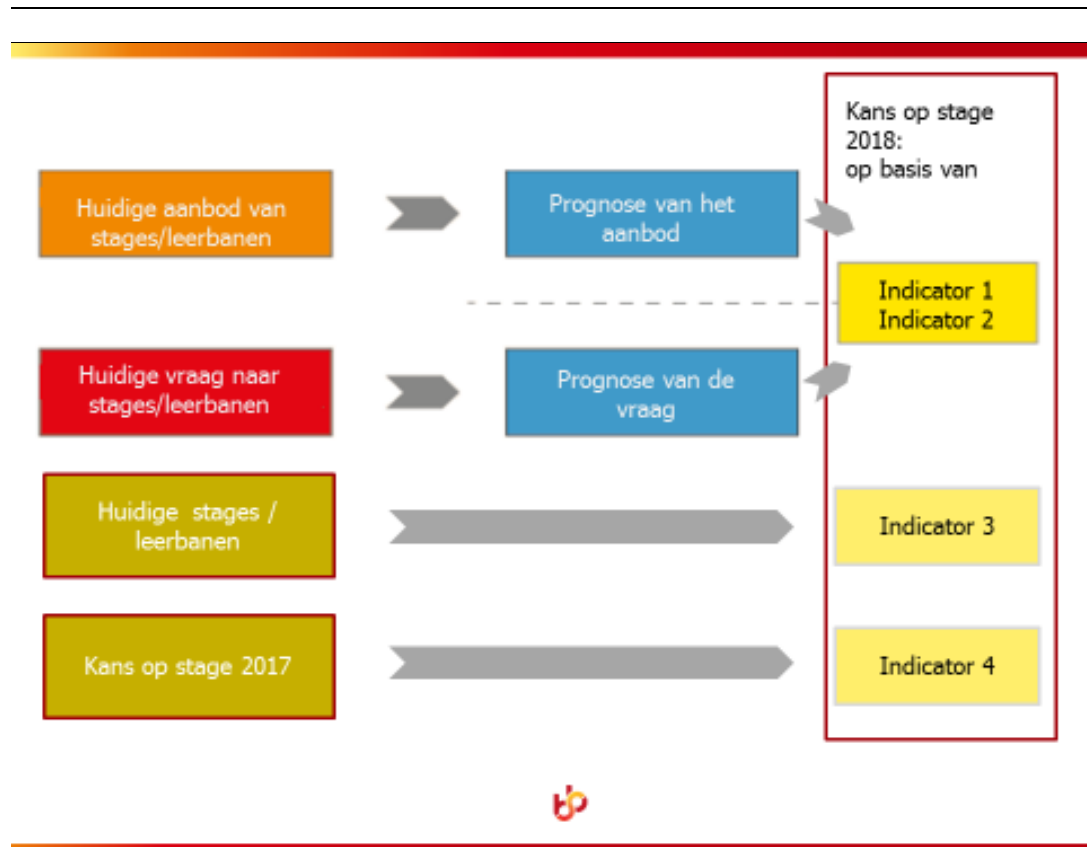
¹ Dit is een van de uitbreidingen die in 2017 is doorgevoerd: in 2016 werd enkel naar het aantal studenten gekeken dat bij publiek bekostigde MBO instellingen stond ingeschreven.



studenten dat in die regio woont en stage loopt bij een bedrijf buiten die regio, en het aantal studenten dat in een andere regio woont en stage loopt bij een bedrijf in die regio. Deze indicator kan bepaald worden op basis van informatie over alle stages en leerbanen die er in het afgelopen schooljaar zijn geweest.

Deze drie indicatoren kunnen vergeleken worden met de kans op stage zoals die vorig jaar is bepaald; dit kan als de vierde indicator worden beschouwd. Dit maakt het mogelijk om een vergelijking met de indicatoren voor dit jaar te maken, en te bepalen bij hoeveel (en welke) kwalificaties sprake is van grote verschuivingen.

figuur 2 Kans op stage: de werkwijze op hoofdlijnen



Bron: SBB, 2016

De definitieve, door SBB te publiceren kans op stage kan vervolgens berekend worden als een gewogen gemiddelde van de eerste drie indicatoren. De gewichten zijn hetzelfde als vorig jaar, om zo de vergelijkbaarheid in de uitkomsten te verhogen. Vervolgens is, ter verificatie, aan de adviseurs van SBB gevraagd hoe herkenbaar zij de aldus berekende kans op stage vinden.

2.2 De kans op stage voor 2018

De kans op stage wordt aan het begin van elk kalenderjaar gepubliceerd. De prognose betreft daarom de prognose voor het komende kalenderjaar; ten tijde van het schrijven van dit rapport was dat 2018. De huidige vraag en aanbod verwijzen in dit geval naar vraag en aanbod in 2017.



Veel van de databronnen waar de kans op stage op gebaseerd wordt (zie volgend hoofdstuk) zijn echter ingedeeld in schooljaren in plaats van kalenderjaren. Berekeningen op basis van deze data hebben daarom in eerste instantie betrekking op schooljaren, niet kalenderjaren. Het gevolg hiervan is dat de berekende prognoses betrekking hebben op het komende schooljaar (in dit geval schooljaar 2017 / 2018). We beschouwen dit als de beste prognose voor het kalenderjaar 2018.

Een alternatieve oplossing zou zijn om niet alleen prognoses voor 2017 / 2018 te maken, maar ook voor 2018 / 2019. De kans op stage kan dan in eerste instantie voor beide schooljaren berekend worden. De kans op stage voor het kalenderjaar 2018 kan vervolgens berekend worden als het gemiddelde van de 'kansen' voor de twee opeenvolgende schooljaren. Theoretisch gezien is dit wellicht zuiverder, maar de resultaten worden minder precies omdat er extra bronnen van onzekerheid worden ingebouwd (omdat prognose voor 2 jaar gemaakt moeten worden in plaats van voor 1 jaar). Daarom hebben we niet voor deze alternatieve oplossing gekozen.

2.3 Kans op stage in zeven stappen

De kans op stage wordt in zeven stappen bepaald. Deze stappen worden nu kort besproken.

2.3.1 Bepaal het huidige aanbod van stages en leerbanen

Het aanbod van stages en leerbanen op een bepaald moment valt in twee delen uiteen: het aanbod dat op dat moment is vervuld en het aanbod dat op dat moment (nog) niet is vervuld. Het aanbod dat al vervuld is, bestaat uit stages en leerbanen waarvoor een BPVO (beroepspraktijkvormingsovereenkomst) is afgesloten. Daarnaast kunnen erkende leerbedrijven vacatures hebben openstaan die nog niet zijn ingevuld.

Bij de aanvang van dit project hadden de meest recente gegevens over het aanbod van stages en leerbanen betrekking op het begin van het lopende schooljaar (2017 / 2018). Het huidige aanbod kan bepaald worden op basis van de volgende informatie:

- Uit registraties over alle BPVOs van het afgelopen schooljaar (2016 / 2017) kan afgeleid worden hoeveel BPVOs aan het einde van dat schooljaar (en dus aan het begin van schooljaar 2017 / 2018) nog lopen.
- Erkende leerbedrijven kunnen hun vacatures op Stagemarkt.nl registreren (een online registratiesysteem van SBB, voorheen Artus). Uit Stagemarkt.nl kan het aantal openstaande vacatures aan het begin van het schooljaar 2017 / 2018 gehaald worden.

2.3.2 Maak een prognose van het aanbod van stages en leerbanen

Indicator 1:

Het huidige aanbod is een standcijfer: het totaal aantal stages en leerbanen dat aan het begin van het schooljaar 2017 / 2018 wordt aangeboden. In de loop van dit schooljaar zal dit aanbod uitgebreid worden. Als tweede stap moet daarom een inschatting gemaakt worden van het totale aanbod voor het schooljaar 2017 / 2018: het totale aantal stages en leerbanen dat in de loop van dit schooljaar door erkende leerbedrijven aangeboden zal worden. Per kwalificatie, type leerweg en arbeidsmarktregio moet bepaald worden in welke mate het huidige aanbod zal toenemen. Deze toename kan van twee groepen bedrijven komen:

- bedrijven die aan het begin van het schooljaar al een aanbod van stages en/of leerbanen hadden (in de vorm van lopende BPVOs en/of op Stagemarkt.nl geregistreerde



- vacatures): zij kunnen in de loop van het schooljaar nog meer BPVOs afsluiten en/of vacatures openen;
- bedrijven die aan het begin van het schooljaar nog geen aanbod van stages en/of leerbanen hadden: zij kunnen in de loop van het schooljaar alsnog stages en/of leerbanen aanbieden (in de vorm van getekende BPVOs en/of openstaande vacatures).

Voor elk van deze groepen wordt een aparte correctiefactor berekend die de verhouding tussen het huidige aanbod en het verwachte aanbod weergeeft: voor de eerste groep is dit de correctiefactor inschrijvingen, voor de tweede groep de correctiefactor ontbreken leerbedrijven (zie paragrafen 4.1 en 5.4).

Indicator 2:

De uitkomst van indicator 1 wordt voor een belangrijk deel door de correctiefactoren bepaald. Deze worden afgeleid op basis van de uitkomsten van de enquête, wat een zekere mate van onnauwkeurigheid introduceert. Daarom is een alternatieve maatstaf bedacht, op basis van het aantal afgesloten BPVOs in het voorafgaande schooljaar. Dit geeft het bereikte evenwicht (tussen vraag en aanbod) in dat schooljaar weer. Via een enquête onder leerbedrijven kan bepaald worden in welke mate de leerbedrijven verwachten dat het aantal BPVOs zal gaan toenemen. Deze "groeivoet BPVOs" kan gebruikt worden om een alternatieve prognose van het aanbod van stages en leerbanen te maken² (in termen van aantal BPVOs).

2.3.3 Bepaal de huidige vraag naar stages en leerbanen

De vraag naar stages en leerbanen wordt uitgeoefend door studenten die ingeschreven staan bij MBO opleidingen in Nederland. Dit kunnen zowel publiek bekostigde als particulier bekostigde opleidingen zijn³. De meest recente gegevens over aantallen ingeschreven studenten hebben betrekking op het voorafgaande schooljaar (2016 / 2017).

Niet alle studenten hoeven elk jaar een stage of leerbaan te volgen. Het doel van deze stap is daarom om te bepalen welk deel van de ingeschreven studenten in het voorafgaande schooljaar daadwerkelijk een stage of leerbaan heeft gelopen: wat is in het schooljaar 2016 / 2017 de verhouding tussen het aantal studenten en het aantal afgesloten BPVOs (per kwalificatie, type leerweg en regio).

Hier wordt per kwalificatie het gemiddeld aantal BPVO's per student berekent. Dit is relevant, omdat de verhouding student en BPVO niet 1 op 1 is. Voor een aantal bol opleidingen hebben de mbo-instellingen geen stage in het onderwijsprogramma staan van de eerstejaars studenten. Pas in het tweede jaar gaan de studenten voor het eerst op stage. Daarnaast zijn er opleidingen waar studenten meerdere stages per jaar lopen om aan voldoende stage-uren te komen. We proberen hier dus niet stagetekorten inzichtelijk te krijgen, maar het benodigde aantal stageplaatsen (huidige vraag naar stages en leerbanen).

2.3.4 Maak een prognose van de vraag naar stages en leerbanen

Om een prognose te maken voor de vraag naar stages en leerbanen in het schooljaar 2017 / 2018, moeten we per kwalificatie, type leerweg en arbeidsmarktregio een inschatting maken van de groei in het aantal afgesloten BPVOs.

² Deze indicator vergelijkt het verwachte aanbod met de verwachte vraag in termen van aantallen BPVOs. Er is ook een variant van deze indicator waarbij de prognose van vraag en aanbod naar uren zijn omgerekend (waarvoor alsnog resultaten van de enquête gebruikt moeten worden), maar deze variant is uiteindelijk niet gebruikt bij het afleiden van de Kans op Stage 2018.

³ Dit is een van de uitbreidingen die in 2017 is doorgevoerd: in 2016 werd enkel naar het aantal studenten gekeken dat bij publiek bekostigde MBO instellingen stond ingeschreven.



Dit kan op verschillende manieren. Een manier is om gebruik te maken van de jaarlijkse ramingen van het ministerie van OCW over het aantal studenten in het MBO (naar type leerweg en sector). Deze ramingen zijn voor alle leerwegen en sectoren beschikbaar. Deze ramingen zijn gebruikt bij het bepalen van Kans op Stage 2017.

Het nadeel is dat deze ramingen alleen op een erg hoog aggregatieniveau (sectoren) beschikbaar zijn en daarom de ontwikkelingen voor individuele kwalificaties niet altijd goed weergeven. Daarom is voor Kans op Stage 2018 besloten om per kwalificatie en leerweg de trendmatige ontwikkeling van het aantal studenten in het bekostigde onderwijs te bepalen, en op basis hiervan een prognose te maken van de vraag naar stages en leerbanen voor 2018.

2.3.5 Bereken de eerste twee indicatoren

De eerste twee indicatoren worden berekend door per kwalificatie, type leerweg en arbeidsmarktregio de prognose van het aanbod te delen door de prognose van de vraag. Hiervoor is het noodzakelijk dat vraag en aanbod op een vergelijkbare manier gedefinieerd worden.

In eerste instantie worden vraag en aanbod gedefinieerd in termen van aantallen (BPVOs en/of vacatures). Zowel indicator 1 als indicator 2 zijn op basis van deze grondslag gedefinieerd. Er is echter een grote variatie in de omvang van BPVOs (zowel tussen verschillende kwalificaties als tussen BOL en BBL). Voor Kans op Stage 2017 is daarom indicator 2 op een alternatieve manier berekend, door zowel de vraag als het aanbod in uren uit te drukken. Voor de vraag is daar voldoende informatie voor beschikbaar, voor een deel van het aanbod ontbreekt informatie hierover in de beschikbare data. Voor Kans op Stage 2017 is toen een vraag in de enquête gesteld waarmee de gemiddelde duur van een stageplaats of leerbaan (uren per week) bepaald kon worden. Deze alternatieve variant van indicator 2 is uiteindelijk niet gebruikt bij het bepalen van de kans op stage.

Deze twee indicatoren worden zo berekend dat ze aan de volgende randvoorwaarden voldoen:

- Ze zijn per definitie positief
- Een score van 1,0 staat voor evenwicht.
- Een score lager dan 1,0 staat voor een relatief lage kans op een stage of leerbaan (de vraag naar stages is bijvoorbeeld hoger dan het aanbod).
- Een score hoger dan 1,0 staat voor een relatief hoge kans op een stage of leerbaan (het aanbod van stages is hoger dan de vraag).

2.3.6 Bereken de derde indicator: de regionale mobiliteit

Voor het afgelopen schooljaar is informatie beschikbaar over (vrijwel) alle BPVOs uit dat schooljaar. Per BPVO is niet alleen informatie beschikbaar over de opleiding (kwalificatie en type leerweg), maar ook over de arbeidsmarktregio waar de student woont (staat ingeschreven) en de arbeidsmarktregio waarin het leerbedrijf gevestigd is waar de student zijn of haar stage of leerbaan volgt.

Omdat deze informatie voor vrijwel alle BPVOs beschikbaar is, kunnen per opleiding, type leerweg en arbeidsmarktregio zowel de bruto instroom als de bruto uitstroom bepaald worden:

- Bruto uitstroom uit regio x: het aantal studenten (per opleiding en type leerweg) dat in regio x woont maar een stage of leerbaan in een andere regio heeft.



- Bruto instroom in regio x: het aantal studenten (per opleiding en type leerweg) dat niet in regio x woont maar er wel een stage of leerbaan heeft

Hiermee kan vervolgens de netto instroom bepaald worden (het verschil tussen de bruto instroom in regio x en de bruto uitstroom uit regio x).

Door dit (per opleiding en type leerweg) te relateren aan het aantal studenten dat in regio x woont, kan de relatieve netto instroom berekend worden. Deze heeft de volgende kenmerken:

- de relatieve netto instroom ligt per definitie tussen -1 en oneindig⁴;
- een waarde tussen -1 en 0 staat voor netto uitstroom;
- een waarde van 0 betekent dat er per saldo geen in- of uitstroom is;
- een positieve waarde staat voor netto uitstroom.

De derde indicator kan dan berekend worden als: relatieve netto instroom + 1. Deze indicator voldoet dan aan dezelfde randvoorwaarden als indicatoren 1 en 2.

2.3.7 Bereken de kans op stage

Na de eerste vijf stappen zijn drie verschillende indicatoren bekend:

- Indicatoren 1 en 2, die beide iets zeggen over de verhouding tussen aanbod en vraag in het lopende schooljaar.
- Indicator 3, die de mate van regionale mobiliteit in het voorafgaande schooljaar weergeeft.

Kans op stage: primair op landelijk niveau bepaald

De kans op stage wordt bepaald voor bijna 500 verschillende kwalificaties, 2 verschillende leerwegen en 35 arbeidsmarktregio's. Op regionaal niveau kunnen er in theorie dus bijna 35.000 verschillende kansen berekend worden.

Ook al is er veel informatie beschikbaar over vraag en aanbod op dit gedetailleerde niveau (kwalificatie, leerweg en regio), het is niet mogelijk om op dit gedetailleerde niveau betrouwbare uitspraken te doen over (regionale verschillen in) de kans op stage. De kans op stage wordt daarom primair op landelijk niveau bepaald, op basis van de scores voor indicatoren 1 en 2. Voor elk van deze indicatoren wordt (per kwalificatie en leerweg) de score op landelijk niveau bepaald. De kans op stage wordt dan als een gewogen gemiddelde van deze indicatoren berekend.

De gewichten zijn hetzelfde als vorig jaar, om zo de vergelijkbaarheid in de uitkomsten te verhogen: de kans voor stage 2018 is voor 25% door indicator 1 bepaald en voor 75% door indicator 2.

Regionale invulling op basis van regionale mobiliteit

De informatie over de regionale mobiliteit is vervolgens gebruikt om de regionale verschillen in de kans op stage te bepalen. De regionale kans op stage is berekend als een gewogen gemiddelde van de landelijke kans op stage en de regionale mobiliteit, waarbij de regionale mobiliteit voor 20% meetelt.

Tot slot worden zowel de nationale als de regionale variant van de kans op stage ingedikt in een variabele met vijf antwoordcategorieën:

- gering;
- matig;

⁴ De minimumscore wordt behaald als alle studenten die in een bepaalde regio wonen hun stage of leerbaan in een andere regio lopen, en geen enkele student uit een andere regio een stage in deze regio volgt. De bruto instroom is dan gelijk aan nul en de bruto uitstroom is gelijk aan het aantal studenten dat in de desbetreffende regio woont. Dan geldt: netto relatieve instroom = $0 - \text{bruto_uitstroom} / \text{aantal_woonachtige_studenten} = 0 - 1 = -1$.



- voldoende;
- ruim voldoende;
- goed.

2.4 Kans op stage voor alle opleidingen bepaald

De zojuist beschreven zes stappen zijn voor alle in 2018 geldige kwalificaties toegepast. De verschillende indicatoren en de kans op stage zijn hiermee voor alle kwalificaties berekend.

Voor de kleinste kwalificaties is de betrouwbaarheid en bruikbaarheid van deze indicatoren en de kans op stage niet duidelijk. Voor een opleiding met landelijk 45 studenten zal de regionale mobiliteit bijvoorbeeld van jaar tot jaar sterk kunnen veranderen, afhankelijk van waar de studenten wonen die dat jaar met hun opleiding stoppen en waar de studenten wonen die zich voor het eerst inschrijven. Dergelijke veranderingen vallen niet te voorspellen. Voor deze opleidingen is het bovendien waarschijnlijk dat ze niet of nauwelijks in de leerbedrijvenquête voorkomen. Dit betekent dat er geen correctiefactoren voor berekend kunnen worden.

SBB wil daarom voor de kleinste opleidingen maar een deel van de informatie publiceren. De uiteindelijke dataset met alle indicatoren en de kans op stage (voor alle kwalificaties, leerwegen en regio's) bevat een indicator die per kwalificatie en leerweg aangeeft of SBB de berekende Kans op Stage 2018 zal publiceren of niet.

De richtlijnen daarbij zijn als volgt:

1. Er zijn voor de kwalificatie landelijk minimaal 100 studenten ingeschreven.
2. Per leerweg zijn minimaal 50 studenten én minimaal 30 procent van het totaal aantal studenten van de kwalificatie (dus bol+bbl tezamen) ingeschreven.
3. Er zijn minimaal 100 leerbedrijven erkend voor de desbetreffende kwalificatie zijn.

Als niet voldaan wordt aan alle drie de richtlijnen, dan worden er geen kansen berekend.



3 De gebruikte data

3.1 Registratiegegevens

3.1.1 Aanbod

Het aanbod van stages en leerbanen (BPVOs) bestaat uit twee onderdelen:

- Aanbod dat nog niet is vervuld: vacatures bij erkende leerbedrijven.
- Aanbod dat is vervuld: bestaande (lopende) BPVOs.

Voor elk van deze twee onderdelen is een dataset met registratiegegevens beschikbaar. In tabel 1 worden de gebruikte datasets benoemd, daarna worden ze kort toegelicht.

tabel 1 Kenmerken brondata aanbod

<i>Naam</i>	<i>Inhoud</i>	<i>Waarvoor gebruikt</i>
Vacatures stagemarkt 20170808.xlsx	<ul style="list-style-type: none">• Aantal in Stagemarkt.nl geregistreerde vacatures op 8 augustus 2017 (vlak voor het begin van het huidige schooljaar)	<ul style="list-style-type: none">• Aanbod van BPVOs• Afleiden diverse correctiefactoren• Bepalen aantal bedrijven uit de populatie met vacatures op Stagemarkt.nl
Lopende bpvos 2016_2017 incl part ow.xlsx	<ul style="list-style-type: none">• Aantal lopende BPVOs in het afgelopen schooljaar (T), voor publiek en particulier bekostigd onderwijs• Gegevens over locatie studenten en leerbedrijven per BPVO• Omvang van de stage (in uren)	<ul style="list-style-type: none">• Aanbod van BPVOs• Regionale mobiliteit• Afleiden diverse correctiefactoren

Bron: Panteia, 2017

Vacatures bij erkende leerbedrijven

Erkende leerbedrijven kunnen vacatures voor stagiaires (BOL, BBL, of leerweg onbekend) op Stagemarkt.nl plaatsen. Gegevens uit Stagemarkt.nl zijn gebruikt om een overzicht te krijgen van alle openstaande vacatures aan het begin van het huidige schooljaar (schooljaar T+1). Over de omvang van de vacatures (voor hoeveel uur / week en hoeveel weken de vacature bedoeld is) is geen informatie bekend.

Lopende BPVOs

SBB heeft een database met gegevens over alle individuele BPVOs. Uit deze database is een selectie gemaakt met *geaggregeerde* gegevens over alle lopende BPVOs in het schooljaar 2016-2017. Deze selectie resulteert in een databestand waarin per record het totaal aantal lopende BPVOs wordt weergegeven voor elke unieke combinatie van

- crebocode;
- type leerweg;
- woonplaats studenten (arbeidsmarktregio);
- vestigingsregio leerbedrijven (arbeidsmarktregio);
- startdatum van de BPVO;
- einddatum van de BPVO;
- omvang van de BPVO (in uren).



Dit zijn zeer specifieke combinaties. Voor de meeste records in het databestand geldt dan ook dat het aantal lopende BPVOs voor die specifieke combinatie precies 1 is. Er zijn echter een aantal combinaties met meerdere lopende BPVOs.

De meeste waarnemingen in dit bestand hebben betrekking op individuele kwalificaties (de crebocode ligt tussen 25000 en 25999). Sommige waarnemingen betreffen inschrijvingen op dossierniveau (de crebocode ligt tussen 23000 en 23999) of op een nog hoger niveau (de crebocode ligt tussen 79000 en 79999). De inschrijvingen op dossierniveau worden gebruikt, de inschrijvingen op hoger niveau niet.

3.1.2 Vraag

De vraag naar stages en leerbanen wordt bepaald door het aantal studenten dat ingeschreven staat bij publiek en particulier bekostigde Nederlandse MBO opleidingen. Hierover heeft het SBB geaggregeerde data ter beschikking gesteld (zie tabel 2). De meeste waarnemingen in deze bestanden hebben betrekking op individuele kwalificaties. Sommige waarnemingen betreffen echter inschrijvingen op dossierniveau.

Voor sommige correcties is het nodig om per kwalificatiedossier te weten hoe het aantal gediplomeerden in dat dossier verdeeld is over de onderliggende kwalificaties. Dit kan bepaald worden op basis van een bestand met het aantal studenten dat een diploma heeft gehaald (per jaar, kwalificatie(dossier), leerweg en regio) (zie tabel 2).

Om een prognose te kunnen maken van de vraag naar stages en leerbanen in het schooljaar 2017 /2018 kan eventueel gebruikt worden gemaakt van referentieramingen die door het ministerie van OCW worden gepubliceerd. Het ministerie van OCW heeft hierover een algemeen rapport gepubliceerd alsmede een aantal tabellen met gegevens per onderwijsniveau⁵. Voor dit onderzoek zijn tabellen met gegevens over het middelbaar beroepsonderwijs relevant, waarin het aantal studenten per schooljaar is uitgesplitst naar leerweg en sector (zie tabel 2).

tabel 2 Kenmerken brondata vraag

<i>Naam</i>	<i>Inhoud</i>	<i>Waarvoor gebruikt</i>
170110 inschrijvingen 2016-2017.xlsx	<ul style="list-style-type: none"> Aantal ingeschreven studenten bij publiek bekostigde MBO opleidingsinstellingen, per crebo, leerweg, niveau en woonplaats student. 	<ul style="list-style-type: none"> Huidige vraag naar BPVOs Diverse correctiefactoren
Inschrijvingen_particulierow_1516.xlsx	<ul style="list-style-type: none"> Aantal ingeschreven studenten bij particulier bekostigde MBO opleidingsinstellingen, per crebo, leerweg, niveau en woonplaats student. 	<ul style="list-style-type: none"> Huidige vraag naar BPVOs Diverse correctiefactoren
170110 gediplomeerden 2015-2016.xlsx	<ul style="list-style-type: none"> Aantal studenten dat diploma heeft behaald, per crebo, leerweg, niveau en woonplaats student. 	<ul style="list-style-type: none"> Huidige vraag naar BPVOs Diverse correctiefactoren

⁵ Het rapport en de bijbehorende tabellen zijn beschikbaar op de volgende website: <https://www.rijksoverheid.nl/documenten/rapporten/2017/09/19/referentieraming-2017> (geraadpleegd oktober 2017)



<i>Naam</i>	<i>Inhoud</i>	<i>Waarvoor gebruikt</i>
Pubmbo17.ods	<ul style="list-style-type: none"> Referentieramingen OCW, met prognoses van aantal studenten in het MBO, per schooljaar, leerweg en sector* 	<ul style="list-style-type: none"> Prognose vraag naar BPVOs

*Bron: Panteia, 2017 * De referentieramingen delen alle kwalificaties in 4 sectoren in (techniek, zorg & welzijn, economie en groen)*

3.1.3 Overig

De databestanden die tot nu toe besproken zijn, hebben allemaal betrekking op de vraag naar of het aanbod van stages en leerbanen (BPVOs). Deze informatie moet een aantal bewerkingen ondergaan, waarvoor additionele datasets nodig zijn.

De indeling van individuele opleidingen in kwalificaties en kwalificatiedossiers (vastgelegd via crebocodes) wordt jaarlijks aangepast. Bij het bepalen van de kans op stage moet aan het begin van het traject bepaald worden wat op dat moment de geldige kwalificatiestructuur is; dit overzicht geeft aan over welke kwalificaties en kwalificatiedossiers de kans op stage berekend gaat worden (zie tabel 3).

SBB heeft een koppelschema waarin de crebocodes uit voorgaande jaren (voor zover mogelijk) ingedeeld worden volgens de huidige kwalificatiestructuur (zie tabel 3). Dit koppelschema is noodzakelijk om informatie uit voorgaande jaren op een correcte manier aan de huidige kwalificatiestructuur te koppelen.

De tabel met Brincodes van alle (publiek en particulier bekostigde) erkende MBO opleidingsinstellingen in Nederland wordt gebruikt om afgesloten BPVOs te kunnen verdelen over publiek en particulier bekostigde onderwijsinstellingen.

Tot slot heeft SBB een bestand aangeleverd waarin is aangegeven voor welke kwalificaties de Kans op Stage 2018 op de website gepubliceerd zal worden.



tabel 3 Kenmerken brondata overig

<i>Naam</i>	<i>Inhoud</i>	<i>Waarvoor gebruikt</i>
171106 indeling crebo naar kd.xlsx	<ul style="list-style-type: none"> • Overzicht van de kwalificaties voor het komende schooljaar (naam en crebocode; crebocode tussen 25000 en 25999) <ul style="list-style-type: none"> • Per kwalificatie, het bijbehorende kwalificatiedossiernummer (crebocode tussen 23000 en 23999) 	<ul style="list-style-type: none"> • Controle op databestanden met crebocodes (zorg dat alle crebocodes overeenkomen met het huidige kwalificatieoverzicht)
Koppeltabel crebo creboHKS SK MS.xlsx	<ul style="list-style-type: none"> • Koppeling tussen crebocodes zoals die in het afgelopen schooljaar zijn gebruikt en crebocodes volgens het huidige kwalificatieoverzicht. 	<ul style="list-style-type: none"> • Controle op databestanden met crebocodes (zorg dat alle crebocodes overeenkomen met het huidige kwalificatieoverzicht)
Crebo HKS naar sector.xlsx	<ul style="list-style-type: none"> • Koppeling tussen crebocodes en indeling in sectoren 	<ul style="list-style-type: none"> • Afleiden prognoses vraag per sector
Koppeltabel Brin Onderwijsinstelling.xlsx	<ul style="list-style-type: none"> • Brin code voor alle publiek en particulier bekostigde erkende MBO opleidingsinstellingen in Nederland 	<ul style="list-style-type: none"> • Indeling van alle erkende MBO opleidingsinstellingen in publiek en particulier bekostigd
Overzicht selecties.xlsx	<ul style="list-style-type: none"> • Of SBB de volgens deze methodiek berekende Kans op Stage 2018 gaat publiceren, per kwalificatie en leerweg 	<ul style="list-style-type: none"> • Om selectie te kunnen maken van opleidingen waarvoor de berekende Kans op Stage 2018 gepubliceerd zal worden.

Bron: Panteia, 2017

3.2 Enquête onder erkende leerbedrijven

3.2.1 De onderzoekspopulatie en het steekproefkader

De SBB had in juli 2017 ongeveer 252.000 erkende leerbedrijven in haar bestand. Het grootste deel hiervan is in Nederland gevestigd, maar in ruim 3,5% van de leerbedrijven gaat het om een buitenlands bedrijf of vestiging (dit zijn iets minder dan 10.000 bedrijven). Elk van deze leerbedrijven zou in principe volgend kalenderjaar een vacature voor een leerbaan of een stage kunnen hebben.

SBB heeft besloten om niet alle erkende leerbedrijven voor deze enquête te benaderen. De volgende groepen bedrijven zijn vooraf uitgesloten:

- Bedrijven die in het buitenland gevestigd zijn.
- Bedrijven die in de twee maanden voorafgaand aan deze enquête aan de BPV-monitor mee hebben gedaan. Deze bedrijven zijn uitgesloten om enquêtemoedigheid en irritatie zoveel mogelijk te beperken.
- Bedrijven die tot een ketenbedrijf behoren én die hebben aangegeven geen enquêtes te willen ontvangen van SBB zijn uitgesloten van deelname (bijvoorbeeld alle vestigingen van landelijke supermarktketens).

Het is belangrijk om onderscheid te maken tussen bedrijven die op 8 augustus 2017 één of meerdere vacatures op Stagemarkt.nl hadden geregistreerd, en bedrijven die dat niet hadden. Voor bedrijven met geregistreerde vacatures die aan de enquête meedoen is het namelijk mogelijk om een vergelijking te maken tussen de informatie op Stagemarkt.nl over het aantal geregistreerde vacatures (per kwalificatie en



leerweg) aan het begin van het schooljaar 2017 / 2018, en het totaal aantal stages of leerbanen dat het bedrijf in de loop van 2017 / 2018 verwacht aan te kunnen bieden (een van de vragen uit de enquête). Deze vergelijking wordt gebruikt bij het maken van de prognose van het aanbod van stages en leerbanen (zie paragraaf 4.1).

De onderzoekspopulatie voor dit onderzoek is daarom gestratificeerd op basis van de volgende twee kenmerken:

- Of bedrijven al dan niet zijn uitgesloten van deelname aan deze enquête. Dit bepaalt het steekproefkader (dat deel van de onderzoekspopulatie waar de steekproef uit getrokken wordt).
- Of bedrijven op 18 augustus 2017 openstaande vacatures op Stagemarkt.nl hadden.

3.2.2 Bepalen van de weegfactoren

Binnen het steekproefkader kunnen twee strata onderscheiden worden. De omvang van de netto steekproef is gelijk verdeeld over deze strata:

- De helft is getrokken uit de deelpopulatie van bedrijven die aan het begin van het schooljaar vacatures op stagemarkt.nl hadden openstaan.
- De helft is getrokken uit de deelpopulatie van bedrijven die aan het begin van het schooljaar *geen* vacatures op stagemarkt.nl hadden openstaan.

De steekproef is hierdoor niet representatief voor de onderzoekspopulatie. Hiervoor is gecorrigeerd, door weegfactoren te berekenen op basis van het per 8 augustus 2017 op Stagemarkt.nl geregistreerde aantal vacatures.

3.2.3 Aantal deelnemende bedrijven

De leerbedrijfenquête is door ruim 5800 leerbedrijven ingevuld. Volgens het bestand met gegevens over vacatures op Stagemarkt.nl waren er op 8 augustus 2017 bijna 58.000 erkende leerbedrijven met openstaande vacatures; dat is iets minder dan een kwart van het totaal aantal erkende leerbedrijven.

tabel 4 Kenmerken brondata enquête

Naam	Inhoud	Waarvoor gebruikt
H4229 Kans op stage leerbedrijven 2017.sav	<ul style="list-style-type: none"> • Resultaten enquête onder erkende leerbedrijven, exclusief organisatienummer, met diverse organisatiekenmerken 	<ul style="list-style-type: none"> • Berekenen diverse correctiefactoren
Leerbedrijven_populatie.sav	<ul style="list-style-type: none"> • Populatiebestand (alle erkende leerbedrijven aan het begin van het huidige schooljaar) 	<ul style="list-style-type: none"> • Bepalen weegfactor voor de leerbedrijfenquête

Bron: Panteia, 2017

3.2.4 Onderwerpen in de enquête

De vragenlijst gaat over de kwalificaties waarvoor het leerbedrijf erkend is. Dit houdt in dat de kwalificaties waarop de vragen gaan, per bedrijf kunnen variëren. Dit is mogelijk omdat per bedrijf bekend is voor welke kwalificaties ze een erkend leerbedrijf zijn. Deze informatie bepaalt ook het aantal vragen dat ze krijgen (een leerbedrijf dat voor 5 kwalificaties erkend is, krijgt ongeveer 5 keer zoveel vragen als een bedrijf dat voor 1 kwalificatie erkend is). Er geldt hierbij een maximum van 12 kwalificaties: als



bedrijven vacatures hebben openstaan voor meer dan 12 kwalificaties, dan beperkt de vragenlijst zich tot 12 vacatures (welke dat zijn wordt random bepaald).⁶

De volgende onderwerpen komen in de vragenlijst aan bod:

1. Per kwalificatie en per leerweg wordt vastgesteld hoeveel mbo-studenten in schooljaar 2016 / 2017 een stage of leerbaan hebben gehad⁷.
2. Per kwalificatie en per leerweg wordt gevraagd, of het bedrijf in schooljaar 2016 / 2017 stages of leerbanen beschikbaar had die niet zijn ingevuld.
3. Per kwalificatie en per leerweg wordt gevraagd, hoeveel stages of leerbanen het bedrijf in schooljaar 2017 / 2018 verwacht aan te kunnen bieden.

⁶ Dit is een verandering t.o.v. het vorige jaar. Toen zijn bedrijven over alle kwalificaties waarvoor ze erkend zijn ondervraagd. Om overbelasting te voorkomen voor bedrijven met veel erkenningen te voorkomen, zijn toen bedrijven die (volgens de voorloper van stagemarkt.nl) vacatures voor meer dan 15 verschillende kwalificaties hadden uitgesloten van de enquête.

⁷ Deze informatie is bekend in registraties van SBB; deze registratiegegevens zijn ter controle aan de bedrijven voorgelegd.



4 Complicaties en correctiefactoren

4.1 Correctiefactor ontbreken leerweg

Voor een deel van de vacatures die in Stagemarkt.nl geregistreerd staan is niet bekend om wat voor leerweg het gaat (voor 30% van alle waarnemingen staat 'BOL/BBL' als leerweg vermeld). Deze informatie kan niet geaggregeerd worden naar het niveau van de twee afzonderlijke leerwegen. Op basis van de leerbedrijvenquête wordt daarom een "correctiefactor ontbreken leerweg" berekend. Deze correctiefactor geeft voor elke kwalificatie weer wat het aandeel van het aantal stages (BOL) is in het totale verwacht aanbod van stages en leerbanen (BOL + BBL).

Deze correctiefactor wordt voor twee bestanden gebruikt:

- in Stagemarkt.nl geregistreerde vacatures waarvan de leerweg niet bekend is, worden over de twee leerwegen verdeeld.
- Informatie in de leerbedrijvenquête voor vacatures waarvoor de leerbedrijven geen voorkeur voor BOL of BBL hebben aangegeven, wordt over de twee leerwegen verdeeld.

4.2 Correctiefactoren inschrijvingen en ontbreken leerbedrijven

Het huidige aanbod van stages en leerbanen wordt bepaald op basis van administratieve gegevens over geregistreerde vacatures en lopende BPVOs (zie paragraaf 2.3.1). Om tot een prognose te komen, moet per kwalificatie, type leerweg en arbeidsmarktregio bepaald worden in welke mate het huidige aanbod zal toenemen. Deze toename kan van twee groepen bedrijven komen:

- bedrijven die aan het begin van het schooljaar al een aanbod van stages en/of leerbanen hadden (in de vorm van lopende BPVOs en/of in Stagemarkt.nl geregistreerde vacatures): zij kunnen in de loop van het schooljaar nog meer BPVOs afsluiten en/of vacatures openen;
- bedrijven die aan het begin van het schooljaar nog geen aanbod van stages en/of leerbanen hadden: zij kunnen in de loop van het schooljaar alsnog stages en/of leerbanen aanbieden (in de vorm van getekende BPVOs en/of openstaande vacatures).

Voor elk van deze groepen wordt een aparte correctiefactor berekend die de verhouding tussen het huidige aanbod en het verwachte aanbod weergeeft: voor de eerste groep is dit de correctiefactor inschrijvingen, voor de tweede groep de correctiefactor ontbreken leerbedrijven (zie paragraaf 2.3.2).

De correctiefactor inschrijvingen gaat uit van het huidige aanbod. Dit bestaat uit lopende BPVOs en in Stagemarkt.nl geregistreerde openstaande vacatures. Voor de openstaande vacatures verwachten we dat leerbedrijven het aantal beschikbare leerbanen en/of stages soms overschatten en soms onderschatten. Dit kan een aantal oorzaken hebben:

- Voor een deel van de vacatures zal gelden dat de behoefte eraan pas in de loop van het jaar ontstaat of onderkend wordt. Dit resulteert in een onderschatting van het aantal vacatures aan het begin van het schooljaar.
- Een deel van de organisaties zal fouten gemaakt hebben bij het registreren van hun vacatures in Stagemarkt.nl. Dit kan zowel in een onder- als overschatting van het aantal vacatures resulteren.
- Een deel van de organisaties heeft een vacature voor een stagiaire, die door stagiaires van meerdere opleidingen vervuld kan worden (bijvoorbeeld timmerman of allround timmerman). Voor beide opleidingen kan dan een vacature in Stagemarkt.nl geregistreerd worden. Dit resulteert in een overschatting van het aanbod van stages en leerbanen. Stel, bijvoorbeeld, dat een bedrijf in de loop van het jaar 2 stagiaires wil



plaatsen (bijv. 1 stagiair timmerman voor sept / dec 2017 en 1 stagiair timmerman voor maart / juni 2018). Dit bedrijf zou dan in Stagemarkt.nl 4 vacatures geregistreerd kunnen hebben (2 opleidingen per stagiair, voor 2 stagiaires).

Om voor deze oorzaken te corrigeren is een "correctiefactor inschrijvingen" nodig. Deze correctiefactor wordt gebaseerd op gegevens van bedrijven uit de leerbedrijvenquête die ook in het huidige aanbodbestand zitten (dit zijn bedrijven met lopende BPVOs en/of in Stagemarkt.nl geregistreerde vacatures). Voor deze bedrijven wordt de verhouding bepaald tussen het huidige aanbod (op basis van beschikbare administratieve gegevens) en het verwachte aanbod (op basis van de antwoorden op de vraag uit de enquête). Deze verhouding is de correctiefactor inschrijvingen.

Een deel van de erkende leerbedrijven zit niet in het huidige aanbodbestand. Een deel van deze bedrijven kan in de loop van het jaar alsnog stages en/of leerbanen aanbieden. De "correctiefactor ontbreken leerbedrijven" geeft het gemiddelde verwachte aanbod (aantal stages en/of leerbanen) per bedrijf weer, voor de groep bedrijven die niet in het huidige aanbodbestand zitten. Deze correctiefactor kan berekend worden op basis van de informatie van bedrijven uit de leerbedrijvenquête die niet in het huidige aanbodbestand zitten. Door deze correctiefactor te vermenigvuldigen met het totaal aantal bedrijven dat niet in het huidige aanbodbestand zit wordt een inschatting verkregen van het totale aanbod dat nog van deze groep bedrijven verwacht wordt.

4.3 Correctiefactor omvang BPVOs

Voor indicator 2 wordt een prognose gemaakt van het aanbod van stages en leerbanen, in termen van aantal BPVOs (in plaats van aantal uren). Het uitgangspunt hiervoor is het aantal BPVOs in het voorafgaande schooljaar. Dit wordt vermenigvuldigd met een "groei voet aantal BPVOs". Deze wordt berekend op basis van de leerbedrijvenquête, door een vergelijking te maken van het aantal stages en leerbanen in het voorafgaande schooljaar en het verwachte aantal stages en leerbanen in het huidige schooljaar.

4.4 Correctiefactor dossierniveau

Diverse administratieve bestanden bevatten gegevens over stages en leerbanen waarvoor wel het kwalificatiedossier bekend is (crebocode tussen 23000 en 23999) maar niet het juiste kwalificatieniveau (crebocode tussen 25000 en 25999). Bij het bepalen van de huidige en verwachte vraag en aanbod willen we deze informatie wel kunnen gebruiken. Hiervoor is het nodig om, per kwalificatiedossier en leerweg, een verdeelsleutel te maken die weergeeft hoe de vraag of aanbod op dossierniveau verdeeld moet worden over de onderliggende kwalificaties. Hiervoor wordt de "correctiefactor dossierniveau" afgeleid. Hiervoor wordt gebruik gemaakt van informatie over het aantal gediplomeerden: per kwalificatiedossier wordt bepaald hoe het aantal gediplomeerden van dat dossier verdeeld is over de onderliggende kwalificaties.

4.5 Aantal BPVOs per student: enkel voor publiek bekostigd onderwijs

Om de huidige vraag naar stages en leerbanen te bepalen, maken we een inschatting van het aantal BPVOs per student (per kwalificatie en leerweg). Omdat de verwachting bestaat dat BPVOs in het particulier bekostigd onderwijs minder goed geregistreerd



worden, wordt het aantal BPVOs per student enkel bepaald op basis van gegevens (m.b.t. aantallen BPVOs en ingeschreven studenten) over publiek bekostigd onderwijs. Vervolgens wordt aangenomen dat (per kwalificatie x leerweg) het aantal BPVOs per student voor particulier bekostigd onderwijs vergelijkbaar is met het publiek bekostigd onderwijs.

Hierdoor is het noodzakelijk dat bestanden met registratiedata over de vraag (vanuit studenten) uitgesplitst moeten worden naar kwalificatie x leerweg x arbeidsmarktregio x bekostiging; dit is gedetailleerder dan de registratiedata over het aanbod (vanuit leerbedrijven), want daarvoor is de uitsplitsing naar bekostiging niet nodig.

4.6 Buitenlandse stages

Het overgrote deel van alle stages wordt in Nederland gerealiseerd, maar een klein deel speelt in het buitenland. Voor sommige kwalificaties is het aandeel van buitenlandse stages echter hoger. Om hier rekening mee te houden wordt de Kans op Stage 2018 berekend op basis van alle stages (binnenlands en buitenlands).

Dit is mogelijk omdat bestanden met registratiedata informatie bevatten over alle stages, zowel binnenlandse als buitenlandse (de enquête onder leerbedrijven is echter beperkt tot Nederlandse vestigingen van leerbedrijven). Bij het inlezen en verwerken van deze registratiedata worden alle stages over 36 verschillende arbeidsmarktregio's verdeeld: 35 binnenlandse arbeidsmarktregio's en 1 regio 'buitenland'. Bij het berekenen van indicatoren 1 en 2 wordt op deze manier rekening gehouden met zowel binnenlandse als buitenlandse stages.

Voor het bepalen van de regionale mobiliteit (indicator 3) speelt dit geen rol: deze wordt enkel voor de 35 binnenlandse arbeidsmarktregio's bepaald.

4.7 Geen correctie voor toe- en uittreding

Een deel van het aanbod in het huidige schooljaar kan afkomstig zijn van bedrijven die aan het begin van het schooljaar nog niet als leerbedrijf erkend zijn. Met deze bedrijven wordt geen rekening gehouden bij het bepalen van het aanbod van stages en leerbanen (en ook niet bij de verschillende correctiefactoren). Om hier rekening mee te houden zou een "correctiefactor toetreding" afgeleid kunnen worden, die aangeeft met welke fractie het huidige aanbod aangepast moet worden om rekening te houden met nieuwe leerbedrijven.

Een deel van de organisaties verliest in de loop van het schooljaar haar erkenning als leerbedrijf (omdat ze de bedrijfsactiviteiten beëindigen, of om een andere reden). Een deel van deze organisaties kunnen in het huidige aanbodbestand zitten (omdat ze nog lopende BPVOs hebben en/of vacatures in Stagemarkt.nl hebben geregistreerd). Dit kan zorgen voor een overschatting van het verwachte aanbod voor het huidige schooljaar. Om hiervoor te corrigeren zou een correctiefactor uittreding afgeleid kunnen worden. Deze correctiefactor zou aan moeten geven met welke fractie het huidige aanbod aangepast moet worden om rekening te houden met het voortijdige vertrek van erkende leerbedrijven.

Beide correctiefactoren zouden in één correctiefactor gecombineerd kunnen worden: een "correctiefactor toe- en uittreding". In overleg met SBB is besloten deze correctiefactor (nog) niet uit te werken.





5 De methodiek verder uitgewerkt

De berekening van de kans op stage voor 2018 is vastgelegd in 12 verschillende SPSS syntaxen (programma's). Deze syntaxen zijn zo opgezet dat ze in een specifieke volgorde uitgevoerd moeten worden⁸:

- In de eerste syntax (syntax _1_) worden alle te gebruiken datasets ingelezen.
- In de volgende drie syntaxen (syntax_2_ t/m syntax_4_) wordt het huidige aanbod van stages en leerbanen afgeleid. Dit wordt gecombineerd met het afleiden van diverse correctiefactoren.
 - In syntax_2_ wordt een eerste versie van het aanbodbestand gemaakt door de twee databestanden met gegevens over lopende BPVOs en vacatures te combineren.
 - Voor de meeste correcties geldt dat gewogen resultaten van de enquête nodig zijn. De weegfactoren voor de enquête worden in syntax_3_ bepaald. In syntax_4_ worden vervolgens alle analyses op de enquête uitgevoerd die nodig zijn voor de verschillende correcties. In syntax_4_ wordt ook het bestand met het huidige aanbod van stages en leerbanen afgeleid.
- In syntax_5_ wordt per kwalificatiedossier en leerweg de verdeling van het aantal gediplomeerden over de onderliggende kwalificaties bepaald. Hiermee wordt de correctiefactor dossier berekend.
- Syntax_6_ begint met het bepalen van de huidige vraag naar stages en leerbanen: per kwalificatie, leerweg en regio wordt het aantal inschrijvingen in het afgelopen schooljaar bepaald, door inschrijvingen op kwalificatieniveau te combineren met inschrijven op kwalificatiedossierniveau.
- In syntax_7_ en syntax_8_ wordt de prognose van het aanbod van stages en leerbanen (ten behoeve van indicator 1) gemaakt. Eerst wordt in syntax_7_ de omvang van stages en leerbanen bepaald (uren per week en weken per jaar), vervolgens worden in syntax_8_ verschillende bestanden gecombineerd om de prognose van het aanbod van stages en leerbanen te bepalen.
- In syntax_9_ en syntax_10_ worden de huidige en verwachte vraag naar stages en leerbanen afgeleid. Er zijn twee varianten van syntax_10_: een waarbij de verwachte vraag naar leerbanen op de referentieramingen van OCW wordt gebaseerd, en een waarbij de verwachte vraag naar leerbanen op trendprognoses van Panteia wordt gebaseerd.
- In syntax_11_ wordt de regionale mobiliteit bepaald.
- De eerder afgeleide bestanden worden in syntax_12_ gebruikt om de vier indicatoren te bepalen. Vervolgens wordt de kans op stage berekend, op basis van de scores op deze 4 indicatoren. Aan het eind van deze syntax wordt een definitief databestand opgeslagen met daarin per crebo, leerweg en arbeidsmarktregio de scores op de vier indicatoren, de kans op stage, en diverse kenmerken die gebruikt kunnen worden om op de website van Kans op Stage extra achtergrondinformatie te geven. Er zijn twee varianten van syntax_12_: een waarbij de verwachte vraag naar leerbanen op de referentieramingen van OCW is gebaseerd, en een waarbij de verwachte vraag naar leerbanen op trendprognoses van Panteia is gebaseerd.

5.1 _1_hernoem.sps

In deze syntax worden alle te gebruiken datasets ingelezen en krijgen ze een nieuwe, interne naam. Deze interne naam wordt gebruikt in alle volgende syntaxen.

Als de kans op stage in de toekomst opnieuw berekend moet worden, op basis van nieuwe datasets, dan hoeven de namen van de nieuwe datasets alleen maar in deze

⁸ Naast deze syntaxen zijn er nog een aantal syntaxen waarin de labels van de verschillende opleidingen en arbeidsmarktregio's worden vastgelegd en de koppelingen tussen CREBO-indelingen van verschillende jaren en verschillende niveaus (kwalificatie, kwalificatiedossier, sector, sectorkamer, sectorunit en marktsegment). Deze worden hier niet besproken.



syntax aangepast te worden. De interne namen van de gebruikte datasets blijven immers hetzelfde.

Bij het inlezen van de verschillende datasets worden de volgende correcties en controles uitgevoerd:

- veel databronnen zijn in de vorm van Excel bestanden; deze worden omgezet in SPSS datasets;
- alle variabelen worden van de juiste labels voorzien (deze labels kunnen jaarlijks wijzigen, bijvoorbeeld bij verschillende kwalificaties en/of kwalificatiedossiers);
- bij variabelen die in meerdere datasets voorkomen wordt ervoor gezorgd dat ze overal dezelfde naam krijgen;
- de meetniveaus van variabelen worden gecontroleerd; vaak is het nodig om variabelen die in de brondata als stringvariabele (tekst) zijn gedefinieerd om te zetten naar numerieke variabelen;
- als in de toekomst blijkt dat de namen van sommige variabelen in de brondata zijn veranderd, dan zal dat in deze stap worden aangepast (alle relevante variabelen moeten elk jaar dezelfde naam hebben).

Met betrekking tot het aantal ingeschreven studenten zijn twee bestanden beschikbaar: een bestand met inschrijvingen in publiek bekostigd onderwijs en een bestand met inschrijvingen in particulier bekostigd onderwijs. Deze twee bestanden worden gecombineerd tot één bestand met alle ingeschreven studenten. Voor alle inschrijvingen is de leerweg bekend ("BOL of BBL" komt niet voor in dit bestand). Een deel van de studenten die particulier bekostigd onderwijs volgen is als examendeelnemer of OV ingeschreven (27% van alle inschrijvingen in het particulier bekostigd onderwijs), deze worden uit de dataset verwijderd.

Voor acht paren van kwalificaties is besloten om deze voor de berekening van de Kans op Stage samen te voegen, omdat ze zo sterk op elkaar lijken (bijvoorbeeld crebo's 25160 en 25528, die beide opleiden tot leidinggevende team/afdeling/project). Voor elk van deze paren is de koppeltabel tussen oude en nieuwe crebo's (zie paragraaf 3.1.3) zo aangepast dat het laagste crebo vervangen wordt door de hoogste. Dit zorgt ervoor dat de data van beide crebo's steeds bij elkaar worden opgeteld.

5.2 **_2_brongegevens vacatures corrigeren.sps**

In deze syntax wordt de informatie over openstaande vacatures en lopende BPVOs gecombineerd tot één bestand. Dit vormt een eerste versie van het aanbodbestand.

5.2.1 *Dataset openstaande vacatures inlezen*

Het databestand met informatie over openstaande vacatures ingelezen. Hierop worden enkele controles uitgevoerd. Een van deze controles is of er per leerbedrijf, kwalificatie (crebocode) en type leerweg maar 1 record in de dataset voorkomt. Dit blijkt altijd het geval te zijn.

5.2.2 *Voeg gegevens over lopende BPVOs toe*

Het bestand met (gecorrigeerde) gegevens over openstaande vacatures wordt uitgebreid met het bestand met gegevens over lopende BPVOs. Dit zijn alle waarnemingen uit het bestand met alle BPVOs uit het vorige schooljaar die aan de volgende eisen voldoen:

- ze hebben betrekking op een BOL of BBL traject;
- de kwalificatie en/of het kwalificatiedossier is bekend (de crebocode is aanwezig en ligt tussen 23000 en 25999);



- twee weken na het begin van het huidige schooljaar⁹ lopen ze nog.

Iets minder dan 70.000 records uit het bestand met lopende BPVOs van het vorige schooljaar voldoen aan deze eisen (80% BBL, 20% BOL). Dit is een toename ten opzichte van vorig jaar, toen er 65.000 records beschikbaar waren.

Dit onderdeel van de syntax bevat verwijzingen naar specifieke jaartallen. Deze verwijzingen moeten elk jaar worden aangepast.

5.2.3 Creëer eerste versie aanbodbestand

Aan het einde van `syntax_2_` is een dataset gecreëerd die per leerbedrijf, kwalificatie(dossier) en type leerweg maximaal twee records kan bevatten: een record met het aantal nog niet gerealiseerde BPVOs (aantal openstaande vacatures) en een record met het aantal gerealiseerde BPVOs (aantal lopende BPVOs aan het begin van het schooljaar). Voor het bepalen van het totale aanbod maken we geen onderscheid tussen wel of niet gerealiseerde BPVOs. Aan het einde van deze syntax worden de openstaande vacatures en lopende BPVOs daarom (per leerbedrijf, kwalificatie(dossier) en type leerweg) bij elkaar opgeteld.

Hiermee ontstaat een bestand dat per leerbedrijf, kwalificatie(dossier) en type leerweg één record heeft met daarin het aanbod (in termen van BPVOs). Dit is de eerste versie van het aanbodbestand. Dit bestand bevat informatie over ruim 346.000 lopende BPVOs en/of vacatures. Dit bestand bevat ook informatie over de arbeidsmarktregio waarin een (vestiging van een) leerbedrijf is gevestigd. Er worden 36 arbeidsmarktregio's onderscheiden: 35 binnenlandse arbeidsmarktregio's en het buitenland (dat als 1 regio wordt beschouwd).

5.3 **_3_weegfactoren leerbedrijvenquête.sps**

Het doel van deze syntax is om de juiste weegfactoren af te leiden voor de leerbedrijvenquête. Dit gebeurt door een vergelijking te maken tussen het aantal bedrijven dat de enquête heeft ingevuld en het aantal bedrijven in de onderzoekspopulatie. Hierbij moet rekening gehouden worden met de stratificatie van de steekproef.

5.3.1 De stratificatie van de populatie

Het startpunt is een bestand met een overzicht van alle door SBB geregistreerde en erkende leerbedrijven (peildatum: juli 2017). Dit bestand bevat 252.000 leerbedrijven. Een deel van deze leerbedrijven zijn uitgesloten van deelname aan deze enquête. De resterende bedrijven vormen het steekproefkader (dat deel van de onderzoekspopulatie waar de steekproef uit getrokken wordt). Binnen het steekproefkader worden twee strata onderscheiden:

- bedrijven die op 8 augustus 2017 vacatures op `stagemarkt.nl` hadden openstaan.
- bedrijven die op 8 augustus 2017 *geen* vacatures op `stagemarkt.nl` hadden openstaan.

Uit beide strata zijn evenveel bedrijven benaderd om aan de enquête mee te doen.

Om de weegfactoren te bepalen moet per stratum bepaald worden hoeveel bedrijven er in de totale populatie zitten (dus inclusief de bedrijven die uitgesloten zijn van deelname aan de enquête) en hoeveel bedrijven de enquête hebben ingevuld.

⁹ Strikt genomen hadden we de grens bij het begin van het huidige schooljaar kunnen leggen en niet twee weken daarna. Het risico hiervan is dat trajecten die vrijwel geheel in het vorige schooljaar speelden maar door omstandigheden pas in de eerste (of tweede) week van het nieuwe schooljaar zijn afgerond worden meegeteld.



Hiervoor moeten zowel het populatiebestand als het enquêtebestand gestratificeerd worden.

Om het populatiebestand te kunnen stratificeren, moet per bedrijf bepaald worden of ze per 8 augustus 2017 vacatures in Stagemarkt.nl hadden geregistreerd. Hiervoor is het populatiebestand gekoppeld aan het bronbestand met vacatures voor stageplaatsen en leerbanen.

Een deel van de bedrijven uit het bronbestand (en zelfs een paar bedrijven die aan de enquête hebben meegedaan) komen niet voor in het bestand met geregistreerde en erkende leerbedrijven. Misschien komt dit omdat er een paar maanden zit tussen de peildatum waarop het bestand met geregistreerde en erkende leerbedrijven is bepaald, de peildatum waarop het aantal geregistreerde vacatures in Stagemarkt.nl is bepaald en de periode waarin de enquête is gehouden.

Op basis van deze informatie kan het populatiebestand gestratificeerd worden, waarna het aantal bedrijven per stratum bepaald kan worden.

5.3.2 Bepalen van de weegfactoren

Om de weegfactoren voor de steekproef te bepalen wordt de steekproef op dezelfde manier gestratificeerd als het populatiebestand, waarna per stratum een vergelijking gemaakt kan worden tussen de relatieve omvang van dat stratum in de steekproef en in de populatie.

De weegfactoren zijn 0,67 (wel vacatures) en 1,17 (geen vacatures).

5.3.3 Bepalen van aantal lopende BPVOs per bedrijf

Om de correctiefactor ontbreken leerbedrijven te kunnen bepalen, moet per bedrijf het aantal lopende BPVOs worden bepaald (voor BOL, BBL, voor BPVOs waarvoor de leerweg niet bekend is, en in totaal). Dit gebeurt aan het einde van deze syntax.

5.4 _4_leerbedrijvenenquête.sps

In deze syntax worden de (gewogen) resultaten van de leerbedrijvenenquête gebruikt om diverse correctiefactoren te berekenen. Deze worden in diverse (tijdelijke) datasets opgeslagen die later toegevoegd kunnen worden aan databestanden met informatie over (vooral) het aanbod van stages en leerbanen.

5.4.1 Lees resultaten enquête in en pas structuur aan

Het bestand met de resultaten van de enquête bevat per bedrijf een record. Dit record bevat de antwoorden op vragen over meerdere crebo's en leerwegen. Dit bestand moet gecombineerd worden met de eerste versie van het aanbodbestand, waarin per bedrijf meerdere records voor kunnen komen (per crebo en leerweg een apart record).

De eerste stap van deze syntax is daarom het aanpassen van de structuur van het enquêtebestand.

5.4.2 Correctiefactor ontbreken leerweg

In vraag drie van de enquête wordt aan bedrijven gevraagd of ze, per kwalificatie en per leerweg, kunnen aangeven hoeveel stages of leerbanen ze voor het huidige schooljaar (2017 / 2018) verwachten aan te kunnen bieden.



Door de antwoorden op deze vraag te aggregeren naar het niveau van kwalificaties en leerwegen kan per kwalificatie en per leerweg het verwachte aanbod voor het huidige schooljaar bepaald worden. De correctiefactor ontbreken leerweg wordt dan berekend als het aandeel verwachte stageplaatsen (BOL) in het verwachte aanbod voor BOL en BBL samen. Er wordt geen rekening gehouden met het aantal verwachte stageplaatsen waarvoor niet bekend is of het BOL of BBL wordt. Dat is mogelijk omdat voor 77% van de antwoorden bekend is of het BOL of BBL wordt.

Voor kwalificaties met weinig studenten is het niet mogelijk om de correctiefactor op het niveau van de kwalificaties te berekenen (de enquête bevat te weinig waarnemingen voor deze kwalificatie). Voor deze kwalificaties wordt de correctiefactor op het niveau van kwalificatiedossiers bepaald.

Vervolgens worden deze correctiefactoren gebruikt om in de eerste versie van het aanbodbestand de aantallen openstaande vacatures (per kwalificatie en regio) waarvoor de leerweg niet bekend is, over BOL en BBL te verdelen. Hierna wordt een gecorrigeerde versie van het huidige aanbodbestand opgeslagen.

5.4.3 Correctiefactor inschrijvingen

Deze correctiefactor wordt bepaald op basis van bedrijven die in het huidige aanbodbestand zitten en aan de leerbedrijvenenquête hebben meegedaan. Voor zover deze bedrijven aan het begin van het huidige schooljaar vacatures in Stagemarkt.nl hadden geregistreerd, kunnen we enkel gegevens gebruiken van bedrijven die hierbij aangeven op welk type leerweg deze vacatures betrekking hebben (deze informatie is noodzakelijk om de correctiefactor per leerweg te kunnen bepalen).

Voor deze selectie van bedrijven wordt (per leerweg) het huidige aanbod (lopende BPVOs en openstaande vacatures voor alle kwalificaties) vergeleken met het verwachte aanbod. Voor het verwachte aanbod worden de antwoorden op vraag 3 uit de leerbedrijvenenquête gebruikt.

Voor het huidige aanbod is de correctiefactor ontbreken leerweg al toegepast. Dit zorgt ervoor dat het gehele huidige aanbod over BOL en BBL is verdeeld. Deze correctiefactor wordt daarom ook gebruikt om het verwachte aanbod over BOL en BBL te verdelen. Als dit gebeurd is kan de correctiefactor berekend worden, door per kwalificatiedossier en leerweg het totale verwachte aanbod (aantal verwachte stages of leerbanen) te delen door het totale huidige aanbod (aantal openstaande vacatures en lopende BPVOs).

Voor 56% van de combinaties kwalificatiedossier x leerweg zijn er minder dan 75 waarnemingen in de enquête beschikbaar. Voor deze opleidingen wordt de correctiefactor inschrijvingen op het niveau van enkel de leerweg bepaald.

De correctiefactor inschrijvingen voor BOL is gemiddeld 3,7. Er zijn twee uitschieters waarbij de correctiefactor hoger dan 10 is (14 en 19). Deze waarden worden door 10 vervangen. Voor BBL is de correctiefactor inschrijvingen gemiddeld 4,3 met 7 uitschieters van 12 tot 35. Deze waarden worden door 12 vervangen.

5.4.4 Correctiefactor ontbreken leerbedrijven

Deze correctiefactor wordt berekend op basis van de groep bedrijven die aan de leerbedrijvenenquête hebben meegedaan maar niet in het aanbodbestand zitten.



Voor deze groep bedrijven wordt, per kwalificatiedossier en leerweg, het gemiddeld aantal stages of leerbanen bepaald dat ze in de loop van het huidige schooljaar verwachten (op basis van vraag 3). Ook hier wordt het aantal stages of leerbanen waarvoor de leerweg nog onbekend is, over BOL en BBL verdeeld op basis van de eerder afgeleide correctiefactor ontbreken leerweg.

Een klein deel van de waarnemingen (7 van de ruim 4500) betreft bedrijven die voor het desbetreffende kwalificatiedossier en leerweg meer dan 25 plaatsen verwachten. Wellicht kloppen deze opgaven, maar ze komen zo weinig voor en hebben zo'n groot effect op de resultaten dat ze als uitbijter behandeld worden (en dus niet gebruikt worden bij de verdere berekeningen).

Voor 82% van de combinaties kwalificatiedossier x leerweg zijn er minder dan 75 waarnemingen in de enquête beschikbaar. Voor deze opleidingen wordt de correctiefactor inschrijvingen op het niveau van sectorkamer x leerweg bepaald.

Het gemiddeld aantal verwachte stages of leerbanen (gemiddeld per erkend leerbedrijf) kent voor BBL geen uitschieters, maar wel een paar voor BOL (waar het gemiddelde een paar keer hoger is dan 1,9: dat zou betekenen dat elk erkend leerbedrijf dat aan het begin van het jaar nog geen BPVO had lopen en nog geen vacature open had staan, bijna 2 stages verwacht te openen). Het gemiddelde is daarom op 1,9 begrensd.

Tot slot wordt het gemiddeld aantal verwachte stages of leerbanen (gemiddeld per erkend leerbedrijf) vermenigvuldigd met het totaal aantal voor dat kwalificatiedossier erkende leerbedrijven dat niet in het aanbodbestand zit. Het resultaat (een prognose van het aanbod van stages en leerbanen afkomstig van leerbedrijven die niet in het huidige aanbodbestand zitten) wordt als een apart databestand opgeslagen.

Op basis van deze analyse zorgt de correctie ontbreken leerbedrijven voor een uitbreiding met 465.500 leerbanen (BBL) en 698.300 stages (BOL). Dat is een forse stijging ten opzichte van Kans op Stage 2017. Veel meer dan vorig jaar.

5.4.5 Correctiefactor uren_week

Voor Kans op Stage 2017 zijn in de enquête enkele vragen gesteld, op basis waarvan een inschatting gemaakt kon worden van het gemiddeld aantal uren per week dat een student aan stages en leerbanen geacht wordt te werken. Deze informatie werd gebruikt om een variant van indicator 2 te berekenen, waarbij zowel de verwachte vraag als het verwachte aanbod in uren werden uitgedrukt. Deze variant is uiteindelijk niet gebruikt bij het afleiden van de Kans op Stage 2017, en ook niet voor Kans op Stage 2018.

Om tijdens de analyses voor Kans op Stage 2018 toch de mogelijkheid te hebben om indicator 2 op deze alternatieve manier te berekenen, wordt de correctiefactor uren_week van vorig jaar dit jaar opnieuw ingelezen en gebruikt.

5.4.6 Groeivoet aantal BPVOs

Het aanbod van stages en leerbanen is de optelsom van vervuld aanbod (afgesloten BPVOs) en onvervuld aanbod (openstaande vacatures). Voor het voorafgaande schooljaar zijn hierover twee vragen in de enquête gesteld (vragen 1 en 2). Over het verwachte aantal stages en leerbanen voor het huidige schooljaar is 1 vraag in de enquête opgenomen (vraag 3). Voor dit verwachte aantal plaatsen is niet duidelijk of deze allemaal vervuld zullen worden of niet. Daarom is de groeivoet bepaald op basis



de verhouding tussen de antwoorden op vraag 3 enerzijds, en de antwoorden op vragen 1 en 2 anderzijds.

Het al dan niet opnemen van vraag 2 maakt veel uit voor de resulterende groeivoet: als vraag 3 enkel met vraag 1 vergeleken wordt dan is de resulterende groeivoet veel hoger dan wanneer de antwoorden van vragen 1 en 2 bij elkaar opgeteld worden.

De groeivoet wordt per leerweg berekend als: het verschil tussen het verwachte aanbod (vraag 3) en het aanbod in het voorafgaande jaar (vragen 1 en 2), gedeeld door het aanbod in het voorafgaande jaar (vragen 1 en 2)¹⁰. Dit gebeurt op verschillende aggregatieniveaus: per kwalificatie, per kwalificatiedossier, per marktsegment en voor alle opleidingen.

De uiteindelijke groeivoet per kwalificatie wordt bepaald door het aantal waarnemingen:

- als er voor een kwalificatie minstens 55 waarnemingen zijn, dan wordt de correctiefactor op kwalificatieniveau gebruikt;
- als dit niet het geval is, maar voor het bijbehorende kwalificatiedossier minstens 55 waarnemingen zijn, dan wordt de correctiefactor op kwalificatiedossierniveau gebruikt;
- als dit niet het geval is, maar voor het bijbehorende marktsegment minstens 55 waarnemingen zijn, dan wordt de correctiefactor op marktsegmentniveau gebruikt;
- als er ook op het niveau van het marktsegment minder dan 55 waarnemingen zijn, dan wordt het gemiddelde voor alle opleidingen (per type leerweg) gebruikt.

Bijzonder hoge en lage groeivoeten zijn vervolgens aangepast: groeivoeten van meer dan 50% zijn op 50% gesteld, groeivoeten van minder dan -50% (halvering van het aanbod) zijn op -50% gesteld.

Om de te hanteren grenswaarde te kunnen bepalen, zijn voor Kans op Stage 2017 indicatoren 1 en 2 (en de uiteindelijke kans op stage) berekend op basis van verschillende grenswaarden. Vervolgens zijn de resulterende waarden voor deze indicatoren (en de kans op stage) voor een selectie van opleidingen met elkaar vergeleken. Een grenswaarde van 55 waarnemingen leverde toen de meest plausibele resultaten op. Voor Kans op Stage 2018 is deze grenswaarde overgenomen.

De resulterende correctiefactor wordt in een apart databestand opgeslagen.

5.5 **_5_correctiefactor dossier.sps**

Het bestand met informatie over studenten die in schooljaar 2015/2016 hun diploma haalden wordt geopend. De correctiefactor dossierniveau wordt vervolgens per kwalificatie en leerweg berekend als de verhouding tussen:

- het aantal studenten dat een diploma heeft gehaald voor deze *kwalificatie* en leerweg;
- het aantal studenten dat een diploma heeft gehaald voor het bovenliggende *kwalificatiedossier* en leerweg.

De resulterende correctiefactor wordt in een apart databestand opgeslagen.

¹⁰ In formulevorm: $(\text{verwacht aanbod} - \text{aanbod voorafgaand jaar}) / (\text{aanbod voorafgaand jaar})$. Dit is equivalent met $(\text{verwacht aanbod}) / (\text{aanbod voorafgaand jaar}) - 1$.



5.6 **_6_aantal_inschrijvingen.sps**

De berekende correctiefactor dossierniveau wordt in deze syntax gebruikt om inschrijvingen op kwalificatiedossierniveau over de onderliggende kwalificaties te verdelen.

Uitgangspunt hiervoor is het bestand met alle inschrijvingen (zowel publiek bekostigd als particulier bekostigd onderwijs) op kwalificatieniveau. Door alle inschrijvingen te aggregeren naar kwalificatiedossierniveau wordt bepaald hoeveel inschrijvingen er per kwalificatiedossier zijn. Dit gebeurt op twee manieren:

- Inclusief de inschrijvingen op dossierniveau
- Exclusief de inschrijvingen op dossierniveau (dit gebeurt door eerst het aantal inschrijvingen per kwalificatie x leerweg x woonregio x bekostigingswijze te bepalen. Vervolgens wordt dit geaggregeerd naar het niveau van kwalificatiedossier x leerweg, maar dan enkel voor de inschrijvingen op kwalificatieniveau)

Het verschil tussen deze twee totalen (het totaal inclusief en exclusief inschrijvingen op dossierniveau) is precies het aantal inschrijvingen op dossierniveau, dat over de onderliggende kwalificaties verdeeld moet worden.

Hiervoor kan de correctiefactor dossierniveau gebruikt worden. Deze correctiefactor geeft per kwalificatie en leerweg weer welk deel van de dossierinschrijvingen aan deze kwalificatie en leerweg moet worden toegerekend. De dataset met inschrijvingen is echter op een lager aggregatieniveau gedefinieerd: resultaten zijn niet alleen naar kwalificatie en leerweg, maar ook nog naar regio verbijzonderd. Bij het toerekenen van de dossierinschrijvingen aan de inschrijvingen (per kwalificatie, leerweg en regio) moet dus ook een verdeelsleutel gebruikt worden om de aan een kwalificatie en leerweg toegekende inschrijvingen over de regio's te verdelen. Hiervoor gebruiken we, per kwalificatie en leerweg, de verdeling van het aandeel inschrijvingen over de regio's.

Voor de verdeling van de dossierinschrijvingen over de onderliggende kwalificaties worden zodoende (per kwalificatie, leerweg, regio en bekostigingswijze) de volgende drie variabelen met elkaar vermenigvuldigd:

- het aantal inschrijvingen op dossierniveau
- de correctiefactor dossierniveau
- per kwalificatie en leerweg, het aandeel inschrijvingen per regio.

Aan het einde van deze syntax wordt bepaald worden voor welke kwalificaties er minstens 80 inschrijvingen zijn en voor welke minstens 100.

5.7 **_7_omvang BPVOs.sps**

In deze syntax wordt de omvang van BPVOs in het afgelopen schooljaar bepaald: het gemiddeld aantal uren per week en het gemiddeld aantal weken per jaar van stages en leerbanen waarvoor in het afgelopen schooljaar BPVOs zijn afgesloten.

Uitgangspunt hiervoor is het (in syntaxen `_1_` en `_2_` aangepaste) bestand met informatie over alle afgesloten BPVOs. Dit bestand bevat informatie over onder andere de startdatum, de geplande en werkelijke einddatum, en de omvang van de stage in uren.

De geplande en werkelijke duur van de stageperiode wordt bepaald door het aantal dagen te bepalen tussen de startdatum van de BPVO enerzijds, en de geplande en



werkelijke einddatum anderzijds. Dit is inclusief weekenden en mogelijke zon- en feestdagen. We corrigeren niet voor mogelijke zon- en feestdagen, maar wel voor weekenden. Dit doen we door het aantal geplande dagen te vermenigvuldigen met 5/7. Dit levert een inschatting op van het aantal werkdagen van de stage.

Het gemiddeld aantal uren per werkdag (per stage of leerbaan) wordt berekend als de verhouding tussen de omvang van de stage in uren (dit is het totaal aantal voor deze stage / leerbaan te werken uren) en de geplande duur van de stage (in werkdagen).

Vervolgens wordt het gemiddeld aantal uren per week bepaald als 5 * het gemiddeld aantal uren per werkdag. Stages en leerbanen met een bijzonder hoog of bijzonder laag aantal uren per werkdag (minder dan 0,4 uur¹¹ of meer dan 8,0 uur) blijven hierbij buiten beschouwing.

Ook het aantal weken dat een stage of leerbaan duurt wordt bepaald; hierbij wordt niet zozeer het totaal aantal weken geteld, maar het aantal weken *in het afgelopen schooljaar*. Dit is bijvoorbeeld relevant voor BBL leerbanen die een looptijd van meerdere jaren kunnen hebben.

Deze gegevens zijn voor beide type leerwegen afgeleid. Zowel voor BOL als voor BBL laten de gegevens een behoorlijke variatie zien, zowel qua aantal uren per week als aantal weken per jaar. In overleg met SBB is besloten om voor de BBL leerwegen uit te gaan van een vaste werkbelasting:

- elke leerbaan duurt precies 1 schooljaar; dit komt neer op 52 weken per jaar
- elke leerbaan telt evenveel uren per week; we gaan uit van 3 dagen / week (omdat de data laten zien dat BPVOs die voor BBL leerwegen zijn afgesloten gemiddeld ongeveer 3 dagen per week in beslag nemen).

Het resultaat van deze syntax is een apart databestand waarin per kwalificatie en per leerweg twee kenmerken van de omvang van stages en leerbanen zijn opgenomen:

- het gemiddeld aantal uren per week
- het gemiddeld aantal weken per jaar

Voor BBL leerbanen is dit voor alle kwalificaties hetzelfde, voor BOL stages kan dit variëren.

5.8 **_8_aanbod_jaar_Tplus1.sps**

Het doel van deze syntax is om het aanbod van stages en leerbanen voor het lopende schooljaar te bepalen. Dit bestand wordt op een aantal manieren aangepast en uitgebreid.

5.8.1 *Bepaal gemiddelde omvang huidig aanbod*

Voor de lopende BPVOs kan de omvang van het huidige aanbod bepaald worden op dezelfde manier als in syntax `_7_` is gedaan; het enige verschil is dat nu gekeken wordt naar het deel van de lopende BPVOs dat in het lopende schooljaar valt (in syntax `_7_` is gekeken naar de omvang van BPVOs in het voorafgaande schooljaar).

Voor de openstaande vacatures ontbreekt deze informatie; hiervoor wordt gebruik gemaakt van de correctiefactor `uren_week` die in syntax `_4_` is afgeleid (op basis van informatie uit de leerbedrijvenquête).

¹¹ Dit komt overeen met minder dan 2 uur per week.



Hierna wordt, per kwalificatie en leerweg, de gemiddelde omvang per stage en leerbaan bepaald. Dit is het product van het gemiddeld aantal uren per week en het gemiddeld aantal weken per jaar.

Voor BBL is besloten om uit te gaan van een vaste werkbelasting per leerbaan: 1 leerbaan = 1 schooljaar. De gegevens laten zien dat de aanname dat een BBL-er gemiddeld genomen 4 dagen/week (32 uur) werkt niet klopt; we stellen het daarom gelijk aan 3 dagen/week (=24 uur/week), wat beter overeenkomt met de resultaten.

5.8.2 Prognose van groei bij leerbedrijven die al in huidige aanbodbestand zitten
In `syntax_4_` is de correctiefactor inschrijvingen bepaald. Voor de bedrijven uit de leerbedrijvenquête die in het huidige aanbodbestand zitten, geeft deze factor de verhouding weer tussen het verwachte aanbod (aantal verwachte stages of leerbanen) en het huidige aanbod (aantal openstaande vacatures en lopende BPVOs).

Deze factor wordt nu (per kwalificatie, leerweg en regio) vermenigvuldigd met het totale huidige aanbod. Hierdoor ontstaat een prognose van het aanbod van stages en leerbanen bij leerbedrijven die al in het huidige aanbodbestand zitten.

5.8.3 Prognose van groei bij leerbedrijven die nog niet in huidige aanbodbestand zitten

De correctiefactor ontbreken leerbedrijven is ook in `syntax_4_` bepaald, op basis van de bedrijven uit de leerbedrijvenquête die *niet* in het huidige aanbodbestand zitten. Deze correctiefactor is per sectorkamer en leerweg bepaald en geeft (in absolute cijfers) weer met hoeveel het huidige aanbod verhoogd (of verlaagd) moet worden.

Om deze correctie over alle kwalificaties, leerwegen en regio's te verdelen, wordt eerst per sectorkamer en leerweg de relatieve verdeling van het huidige aanbod over de onderliggende kwalificaties en regio's bepaald. Dit levert verschillende verhoudingen op. Door deze verhoudingen met de correctiefactor ontbrekende leerbedrijven te vermenigvuldigen ontstaat per kwalificatie, leerweg en regio een prognose van het aanbod voor het lopende schooljaar door leerbedrijven die nog niet in het huidige aanbodbestand zitten.

5.8.4 Prognose totale aanbod op nationaal niveau

De twee hiervoor beschreven correcties kunnen bij elkaar opgeteld worden. Dit levert een prognose van het totale aanbod van stages en leerbanen voor het lopende schooljaar van alle erkende leerbedrijven. Deze gegevens zijn per regio beschikbaar maar nog niet nationaal. Daarom worden de prognoses per kwalificatie en leerweg geaggregeerd om tot een prognose van het aanbod op nationaal niveau te komen.

5.9 **`_9_vraag_jaar_T.sps`**

5.9.1 Combineer bestanden met aantal inschrijvingen en gemiddeld aantal uren per BPVO

De basis voor het huidige vraagbestand is in `syntax_6_` gelegd. Hier is een bestand gemaakt waarin per kwalificatie en leerweg het totaal aantal ingeschreven studenten voor het afgelopen schooljaar is opgenomen.

Dit bestand wordt geopend, waarna informatie over de gemiddelde omvang van BPVOs (wat in `syntax_7_` is bepaald) wordt toegevoegd.



5.9.2 De verhouding tussen aantal inschrijvingen en de vraag naar stages

Het openstaande bestand bevat informatie over het aantal ingeschreven studenten in het afgelopen schooljaar. Dit wordt verder uitgebreid met informatie over het aantal afgesloten BPVOs in het afgelopen schooljaar (dit is in `syntax_5_` bepaald).

Per kwalificatie en leerweg is het nu mogelijk om het gemiddeld aantal afgesloten BPVOs per ingeschreven student te bepalen. Dit wordt alleen gedaan voor studenten die publiek bekostigd onderwijs volgen. De indruk is dat bij particulier bekostigd onderwijs de registratie van BPVO's niet zo compleet is als bij publiek bekostigd onderwijs. Dit zou lijden tot een onderschatting van het gemiddeld aantal afgesloten BPVOs per ingeschreven student in het particulier bekostigd onderwijs. In plaats daarvan nemen we aan dat het gemiddeld aantal afgesloten BPVOs per ingeschreven student voor studenten in het particulier bekostigd onderwijs hetzelfde is als voor student in het publiek bekostigd onderwijs.

Voor de leerweg BOL wordt deze verhouding gebruikt om (per kwalificatie *en regio*) de vraag naar BPVOs te bepalen: de vraag naar BPVOs wordt berekend door het aantal ingeschreven studenten met deze verhouding te vermenigvuldigen. Het resultaat is vrijwel gelijk aan het aantal afgesloten BPVOs (immers, de verhouding is het aantal afgesloten BPVOs gedeeld door het aantal ingeschreven studenten).

Voor de leerweg BBL wordt de vraag naar BPVOs gelijkgesteld aan het aantal ingeschreven studenten. Dit is gedaan omdat een student zich niet voor een BBL opleiding kan inschrijven zonder een leerbaan bij een erkend leerbedrijf te hebben. Het is daarom opvallend dat gemiddeld genomen voor de BBL leerweg het aantal BPVOs per student 1,4 is en niet 1,0. Bij een volgende editie van de Kans op Stage zou dit verder onderzocht kunnen worden.

5.10 `_10_vraag_jaar_Tplus1.sps`

Er zijn twee varianten van deze syntax ontwikkeld. Voor de berekening van Kans op Stage 2018 is de voorkeur gegeven aan de tweede variant (op basis van prognoses van Panteia over het aantal inschrijvingen).

5.10.1 `_10_vraag_jaar_Tplus1_prognoses_DUO.sps`

Het Ministerie van OCW publiceert ramingen over het aantal studenten in het MBO, naar type leerweg en sector. Op basis van deze ramingen kan per leerweg en sector de groeivoet van het aantal studenten bepaald worden. We nemen aan dat het aantal BPVOs per student niet verandert. Hierdoor volstaat het om de vraag naar stages in leerbanen in het afgelopen schooljaar te vermenigvuldigen met de groeivoet van het aantal studenten om een prognose te maken van de vraag naar stages en leerbanen voor het huidige schooljaar.

5.10.2 `_10_vraag_jaar_Tplus1_prognoses_Panteia.sps`

De ramingen van het Ministerie van OCW zijn op een hoog aggregatieniveau gemaakt. Er zit hierdoor niet veel variatie in de ontwikkeling per kwalificatie. In werkelijkheid komen er regelmatig grote veranderingen voor (zowel groei als krimp). Panteia heeft daarom alternatieve ramingen gemaakt, door per crebo x leerweg een prognose te maken. Deze ramingen zijn in Excel gemaakt. De resulterende groeivoeten worden in SPSS ingelezen.

We nemen aan dat het aantal BPVOs per student niet verandert. Hierdoor volstaat het om de vraag naar stages in leerbanen in het afgelopen schooljaar te vermenigvuldigen



met de groeivoet van het aantal studenten om een prognose te maken van de vraag naar stages en leerbanen voor het huidige schooljaar.

De ramingen zijn in Excel berekend als het gemiddelde van de groeivoeten van het aantal inschrijvingen over de afgelopen 5 jaar. Voor kwalificaties met meer dan 50 inschrijvingen in 2016/2017 is een maximale groei van 50% gehanteerd; voor kwalificaties met 50 of minder inschrijvingen is het maximum bij 100% gelegd (een verdubbeling van het aantal inschrijvingen). Niet voor alle kwalificaties zijn inschrijvingen over 5 jaar bekend; als dat niet het geval is wordt gekeken naar de afgelopen 3 jaar¹². Als het niet mogelijk is om een groeivoet op kwalificatieniveau te berekenen dan wordt gekeken naar de groeivoet op dossierniveau (over de afgelopen 3 jaar). Als ook dat niet mogelijk is, dan wordt het aantal inschrijvingen constant gehouden.

5.11 **_11_regionale mobiliteit.sps**

De regionale mobiliteit wordt bepaald op basis van het bestand met informatie over alle BPVOs uit het afgelopen schooljaar. Alleen waarnemingen over de leerwegen BOL en BBL, waarvoor de kwalificatie, woonregio en regio van leerbedrijf bekend zijn, kunnen hiervoor gebruikt worden.

5.11.1 *Bereken bruto en netto in- en uitstroom*

Per crebo en per leerweg worden de volgende kengetallen per regio berekend:

- studenten_#: aantal studenten dat in regio # woont (de noemer)
- bruto_uitstroom_abs_regio_#: aantal studenten dat in regio # woont maar elders werkt
- bruto_instroom_abs_regio_#: aantal studenten dat NIET in regio # woont maar er wel werkt

(_abs staat voor absoluut: het gaat hier om het absolute aantal studenten)

Hiermee kunnen in- en uitstroomcijfers berekend worden:

- Bruto_uitstroom_rel_regio # = $\text{bruto_uitstroom_abs_regio_} \# / \text{studenten_} \#$
 - Bruto_instroom_rel_regio # = $\text{bruto_instroom_abs_regio_} \# / \text{studenten_} \#$
 - Netto_instroom_rel_regio # = $\text{bruto_instroom_rel_regio_} \# - \text{bruto_uitstroom_rel_regio_} \#$
- (_rel staat voor relatief: het gaat hier om een verhouding)

Voordat deze kengetallen berekend worden, wordt de correctiefactor dossier gebruikt om rekening te kunnen houden met informatie die enkel op kwalificatiedossierniveau beschikbaar is.

5.11.2 *Bereken indicator 3 voor regionale mobiliteit*

Indicator drie moet (net als indicatoren 1 en 2) aan de volgende randvoorwaarden voldoen:

- Een score van 1,0 staat voor evenwicht.
- Een score lager dan 1,0 staat voor een relatief lage kans op een stage of leerbaan; in het geval van regionale mobiliteit wil dit zeggen dat er meer uitstroom dan instroom is.
- Een score hoger dan 1,0 staat voor een relatief hoge kans op een stage of leerbaan; in het geval van regionale mobiliteit wil dit zeggen dat er meer instroom dan uitstroom is.

Het uitgangspunt voor indicator 3 is de relatieve netto instroom. Deze variabele heeft de volgende kenmerken:

- netto_instroom_relatief ligt per definitie tussen -1 en oneindig;
- een waarde tussen -1 en 0 staat voor netto uitstroom;
- een waarde van 0 betekent dat er per saldo geen in- of uitstroom is;

¹² Waarbij dezelfde maxima zijn toegepast als bij 5 jaar.



- een positieve waarde staat voor netto uitstroom.

Indicator 3 wordt nu als volgt gedefinieerd:

- Indicator 3 = netto_instroom_relatief + 1;
- Als netto_instroom_relatief > 1 geldt (en dus indicator 3 > 2), dan wordt indicator 3 op 2 gesteld.

5.12 **_12_indicatoren_prognoses_Panteia.sps**

In deze syntax worden de drie bestanden met de prognose van de vraag (op basis van de ramingen van Panteia m.b.t. de groei van het aantal inschrijvingen), prognose van het aanbod en de regionale dynamiek samengevoegd. Hierna kunnen de eerste drie indicatoren berekend worden en vervolgens de kans op stage.

5.12.1 *Combineer prognose aanbod met prognose vraag*

De volgende twee bestanden worden gecombineerd:

- bestand met prognose van het aanbod van stages en leerbanen (gecreëerd in syntax_8_);
- bestand met prognose van de vraag naar stages en leerbanen (gecreëerd in syntax_10_).

De bestanden worden gekoppeld op basis van crebocode, type leerweg en arbeidsmarktregio.

Vervolgens wordt gecontroleerd of voor alle waarnemingen in deze dataset het kwalificatienummer, kwalificatiedossier, marktsegment en sectorkamer bekend zijn. Voor sommige waarnemingen blijkt het marktsegment te ontbreken. Dit wordt alsnog bepaald (op basis van crebocode).

5.12.2 *Bepaal het aantal BPVOs in het lopende schooljaar*

Er is nog 1 variabele die ontbreekt om de indicatoren te kunnen berekenen: voor indicator 2 moet nog een prognose gemaakt worden van het aanbod in termen van aantal BPVOs. Dit wordt nu berekend, door het aantal BPVOs in het afgelopen schooljaar (bepaald in syntax_5_) te vermenigvuldigen met de groeivoet van BPVOs.

De groeivoet van het aantal BPVOs is in syntax_4_ bepaald, op basis van de resultaten van de leerbedrijvenquête. Voor veel kwalificaties geldt dat deze groeivoet ontbreekt. Dit zijn vooral kwalificaties met relatief weinig studenten; deze kwalificaties komen maar weinig of niet in de leerbedrijvenquête voor, waardoor het niet mogelijk was om de groeivoet te bepalen. Voor deze kwalificaties zijn wel de groeivoeten op hoger aggregatieniveau beschikbaar (per kwalificatiedossier en leerweg, per marktsegment en leerweg, en per leerweg).

De groeivoet van het aantal BPVOs wordt daarom opnieuw berekend, op basis van dezelfde syntax als in syntax_4_, maar op basis van de uitgebreidere dataset.

5.12.3 *Berekening van indicatoren 1, 2 en 3*

Indicatoren 1 en 2 worden nu berekend. Indicator 1 wordt op basis van de volgende formule berekend:

- Indicator_1 = aanbod_jaar_Tplus1/vraag_jaar_Tplus1_crebo (de prognose van het aanbod in het huidige schooljaar (jaar T plus 1), in BPVOs; gedeeld door de prognose van de vraag in het huidige schooljaar, in BPVOs).

Indicator 2 is in eerste instantie als volgt gedefinieerd:



- $Indicator_2_bpvos = BPVOs_jaar_Tplus1/vraag_jaar_Tplus1_crebo$ (de prognose van het aantal BPVOs dat in het huidige schooljaar afgesloten zal worden; gedeeld door de prognose van de vraag in het huidige schooljaar, in BPVOs).

Een risico van deze indicator is dat de omvang van de BPVOs aan de aanbodzijde niet overeen hoeft te komen met de omvang van de BPVOs aan de vraagzijde. Dat zou in een systematische over- of onderschatting van indicator 2 kunnen resulteren. Daarom is er een tweede versie van indicator 2 gedefinieerd, waarbij zowel aan de aanbod- als aan de vraagzijde het aantal BPVOs (per kwalificatie en leerweg) vermenigvuldigd wordt met de gemiddelde verwachte duur van de BPVOs:

- $Indicator_2_uren = (BPVOs_jaar_Tplus1*uren_jaar_aanbod)/(vraag_jaar_Tplus1_crebo_uren)$.

Indicator 3 is al in `syntax_11_` berekend en wordt nu aan de dataset toegevoegd.

5.12.4 *Berekening van de kans op stage*

De kans op stage wordt in eerste instantie op landelijk niveau bepaald. De indicator voor de regionale mobiliteit speelt hier nog geen rol bij.

Op landelijk niveau zijn er drie indicatoren beschikbaar: `indicator_1`, `indicator_2_bpvos` en `indicator_2_uren`. Deze indicatoren zijn zo berekend dat ze aan de volgende voorwaarden voldoen:

- Ze zijn per definitie positief
- Een score van 1,0 staat voor evenwicht tussen vraag en aanbod.
- Een score lager dan 1,0 staat voor een relatief lage kans op een stage of leerbaan (de vraag naar stages is hoger dan het aanbod).
- Een score hoger dan 1,0 staat voor een relatief hoge kans op een stage of leerbaan (het aanbod van stages is hoger dan de vraag).

De kans op stage wordt in eerste instantie berekend als een gewogen gemiddelde van de verschillende indicatoren. Dit gewogen gemiddelde zal dus ook een continue variabele zijn die aan bovenstaande kenmerken voldoet.

De variantie en spreiding van de verschillende indicatoren laat grote onderlinge verschillen zien. Als hier niet voor gecorrigeerd wordt, dan zal de spreiding in het gewogen gemiddelde vooral bepaald worden door de indicator met de hoogste spreiding. Dit is niet wenselijk, daarom worden de eerste drie indicatoren gestandaardiseerd.

De gewichten van de verschillende indicatoren zijn gelijk gehouden aan de gewichten die voor Kans op Stage 2017 zijn bepaald. De kans op stage is daardoor volgens de volgende formule berekend:

$$Kans_op_stage = gewicht * indicator_2_bpvos + (1 - gewicht) * [\alpha * indicator_1 + (1 - \alpha) * indicator_2_uren]$$

waarbij

- $gewicht = 0.75$ (dit is het gewicht van `indicator_2_BPVOs` in de kans op stage);
- $\alpha = 1$ (dit is het relatieve gewicht van `indicator_1` ten opzichte van `indicator_2_uren`).

De keuze voor $\alpha=1$ houdt in dat `indicator_2_uren` niet gebruikt is bij het bepalen van de kans op stage: op landelijk niveau is de kans op stage voor 75% bepaald door `indicator_2_bpvo` en voor 25% door `indicator_1`.



Op regionaal niveau is de kans op stage volgens de volgende formule afgeleid:

$$\text{Kans_op_stage_regionaal} = \text{gewicht_regionaal} * \text{kans_op_stage} + (1 - \text{gewicht_regionaal}) * \text{indicator_3}.$$

Waarbij

- gewicht_regionaal = 0.8 (dezelfde waarde als voor Kans op Stage 2017).

Zowel de nationale als de regionale variant van de kans op stage zijn nu nog continue variabelen: ze kunnen elke positieve waarde aannemen (al liggen ze vooral tussen 0 en 2). De laatste stap is om de kans op stage in te delen in een variabele met vijf antwoordcategorieën:

- "gering": bij een waarde tussen 0 en 0,7;
- "matig": bij een waarde tussen 0,7 en 0,9;
- "voldoende": bij een waarde tussen 0,9 en 1,1;
- "ruim voldoende": bij een waarde tussen 1,1 en 1,3;
- "goed": bij een waarde van 1,3 of meer.

