



Discussion paper

Tijdelijke en vaste contracten

Verschillen tussen de schattingen uit de Polisadministratie en de Enquête beroepsbevolking verklaard?

Bart Bakker
Gerda Gringhuis
Jeffrey Hoogland
Frank van der Linden
John Michiels
Jeroen Pannekoek
Sander Scholtus
Wendy Smits

Inhoudsopgave

1. Inleiding 5

2. Data en methoden 8

- 2.1 De opzet van het onderzoek 8
- 2.2 De gebruikte methoden 10

3. De mogelijke verklaringen 14

4. Verklaring van inconsistenties tussen Polis en EBB 17

- 4.1 De opzet van de Multinomiale Logistische Regressieanalyse 17
- 4.2 De resultaten van de MLR voor alle leeftijdsgroepen samen 19
- 4.3 De resultaten van de MLR voor 15- tot 25-jarigen 21
- 4.4 De resultaten van de MLR voor 25- tot 55-jarigen 23
- 4.5 De resultaten van de MLR voor 55-jarigen en ouder 25

5. De schatting van de onnauwkeurigheden 26

- 5.1 De replicatie van de meetmodellen van 2007 en 2009 26
- 5.2 De verbetering van het meetmodel voor de 25- tot 55-jarigen 31
- 5.3 De uitkomsten van het meetmodel aanschouwelijk gemaakt 35
- 5.4 Het meetmodel voor de 55-jarigen en ouder 37
- 5.5 Het meetmodel voor de 15- tot 25-jarigen 40
- 5.6 De precisering van de ontwikkeling in de onnauwkeurigheden in 2017 en 2018 43

6. Waar hangen de onnauwkeurigheden mee samen? 47

- 6.1 Waar hangen de onnauwkeurigheden in de meting van arbeidscontract in de EBB en de Polis mee samen voor de 25- tot 55-jarigen 48
- 6.2 Waar hangen de onnauwkeurigheden in de meting van arbeidscontract in de EBB en de Polis mee samen voor de 55- tot 75-jarigen 50
- 6.3 Waar hangen de onnauwkeurigheden in de meting van arbeidscontract in de EBB en de Polis mee samen voor de 15- tot 25-jarigen 51

7. Conclusies en discussie 52

- 7.1 De ontwikkeling van de meetmodellen 53
- 7.2 De grootte van de onnauwkeurigheden in de meting van arbeidscontract van werknemers in de EBB en in de Polis 54
- 7.3 De ontwikkeling van de onnauwkeurigheden en de schattingen uit de Polis en de EBB 55
- 7.4 Mogelijke verklaring van de onnauwkeurigheden in de Polis en de EBB 56
- 7.5 Aanbevelingen 58

Referenties 59

Bijlage 1. De variabele soort arbeidscontract 61

Bijlage 2. De procedure voor het schatten van de invloed van covariaten op onnauwkeurigheden in de meting in Polis en EBB 64

Bijlage 3. De variabele softwarecluster 70

Bijlage 4. De overige variabelen 74

Bijlage 5. De MLR voor 15- tot 25-jarigen 76

Bijlage 6. De MLR voor 25- tot 55-jarigen 81

Bijlage 7. De MLR voor 55-jarigen en ouder 86

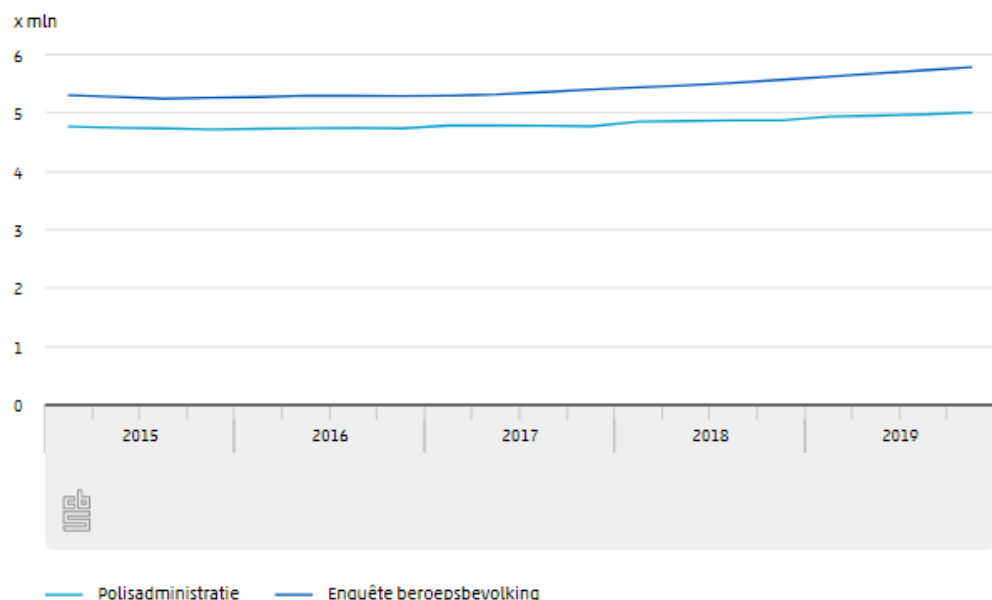
Bijlage 8. De fit van de verschillende HMM-modellen 91

Bijlage 9. Bijdrage van de covariaten aan de onnauwkeurigheden in de meting in de Polis en in de EBB 92

1. Inleiding¹

Het CBS publiceert over de vaste en flexibele arbeidscontracten van werknemers zowel uit de Polis (Polis) als uit de Enquête beroepsbevolking (EBB). De uitkomsten over vast en flexibel werk vanuit de Polis en de EBB laten verschillen zien. Vanuit de Polis wordt gepubliceerd over banen van werknemers. Bij de EBB gaat het om werkzame personen. Personen kunnen meerdere banen hebben. Voor dit onderzoek is gecorrigeerd voor populatieverschillen tussen de beide bronnen, zodat de populaties vergelijkbaar worden. Ook de definitie van vaste en flexibele arbeidsrelaties in beide bronnen is zo gelijk mogelijk gemaakt. De uitkomsten die in dit rapport worden genoemd hebben betrekking op data waarbij de populaties en definities van beide bronnen zijn geharmoniseerd.

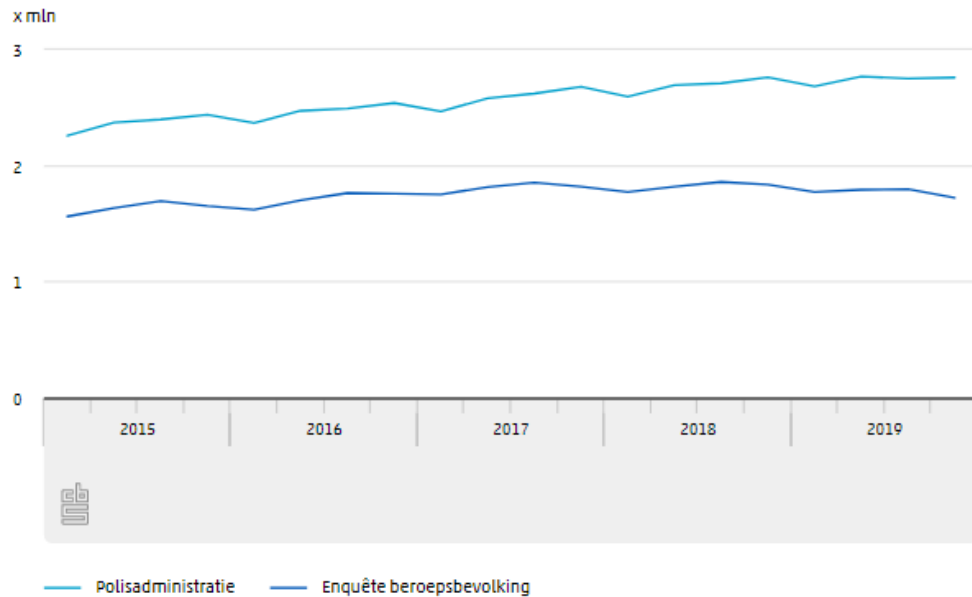
1.1.1 Aantal werknemers met een vaste arbeidsrelatie



Ook als er wordt gecorrigeerd voor populatie- en definitieverschillen verschillen beide bronnen in de schatting van het aantal en aandeel vaste en flexibele werknemers. In 2010, derde kwartaal had 29,8 procent van alle werknemers in de Polis een flexibele baan, in de EBB was dat 21,1 procent. Het gaat dus om een verschil van 8,7 procentpunt. Vanaf 2018 lopen ook de ontwikkeling van de schattingen van het aantal en aandeel vaste en flexibele werknemers uiteen. De EBB laat een sterkere toename zien van het aantal werknemers met een vast arbeidscontract dan de Polis (figuur 1.1.1), en andersom, laat de Polis een sterkere toename zien van het aantal flexibele arbeidscontracten dan de EBB (figuur 1.1.2). In 2018, vierde kwartaal bedroeg de schatting van het aandeel flexibele werknemers in de Polis en EBB respectievelijk 35,7 en 25,2 procent: een verschil van 10,5 procentpunt. Deze cijfers zijn eigen berekeningen na correctie voor de populatieverschillen tussen de EBB en de Polis.

¹ De auteurs danken de personen die commentaar hebben gegeven op eerdere versies van dit discussion paper. Dit discussion paper is gereviewd door Reinoud Stoel (CBS).

1.1.2 Aantal werknemers met een flexibele arbeidsrelatie



Er zijn in het verleden verschillende onderzoeken gedaan naar mogelijke oorzaken van de verschillen (bijv. Bruggink et al., 2008; Mars, 2011; Gringhuis, 2011; Gringhuis et al., 2012). Uit deze onderzoeken kwam naar voren dat de onnauwkeurigheid bij sommige groepen (jongeren, kleine baantjes, bepaalde beroepen) wat groter is dan bij anderen maar dat de verschillen niet aan een enkele groep te wijten zijn.

Een mogelijke verklaring voor de verschillen in de schattingen van het aantal en aandeel vaste en flexibele werknemers is dat beide bronnen onnauwkeurigheden in de meting bevatten. Onnauwkeurigheid in de EBB kan bijvoorbeeld ontstaan doordat respondenten de vragen over het type dienstverband niet goed begrijpen of omdat ze niet goed weten wat voor arbeidscontract ze hebben. In de Polis kun je daarbij denken aan administratieve vertraging die onnauwkeurigheid veroorzaakt (Bakker & Daas, 2012). Pavlopoulos en Vermunt (2015) en Pankowska et al. (2018) hebben schattingen gemaakt van de (on)nauwkeurigheid met behulp van zogenoemde meetmodellen. Daarbij zijn Hidden Markov-modellen (HMM) toegepast op een gekoppeld bestand van de Polis en de EBB uit 2007 (Pavlopoulos & Vermunt, 2015) en 2009 (Pankowska et al., 2018).

Uitgangspunt van de methode is dat er per persoon op meerdere tijdstippen zowel een meting van het type contract in de EBB als in de Polis beschikbaar is: vanuit de Polis in principe elke maand en vanuit de EBB voor maximaal vijf waarnemingen met steeds drie maanden ertussen. Door de informatie uit beide bronnen op meerdere tijdstippen te combineren kan per persoon een schatting worden gemaakt van de meest waarschijnlijke contractvorm. Wanneer het model klopt zijn schattingen over alle personen opgeteld een betere benadering van de werkelijkheid dan de schattingen vanuit de EBB of de Polis afzonderlijk.

In deze modellen wordt verondersteld dat de EBB en de Polis hetzelfde meten, in dit geval of er sprake is van een flexibel of vast arbeidscontract. In de EBB wordt echter niet per se het type contract gemeten. Als er informele afspraken zijn dat een contract vast is, dan kan dat beschouwd worden als een vast contract. Dit zal zeker leiden tot enige vertekening. De veronderstelling voor Nederland lijkt echter redelijk omdat informele afspraken naar

verwachting hier niet zo belangrijk zijn. Er is echter geen mogelijkheid om deze veronderstelling te toetsen. Als in de Polis en de EBB niet precies hetzelfde wordt gemeten is er strikt genomen sprake van sprake van afwijking die dan bestaat uit een systematisch deel ten gevolge van de andere operationele definitie (onzuiverheid) en een willekeurig deel (onnauwkeurigheid). In dit rapport is vanwege de leesbaarheid geen onderscheid gemaakt tussen onzuiverheid en onnauwkeurigheid. Dit is samen 'onnauwkeurigheid in de meting' genoemd.

Voor de onderzoeken van Pavlopoulos en Vermunt (2015) en Pankowska et al. (2018) zijn de 25- tot 55-jarigen geselecteerd die in het eerste kwartaal van respectievelijk 2007 en 2009 voor het eerst deelnamen aan de EBB. De modelschattingen van het aandeel flexibele arbeidscontracten ligt tussen de geobserveerde aandelen in de Polis en de EBB. Het percentage overgangen van tijdelijk naar vast in drie maanden tijd wordt volgens de modeluitkomsten door beide bronnen overschat. Replicaties tot en met 2010 laten zien dat de geschatte onnauwkeurigheden constant blijven in de tijd, tenzij een ingrijpende wijziging in de waarneming plaatsvindt (Pankowska et al., 2018).

De onderzoeksresultaten van Pavlopoulos en Vermunt (2015) en van Pankowska et al. (2018) zijn ondertussen verouderd en kennen enkele beperkingen voor wat betreft de onderzochte populatie. Bovendien is sinds 2009 de kwaliteit toegenomen van de gegevens in de loonaangifte en daarmee in de Polis. Ruim tien jaar na invoering is de loonaangifte veel meer ingeburgerd dan drie jaar na invoering. In de verwerking worden minder correcties uitgevoerd dan tien jaar geleden. Daarnaast hebben UWV en Belastingdienst veel meer poortcontroles ingevoerd die ze pas invoeren als fouten minder vaak gemaakt worden, omdat anders teveel aangiftes tegengehouden zouden worden. Verder is de kwaliteit van de Polisdata mede een zaak van ervaring van de salarisadministrateurs. Ook in de waarneming van de EBB is sinds 2009 veel veranderd. Zo is in meerdere stappen mixed mode waarneming ingevoerd en is de vragenlijst meerdere keren aangepast. Hoe de onnauwkeurigheden nu zijn is daarom onbekend. Inmiddels zijn ook nieuwe inzichten ontstaan over de onderlinge samenhang tussen de onnauwkeurigheden in de tijd en de variabelen die hiermee samenhangen. Bovendien zijn in 2020 voor het eerst gegevens beschikbaar gekomen over de softwarepakketten die bedrijven gebruiken voor hun aangiften.

Daarnaast beperken Pavlopoulos en Vermunt (2015) en Pankowska et al. (2018) zich tot de 25- tot 55-jarigen die in het eerste kwartaal van respectievelijk 2007 en 2009 voor het eerst deelnamen aan de EBB. Dit betekent een beperking van de populatie en betekent dat de verandering van de geschatte onnauwkeurigheden gedurende de onderzoeksperiode niet goed in beeld kan worden gebracht. Daarom is het CBS een nieuw onderzoek gestart voor een recentere periode (2016-2018), met een uitgebreider model, met een bredere populatie en met een hogere resolutie in de tijd (de instroom in de EBB van alle kwartalen van de jaren 2016-2018).

Het onderzoek richt zich op het beantwoorden van de volgende vragen:

1. Hoe groot is de onnauwkeurigheid in de meting van de vaste en flexibele arbeidscontracten van werknemers in de Polisadministratie (Polis) en in de Enquête beroepsbevolking (EBB)?
2. Hoe heeft de onnauwkeurigheid in de meting van de vaste en flexibele arbeidscontracten zich ontwikkeld in 2017 en 2018 en kan daarmee het verder uiteenlopen van de schattingen in de Polis en de EBB worden verklaard?
3. Waardoor worden de (ontwikkelingen in de) onnauwkeurigheden verklaard?

2. Data en methoden

2.1 De opzet van het onderzoek

Om te achterhalen waar de verschillen tussen EBB en Polis vandaan komen zijn de volgende stappen ondernomen:

- Er is een databestand samengesteld waarin de relevante Polis- en EBB-data zijn gekoppeld op individueel niveau voor de jaren 2006-2018. Dit bestand is grotendeels afgeleid uit het Stelsel van Sociaal Statistische Bestanden (Bakker, Van Rooijen en Van Toor, 2014). Speciaal voor het onderzoek zijn op basis van de Polis longitudinale baangegevens samengesteld en aan het bestand toegevoegd, zoals begin- en einddatum van de baan en baanduur. Daarnaast is informatie toegevoegd over de software die wordt gebruikt bij de aangifte van de Polisgegevens door bedrijven. Op dit bestand is een voorbewerking gedaan om de juiste variabelen te creëren. Een belangrijke variabele is daarbij de soort arbeidscontract. In bijlage 1 is gedetailleerd aangegeven hoe deze is samengesteld.
- Voordat Polis en EBB zijn gekoppeld zijn de populaties en de definitie van vaste en flexibele arbeidscontracten in beide bronnen zoveel mogelijk geharmoniseerd. Dat betekent dat in de Polis alleen de hoofdbaan is geselecteerd als een werknemer meer dan een baan heeft op een bepaald moment. De hoofdbaan is de baan met het hoogste loon. Daarnaast zijn de data van werknemers die niet in Nederland wonen verwijderd. Het gaat hierbij voornamelijk om seizoensarbeiders. Verder is de populatie werknemers niet beperkt tot de werknemers met een baan van meer dan 12 uur zoals in de eerder uitgevoerde studies, maar alle werknemers zijn geselecteerd. Daarmee wordt aangesloten bij de huidige Europese richtlijnen voor de afbakening van werknemers.
- De invloed van een groot aantal variabelen op de inconsistenties tussen de Polis en de EBB is bepaald met behulp van Multinomiale Logistische Regressieanalyse (MLR). Tegelijkertijd zijn daarbij de effecten van de tijd gemodelleerd. Daarmee wordt duidelijk welke variabelen de verschillen tussen de meting van de soort arbeidscontract in de Polis en de EBB en daarmee mogelijk ook de onnauwkeurigheden verklaren. Deze kunnen vervolgens weer als input dienen voor de verdere ontwikkeling van meetmodellen.
- Vervolgens is het onderzoek van Pavlopoulos & Vermunt (2015) en Pankowska et al. (2018) gerepliceerd voor de 25- tot 55-jarigen die instromen in de EBB in het eerste kwartaal van de verslagjaren 2007, 2009, 2016, 2017 en 2018. Daarmee wordt duidelijk hoe de onnauwkeurigheden zich sinds 2007 hebben ontwikkeld in de Polis en

- de EBB, gegeven de beperkingen van het model dat in die studies is gebruikt. Daarbij is de nieuwe afbakening van werknemers gebruikt.
- In de volgende stap is gezocht naar verbeteringen van het meetmodel voor de 25- tot 55-jarigen voor de verslagjaren 2016-2018. Daarbij worden enerzijds assumpties van het model verzacht en worden variabelen die in de MLR hoog samenhangen met de soort arbeidscontract en inconsistenties tussen de Polis en de EBB toegevoegd aan het model. Deze variabelen worden in de HMM-analyse covariaten genoemd. Deze verbeteringen in het model moeten meer definitief inzicht bieden in de ontwikkeling van de onnauwkeurigheden in de Polis en de EBB tussen 2016 en 2018. Daarbij wordt een tweetrapsprocedure gevolgd. In de eerste stap worden de variabelen toegevoegd aan het model om de initiële kansen op de arbeidscontracten en de overgangskansen mede te bepalen. In de tweede stap wordt vastgesteld welke van deze variabelen de onnauwkeurigheden in de EBB en de Polis bepalen. In bijlage 2 is deze tweestapsprocedure in detail beschreven. Onderdeel van deze bijlage is een beschrijving van een simulatie op fictieve data waarmee aannemelijk wordt gemaakt dat deze procedure goed werkt. De uitkomsten van deze procedure dienen tot slot voor de verdere toetsing van het model: variabelen die de onnauwkeurigheden níet in belangrijke mate bepalen, zouden naar verwachting ook geen sterk verklarende variabelen mogen zijn in de MLR-analyse op de inconsistenties tussen de Polis en de EBB.
 - De procedure voor het vinden van het best passende model voor de 25- tot 55-jarigen is eveneens toegepast op de andere leeftijdsgroepen, de 15- tot 25-jarigen en de 55- tot 75-jarigen. Ook hierbij is uitgegaan van de gekoppelde Polis- en EBB-data uit de verslagjaren 2016-2018.
 - In de laatste stap zijn de beste modellen toegepast op een uitgebreidere dataset waarin naast de instroom uit het eerste kwartaal van de EBB ook de instroom uit de andere kwartalen wordt meegenomen. Daarmee kan de vraag worden beantwoord of de onnauwkeurigheid in de meting een verklaring vormt voor het verder uiteenlopen van de schattingen uit de EBB en de Polis.

Een belangrijke vraag die vooraf gaat aan het beantwoorden van de onderzoeksvragen is de precieze afleiding van de variabele arbeidscontract. Deze afleiding dient voor beide bronnen zo goed als mogelijk op elkaar te zijn afgestemd. Voor dit onderzoek wordt een variabele 'soort arbeidscontract' met de volgende categorieën gecreëerd: (1) werknemers met een vast arbeidscontract (2) werknemers met een flexibel arbeidscontract en (3) overige. De categorieën (flexibele) werknemers die in de EBB en de Polis-administratie kunnen worden onderscheiden sluiten niet helemaal aan op elkaar. De groep werknemers met een vast arbeidscontract zonder vaste uren is niet te onderscheiden in de Polis-administratie. Deze groep wordt daarom in dit onderzoek in de EBB tot de werknemers met een vast arbeidscontract gerekend. Omgekeerd kunnen in de EBB de Directeuren die ook groootaandeelhouder zijn (DGA's) niet worden onderscheiden omdat DGA's meestal aangeven zelfstandige te zijn. Deze worden daarom in dit onderzoek in de categorie 'Overig' ondergebracht.

Dit onderzoek richt zich op de verschillende schattingen van Polis en EBB van het type contract van werknemers. Een eerste meetfout ontstaat echter al als iemand in één van beide bronnen niet als werknemer geregistreerd staat. Wat betreft de mogelijke oorzaken van onnauwkeurigheden zou je een onderscheid moeten maken tussen deze soorten fouten (1) wel werknemer volgens de ene bron, geen werknemer volgens andere bron (2)

werknemer volgens beide bronnen maar verschillend soort arbeidscontract. Daar is in dit onderzoek geen rekening mee gehouden. De niet-werknemers komen in een restcategorie terecht. In bijlage 1 is de precieze operationalisering van de variabele opgenomen alsmede de afleiding hiervan in de Polis en de EBB. Deze variabele vormt de basis voor de analyse met MLR en de HMM.

2.2 De gebruikte methoden

In het onderzoek wordt van een tweetal methoden gebruik gemaakt. De eerste methode is Multinomiale Logistische Regressie (MLR). Dat wordt gebruikt voor het identificeren van kenmerken die sterk samenhangen met inconsistenties tussen de EBB en de Polis bij de meting van vaste en flexibele contracten. De tweede methode bestaat uit het toepassen van zogenaamde Hidden Markov-Modellen (HMM) waarmee wordt geprobeerd een relatie te leggen tussen de veronderstelde werkelijke contractvorm en de gemeten contractvormen volgens de Polis en volgens de EBB.

MLR-analyses

Ten behoeve van de MLR-analyses wordt op basis van de hiervoor gedefinieerde variabele 'Arbeidscontract' een afhankelijke variabele afgeleid met drie categorieën: (1) 'Het arbeidscontract is in de EBB gemeten als vast en in de Polis als flex', (2) 'Het arbeidscontract is in de EBB gemeten als flex en in de Polis als vast', en (3) 'Het arbeidscontract gemeten in de EBB en in de Polis is gelijk aan elkaar'. De eerste twee categorieën beschrijven inconsistenties in de metingen van de EBB en de Polis. De categorie 'Overig' wordt hierbij buiten beschouwing gelaten. Om meer inzicht te verkrijgen in mogelijke verklaringen van de verschillen tussen de Polis en de EBB zijn enkele statistische analyses uitgevoerd waarin zoveel mogelijk is vastgesteld welke mogelijke verklaringen hout snijden. Om te voorkomen dat de resultaten verstoord worden door onnauwkeurigheden die tussen golven gecorreleerd zijn, is per persoon alleen de eerste waarneming in de EBB gebruikt.

Om de mogelijke verklaringen te toetsen is een MLR met de boven beschreven driewaardige doelvariabele gebruikt. In hoofdstuk 3 wordt beargumenteerd welke variabelen zijn gekozen die als onafhankelijke variabele zijn opgenomen. Als bijvoorbeeld een mogelijke verklaring is dat mensen met kleine banen vaker niet precies weten wat hun arbeidscontract is, wordt het aantal gewerkte uren als variabele opgenomen. Op grond van de MLR worden uitspraken gedaan als 'bij mensen met een baan van minder dan 10 uur per week zijn er drie keer zoveel inconsistenties als bij mensen die 32 uur of meer werken'. Dat is dus niet een verklaring in de zin dat er oorzaken van onnauwkeurigheden worden aangewezen. Wel worden aanwijzingen verkregen waar verder gezocht moet worden naar de oorzaken. In de bijlagen 1, 3 en 4 is een gedetailleerde beschrijving gegeven van de gebruikte variabelen.

De analyses zijn steeds gestart met het vaststellen van het belang van ieder van de variabelen. Dit wordt gekwantificeerd door het verschil tussen twee modellen te beoordelen: het model met alle variabelen en het model met één variabele, de desbetreffende variabele, minder. De maat voor dit verschil is de likelihood-ratio toetsingsgrootte. Als dit verschil niet significant is, dan kan die variabele zonder meer uit het model weggelaten worden. Is het verschil wel significant dan draagt die variabele

significant bij aan de verklaring van de doelvariabele. Die significanties zeggen echter niet alles gezien het grote aantal waarnemingen in onze dataset. Grotere waarden van deze grootte corresponderen met grotere verschillen tussen de geschatte kansen op de categorieën van doelvariabele en daarmee op een groter effect van de desbetreffende variabele. Daarmee worden de variabelen wel geordend naar het belang die deze hebben in de verklaring van de verschillen tussen de Polis en de EBB.

Omdat de effecten van de verklarende variabelen voor verschillende leeftijdscategorieën sterk kunnen verschillen, en om aan te sluiten bij eerder onderzoek van Pavlopoulos & Vermunt (2015) en Pankowska et al. (2018), worden aparte modellen geschat voor ieder van de drie leeftijdscategorieën: '15 tot 25 jaar', '25 tot 55 jaar' en '55 jaar en ouder'. Alleen de verklarende variabelen met significante effecten zijn opgenomen. Deze modellen bevatten in ieder geval de variabele 'Kwartaalnummer' die de verschillen in niveaus per kwartaal weergeeft. Hieraan zijn de overige verklarende variabelen toegevoegd voor zover ze significante effecten hebben. Hiervoor wordt een stapsgewijze modelselectie-procedure toegepast.

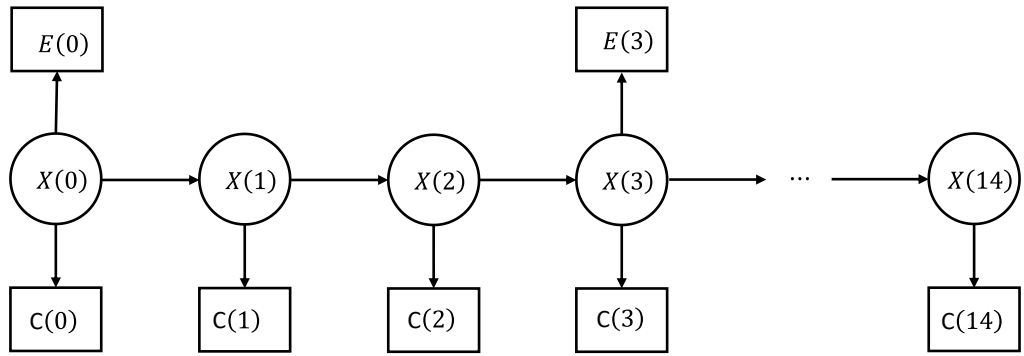
Bij de interpretatie van de parameters van de MLR is het onvoldoende om alleen te kijken naar de uitkomsten van de significantietoetsen. Het aantal waarnemingen is dermate groot dat vrijwel alle parameters significant zijn ook als de effecten klein zijn. Zo is een effect op de kansverhouding (zie hoofdstuk 4 voor een verdere toelichting) van 1,01 wellicht significant volgens de gebruikelijke significantietoets, maar inhoudelijk betekent het dat er vrijwel geen effect is: de kansverhouding is maar 1,01 maal zo groot als de kansverhouding voor de referentiecategorie. Daarom moet gekeken worden naar de relevantie van de parameters. Dit geldt ook voor de mogelijke interactie-effecten. Zoeken naar interactie-effecten op basis van een significantietoets is in dit geval tamelijk zinloos. De kans is groot dat er veel worden gevonden, maar dat die inhoudelijk niet relevant zijn. Bovendien is het altijd erg lastig om interactie-effecten goed te interpreteren. Daarom is er voor dit onderzoek vanaf gezien.

HMM-analyses

Meetmodel

Om de omvang en de ontwikkeling van de onnauwkeurigheden te schatten is een Latente Klasse Model gebruikt en in het bijzonder een zogenaamd Hidden Markov-Model (HMM). In dit type model wordt geprobeerd de samenhangen die er zijn tussen enerzijds de niet-direct meetbare (latente) werkelijke contractvorm en anderzijds de in de EBB en in de Polis gemeten contractvormen nader te beschrijven. De grove basisstructuur van het model is in onderstaande figuur 2.2.1 geïllustreerd, waarin de belangrijkste relaties zijn weergegeven. De variabelen E en C beschrijven de gemeten contractvorm van een onderzoekspersoon volgens gekoppelde gegevens van respectievelijk de EBB en de Polis-administratie gedurende de looptijd van het EBB-panel. Deze variabelen worden bepaald door de werkelijke contractvorm: de variabele X duidt hierbij op de niet-gemeten latente variabele en vertegenwoordigt de werkelijke contractvorm. De betekenis van de latente variabele wordt gevoed door de overeenkomst in de metingen van de EBB en de Polis. De ontwikkeling van deze variabele in de tijd wordt gemodelleerd met een Markovketen. Omdat de waarnemingen van Polis en EBB zich beperken tot de looptijd van het EBB-panel zijn er 15 maandelijkse meetpunten (van $t=0$ tot $t=14$). Bij de EBB wordt om de drie maanden een meting afgenomen, tot een maximum van vijf waarnemingen.

2.2.1 Het HMM-model



Het uiteindelijke gekozen model bevat aanvullende relaties die niet in de figuur zijn opgenomen. Zo ontbreken nog de covariaten die de initiële kansverdeling op $t=0$ en de overgangskansen in de Markovketen voor X mede bepalen. Ook is er een extra kans π_k opgenomen om te corrigeren voor de effecten van niet-waargenomen covariaten, door middel van meerdere parallelle Markovketens (k indiceert individuele ketens), en zijn de mogelijke correlaties tussen meetfouten in een eerdere meting (in de Polis en in de EBB) expliciet in het model verwerkt (mede afhankelijk van covariaten als bijvoorbeeld wel of niet gewerkt in dezelfde baan als bij de vorige meting). Dit leidt tot de volgende uitdrukking voor de kans op de waarnemingen van een afzonderlijk panelrecord i (de vetgedrukte symbolen $\mathbf{E}_i$ en $\mathbf{C}_i$ duiden op vectorgrootheden van reeksen van maandwaarnemingen van respectievelijk EBB en Polis-administratie van persoon i ; $\mathbf{Z}_i$ vertegenwoordigt de tijd-vector met covariaten)²:

$$P(\mathbf{E}_i, \mathbf{C}_i | \mathbf{Z}_i) = \sum_{k=1}^K \sum_{x_0=1}^3 \sum_{x_1=1}^3 \dots \sum_{x_T=1}^3 \pi_k \cdot P(X_{i0} = x_0 | Z_{i0}, k) \cdot \prod_{t=1}^T P(X_{it} = x_t | X_{i(t-1)} = x_{(t-1)}, Z_{it}, k) \cdot P(E_{i0} = e_0 | X_{i0} = x_0, Z_{i0})^{\delta_{i0}} \cdot P(C_{i0} = c_0 | X_{i0} = x_0, Z_{i0}) \cdot \prod_{t=1}^T P(E_{it} = e_t | X_{it} = x_t, X_{i(t-3)} = x_{(t-3)}, E_{i(t-3)} = e_{(t-3)}, Z_{it})^{\delta_{it}} \cdot \prod_{t=1}^T P(C_{it} = c_t | X_{it} = x_t, X_{i(t-1)} = x_{(t-1)}, C_{i(t-1)} = c_{(t-1)}, Z_{it})$$

Het beschreven model vormt een uitbreiding op de eerdere varianten van Pavlopoulos en Vermunt (2015) en Pankowska et al. (2018). Maximum likelihood-schattingen van de parameters in het model, die verband houden met de specificatie van de diverse kansgrootheden in termen van de covariaten, worden berekend via een variant van het Expectation-Maximization (EM)-algoritme (Dempster, Laird and Rubin 1977). Voor de *expectation* stap van dit algoritme wordt gebruik gemaakt van een recursieve methode die specifiek is voor dit type modellen (Baum, Petrie, Soules and Weiss, 1970; Vermunt, Tran, Magidson, 2008). Via het EM-algoritme wordt ook rekening gehouden met ontbrekende waarnemingen in $\mathbf{E}_i$ voor de maanden waarin persoon i niet heeft deelgenomen aan een EBB-peiling. De analyses zijn uitgevoerd met het pakket Latent Gold, versie 5.1.

² De Dirac-delta δ_{it} geeft aan of in de betreffende maand t de onderzoekspersoon i aan een EBB-peiling heeft deelgenomen (zo ja dan is de waarde =1, anders =0).

Uitbreiding HMM-modellering ten opzichte van eerdere studies

Eerst worden de onderzoeken van Pavlopoulos en Vermunt (2015) en Pankowska et al. (2018) gerepliceerd, dat wil zeggen dat de door hen gebruikte modellen gefit zijn op de door hen afgebakende populatie van 25- tot 55-jarigen voor 2007, 2009, 2016, 2017 en 2018. In de eerdere studies is de populatie verder ingeperkt door alleen banen van 12 uur of meer mee te tellen als werkzame personen. Dit onderzoek analyseert de verslagjaren 2007, 2009, 2016, 2017 en 2018 met de nieuwe definitie waarin alle banen meetellen. Door de schattingen van de onnauwkeurigheden uit 2007 en 2009 te vergelijken met 2016, 2017 en 2018, kan de vraag worden beantwoord hoe de onnauwkeurigheden zich hebben ontwikkeld. En passant geeft dit inzicht in de verschillen tussen de oude en de nieuwe definitie.

De modellen zoals die in eerdere studies zijn gebruikt kennen belangrijke beperkingen. Zo is aangenomen dat de onnauwkeurigheden tussen de verschillende waarnemingen van de EBB bij dezelfde persoon niet onderling samenhangen; dat wil zeggen:

$$P(E_{it} = e_t | X_{it} = x_t, X_{i(t-3)} = x_{(t-3)}, E_{i(t-3)} = e_{(t-3)}, Z_{it}) = P(E_{it} = e_t | X_{it} = x_t, Z_{it}).$$

Die aanname is niet plausibel en die zal daarom niet langer worden gedaan. Verder is eerder met een beperkt aantal covariaten gewerkt waarbij die covariaten zijn gekozen die gemakkelijk beschikbaar waren. In de MLR-analyse is nagegaan welke variabelen van belang zijn voor de verklaring van de verschillen tussen Polis en EBB. Dat levert een aantal goede kandidaten op om de modellen te verbeteren. Hierbij wordt per covariaat apart getoetst of het model verbeterd wordt door een effect toe te voegen op de initiële kansen $P(X_{i0} = x_0 | Z_{i0}, k)$ en op de overgangskansen $P(X_{it} = x_t | X_{i(t-1)} = x_{(t-1)}, Z_{it}, k)$.

Selectie en toetsing van HMM-modellen

Of uitbreidingen van het Hidden Markov-model met extra covariaten en relaties tussen reeds eerder opgenomen variabelen een verbetering betekenen wordt in dit onderzoek bepaald met behulp van de Bayesian Information Criterium (BIC). Dit is een voorbeeld van een fitmaat op basis van de likelihood-functie waarbij rekening wordt gehouden met de complexiteit van het model (Schwarz, 1978). Naarmate de BIC kleiner is, past het model beter bij de gegevens.

De toetsing van de HMM-modellen in dit onderzoek kent een technische en inhoudelijke kant. Bij de inhoudelijke toetsing gaat het om de vraag of de gevonden modelresultaten aansluiten bij verwachtingen ten aanzien van de oorzaken van de meetfouten in EBB en Polis. Deze verwachtingen zijn gebaseerd op resultaten van eerder onderzoek en ervaringen met de implementatie van het onderzoeksdesign bij de twee statistiekbronnen. In hoofdstuk 3 wordt hierop verder ingegaan. De uitkomsten van de model-analyses sluiten in het ideale geval aan bij de geformuleerde verwachtingen. Als dat niet het geval is, zou een andere alternatieve en plausibele verklaring gepresenteerd moeten kunnen worden. Een voorbeeld van een dergelijke verwachting is dat een onnauwkeurige opgave van de contractsoort door een EBB-respondent waarschijnlijk in een volgende peiling herhaald wordt, behalve als de werkkring wijzigt. Verbeterd de fitmaat na toevoegen van deze correlatie dan is dit op te vatten als een verbetering van de mate waarin het model bij de gegevens past.

Bij de technische toetsing van de modellen speelt het volgende een rol:

- Het model toegepast op andere jaargangen dan waarop het model is ontwikkeld moet ook goed bij de gegevens passen.

- Kleine veranderingen in de parameters van het model mogen niet tot heel andere resultaten leiden. Als dat wel het geval is, dan is het model niet robuust.
- Uitbreidingen van het model die een geringe impact hebben op de fitmaat mogen niet tot aanzienlijke veranderingen in de modeluitkomsten leiden. Is dit laatste wel het geval dan is moeilijk uit te maken welke van de modellen de 'werkelijkheid' het best weergeeft.
- Met een geschat HMM-model is het mogelijk om voorspelde scores op de latente klassen in de Markovketen te simuleren. Met behulp van deze simulaties kan inzicht worden verkregen in de relatie tussen onnauwkeurigheden in de Polis en de EBB en de covariaten die in de HMM-modellen worden gebruikt, volgens de aanpak die beschreven is in bijlage 2. Covariaten die bepalend zijn voor deze meetfouten zouden naar verwachting ook een belangrijke verklarende rol moeten spelen bij de MLR-analyses op de inconsistenties tussen de Polis en de EBB.

In bijlage 2 is een kleine simulatiestudie beschreven waarbij de bruikbaarheid van de HMM-aanpak is onderzocht voor een vereenvoudigde versie van het model. Als het model een correcte beschrijving geeft van de totstandkoming van de waargenomen arbeidscontracten uit de Polis en de EBB is de conclusie:

- Uit de voorspelde scores op de latente klassen volgt een goede benadering voor de werkelijke percentages mensen met een vast, flexibel of overig arbeidscontract.
- De simulatie-aanpak uit bijlage 2 geeft een goede indruk van de effecten van covariaten op de aanwezigheid van onnauwkeurigheden in de waargenomen arbeidscontracten in de Polis en de EBB.

3. De mogelijke verklaringen

Er zijn veel verklaringen mogelijk die in een nog onbekende combinatie verantwoordelijk zijn voor de verschillen tussen Polis en EBB. Deze zijn geïnventariseerd door raadpleging van literatuur, in een bijeenkomst met deskundigen en in afzonderlijke gesprekken met deskundigen³.

De *vraagstelling in de EBB* is tot 2013 enkele keren gewijzigd. Dat heeft mogelijk consequenties gehad voor het niveau en de ontwikkeling van het aandeel en aantal zoals gemeten in de EBB. Dat kan overigens geen verklaring zijn voor de uiteenlopende ontwikkelingen tussen de Polis en de EBB van na die datum. Daarnaast zijn er veranderingen geweest in de waarneming van de Polis door *verandering van wetgeving of administratieve praktijk*. Een voorbeeld van het laatste is de behandeling van werknemers door een groot bedrijf die systematisch de verkeerde code kregen toegekend. Deze praktijk is pas veranderd toen de Belastingdienst hen daarop wees. Deze bekende

³ Dit waren Nico van Baren (ADP Nederland), Marilijn Labree (ADP Nederland), Dik van Leeuwen (ADP Nederland), Marianne Meijer (CAOP), Christine van Oorschot (zelfstandig adviseur), Dimitris Pavlopoulos (VU), Mark Hartog van Banda, Sonja Kamphuis-Van Veen, Martijn Souren, Rik van der Vliet en Emmy Willems-Pijnenborg (allen CBS)

momenten deelt 2006-2018 op in zes periodes die verschillen qua duur. Dit leidt tot de variabele 'Periode' met zes categorieën.

(1) 1-1-2006 tot 1-1-2010 invoeringstraject van de Polis.

(2) Per 1-1-2010 Invoering van Computer-Assisted Telephone Interviewing (CATI) bij 1e peiling in de EBB, EBB-vraagstelling iedere peiling opnieuw gevraagd met uitzondering van de categorie 'Tijdelijk met uitzicht op vast', wijziging in banenafbakening.

(3) Per 1-1-2012 In de EBB wordt ook voor de categorie 'Tijdelijk met uitzicht op vast' opnieuw gevraagd.

(4) 1-1- 2015 Aanpassing 12 uur naar 1 uur grens van banen in de EBB.

(5) 1-1-2017 Met ingang van deze datum is in de Loonaangifte een extra code voor payrollong toegevoegd. Vanaf dat moment is het pas mogelijk om payrollers in de Polis apart te onderscheiden. Uit eerdere analyses is gebleken dat het grootste deel voorheen als uitzendkracht geregistreerd werd.

(6) Vanaf 1 januari 2018. Bij een groot bedrijf werden tot deze datum vrijwel alle werknemers met een flexibel arbeidscontract geregistreerd, daarna volgens het werkelijke arbeidscontract.

Er wordt een of meer significante effecten van deze categorieën verwacht. Idealiter heeft de variabele 'Kwartaalnummer' na toevoegen van 'Periode' geen categorieën meer met een significant effect: er zijn geen verschillen tussen de effecten van kwartalen binnen de periodes, maar er zijn wel verschillende effecten per periode.

De *vraagstelling in de EBB* is niet duidelijk genoeg. Door de ontwikkeling van het aantal verschillende soorten contracten is dit in toenemende mate problematisch. Dit geldt natuurlijk in de eerste plaats in die bevolkingscategorieën en branches waarin veel flexibele contracten voorkomen. Het kan ook zijn dat de vraagstelling voor sommige bevolkingscategorieën onduidelijk is waardoor het aantal inconsistenties in die categorieën groot is. Mogelijk dat vooral laag opgeleiden, mensen met een migratieachtergrond, jongeren en mensen met kleine of kortdurende baantjes de vragen minder vaak goed begrijpen (zie o.a. Meertens, Krieg & Cremers, 2018). Er wordt daarom verwacht dat opleidingsniveau, migratieachtergrond, leeftijd, aantal contracturen en baanduur een invloed hebben. Dit wordt waarschijnlijk versterkt als de vragen door een ander persoon in het huishouden worden beantwoord. Daarom wordt verwacht dat proxy interviewen het aantal inconsistenties doet toenemen. Er is echter ook een alternatieve verklaring mogelijk voor de effecten van deze variabelen. Het probleem hier kan ook zijn dat de vraagstelling wel duidelijk genoeg is maar dat de situatie waarin iemand zich bevindt onduidelijk is. Dus mensen begrijpen de vraag wel maar weten niet zeker of hun baan vast of flexibel is. Bijvoorbeeld omdat ze het contract niet goed hebben gelezen, omdat de werkgever onduidelijke toezeggingen doet, of omdat de werkgever zich niet aan de afspraken houdt.

Uit eerder onderzoek dat niet is gerapporteerd is gebleken dat mensen die voor het laatst responderen in de EBB, en vaak eerder uitvallen dan de vijfde golf, de grootste mate van inconsistentie lieten zien.⁴ Daarom worden meer inconsistenties verwacht bij *voortijdig uitvallen* en dat EBB-responderen die na de eerste golf besluiten te stoppen, al minder gemotiveerd waren om nauwkeurige antwoorden te geven.

⁴ Persoonlijke noot van Dimitris Pavlopoulos, gebaseerd op eerder onderzoek uitgevoerd door Frank van de Pol.

Bij de EBB is de *waarneemmodus* in de loop der tijd gewijzigd. Aanvankelijk werd in de eerste golf waargenomen via Computer-Assisted Personal Interviewing (CAPI) en de tweede en latere golf via CATI. Vanaf 2012 is mixed-mode-waarneming ingevoerd met Computer-Assisted Web Interviewing (CAWI) kwantitatief als steeds belangrijker mode. Het is bekend dat de waarneemmodus een effect kan hebben op de meting van variabelen (zie bijv. Klausch, Schouten & Hox, 2015). Dit kan komen doordat de modus direct effect heeft op de onnauwkeurigheid in de meting van het arbeidscontract. Het kan ook komen doordat bij de EBB de waarneemmodus niet willekeurig wordt bepaald maar er sprake is van een selectie-effect. Als de mix in de loop der tijd wijzigt kan dit gevolgen hebben voor de ontwikkeling. Daarom wordt een effect van de waarneemmodus verwacht. Welke richting dat effect heeft, kan op basis van de huidige kennis niet op voorhand worden bepaald.

Een foutieve registratie van de soort arbeidscontract bij de bedrijven had tot 2020 meestal geen directe gevolgen voor de betrokkenen. Daarnaast is voor werkgevers mogelijk *de betekenis van het soort arbeidscontract* in de Polis onduidelijk. Dat komt door de toename van het aantal verschillende soorten contracten. Dit geldt natuurlijk in de eerste plaats voor die branches waarin veel tijdelijke contracten en kortdurende arbeidsrelaties voorkomen. Dit kan consequenties hebben voor de schatting van het niveau en de ontwikkeling op basis van de Polis. Hierdoor is het mogelijk dat er fouten aanwezig zijn in de registratie van het arbeidscontract. Daarom wordt verwacht dat deze fouten vaker worden gemaakt in branches waarin veel tijdelijke contracten voorkomen. Andersom geldt ook dat de aard van het arbeidscontract voor werknemers onduidelijk zal zijn. De onduidelijkheid van werkgevers kan die voor de werknemers versterken. Dat zal dan vooral blijken door onnauwkeurigheden in de EBB.

Er wordt verwacht dat er effecten zijn van *de economische activiteit van het bedrijf* (Standaard BedrijfsIndeling, 'SBI-groep') op inconsistenties tussen de Polis en de EBB omdat bij proefanalyses voor het verslagjaar 2016 is gebleken dat er binnen de economische activiteiten met relatief veel jongeren een aanvullend effect is. Vraag is of dit nog steeds het geval is als alle verklarende variabelen zijn toegevoegd. Als dit zo is dan is een mogelijke verklaring dat er een verschillende perceptie is van het arbeidscontract, afhankelijk van de economische activiteit waarin men werkt. Dit effect kan ook worden veroorzaakt door een verklarende variabele die gerelateerd is aan de 'SBI-groep' maar niet is opgenomen in het model.

Bedrijven leveren de informatie aan via standaard *softwarepakketten*. Het is wellicht mogelijk dat er systematische fouten in deze softwarepakketten aanwezig zijn die leiden tot een verkeerde registratie van het arbeidscontract. Ook de defaultsettings kunnen een rol spelen. Omdat een werkgever geen belang had bij de juiste registratie van de variabele, kan het voorkomen dat de defaultsettings ten onrechte niet worden gewijzigd. Informatie over de defaultsettings is helaas niet beschikbaar. Wellicht dat 'SBI-groep' een deel van het mogelijke effect verklaart. Het is immers aannemelijk dat binnen een economische activiteit relatief vaak gebruik wordt gemaakt van dezelfde softwarepakketten.

Er wordt verwacht dat de *bedrijfsgrootte* van invloed is op de kwaliteit waarmee de informatie die nodig is voor de Polis wordt ingevuld. Bij kleine bedrijven is dat een taak die of wordt uitbesteed aan een gespecialiseerd kantoor of door iemand van het bedrijf erbij wordt gedaan. In het laatste geval wordt verwacht dat dat een negatieve invloed heeft op

de kwaliteit. Voor de EBB wordt verwacht dat bedrijfsgrootte van invloed is op de mate waarin informele afspraken gemaakt worden over het arbeidscontract. In grotere bedrijven wordt verwacht dat dat formeler wordt geregeld dan in kleinere bedrijven.

De *conjunctuur* en *werkgelegenheidsstructuur* kunnen invloed hebben op de perceptie van vaste en flexibele arbeidscontracten. Laat de conjunctuur een opgaande lijn zien en is de arbeidsmarkt krap, dan wordt een tijdelijk contract wellicht sneller gezien als een contract dat toch wel omgezet wordt in een vast arbeidscontract. In de EBB en in de Polis kan dit zich vertalen in inconsistenties. De hypothese is dat mensen bij een positieve verwachting van het economisch klimaat voor het komende jaar eerder van mening zijn dat ze een vaste baan hebben, ook als dit formeel niet zo is en er ook geen informele afspraken over gemaakt zijn. Een positieve verwachting voor het economisch klimaat heeft dan een significante toename van de kans op EBB Vast - Polis flex tot gevolg. Het kan ook zijn dat er bij een positieve economische ontwikkeling meer informele afspraken worden gemaakt. Ook dan zijn er meer inconsistenties.

De mensen die volgens de EBB als *zelfstandige* werkzaam zijn, kunnen in de praktijk voor een of meerdere opdrachten werken via een intermediair. In de Polis staan ze dan geregistreerd als uitzendkracht. Als dit toeneemt, kan dit mede de verschillende ontwikkeling van Polis en EBB verklaren.

Enkele van deze oorzaken zijn eerder bekeken door een CBS-werkgroep. Daarbij zijn de mogelijke oorzaken steeds afzonderlijk onderzocht, maar niet in hun onderlinge samenhang. Dat kan betekenen dat de oorzaken door compensatie-effecten niet zichtbaar zijn geworden. Door rekening te houden met de onderlinge samenhangen, kan dit worden voorkomen. Bovendien worden meer mogelijke oorzaken onderzocht. Daarmee wordt een completer beeld gegeven van de onnauwkeurigheden en hun oorzaken.

4. Verklaring van inconsistenties tussen Polis en EBB

4.1 De opzet van de Multinomiale Logistische Regressieanalyse

In hoofdstuk 3 zijn mogelijke verklaringen gegeven voor de onnauwkeurigheden in de meting van het type arbeidscontract in de Polis en de EBB. Die worden in deze paragraaf getoetst door middel van multinomiale logistische regressieanalyse. Daarmee worden bovendien de kandidaat covariaten geselecteerd om in de meetmodellen op te nemen. Samen met de nog vast te stellen samenhang tussen de covariaten en de onnauwkeurigheden in de meting zoals die in de meetmodellen wordt uitgerekend kan dit dienen ter verklaring van de verschillen tussen de EBB en de Polis. Eerst worden de resultaten gepresenteerd voor alle groepen tezamen en vervolgens voor de 15- tot 25-jarigen, dan de 25- tot 55-jarigen en de 55-jarigen en ouder.

De data voor de MLR hebben betrekking op het tijdvak 2010 t/m 2018. De eerste periode van 2006-2010 is niet in de analyses opgenomen omdat de variabele 'Softwarecluster' voor die periode niet beschikbaar is. Het tijdvak 2010-2018 kan opgedeeld worden in 36 kwartalen. Die kwartalen kunnen weer verdeeld worden in vijf perioden. De verklarende variabelen voor deze analyse zijn de volgende dertien variabelen (zie ook bijlagen 1, 3 en 4): (1) 'Kwartaalnummer' (2) 'Periode' (3) 'Leeftijdscategorie' (4) 'Geslacht' (5) 'Opleidingsniveau' (6) 'Migratieachtergrond' (7) 'Proxywaarneming' (8) 'Waarneemmodus' (9) 'Laatste respons' (10) 'Baanduur' (11) 'Contracturen' (12) 'SBI-groep' (13) 'Softwarecluster'.

De doelvariabele *Arbeidscontract* voor de analyse betreft de inconsistente waarneming in het arbeidscontract tussen EBB en Polis. Deze variabele heeft drie categorieën: (1) 'EBB vast, Polis flex', (2) 'EBB flex, Polis vast', en (3) 'EBB en Polis gelijk'. De eerste twee categorieën zijn inconsistenties. Voor deze driewaardige doelvariabele wordt een MLR geschat met de bovenstaande verklarende variabelen. Arbeidsrelaties in de categorie 'Overig' zijn buiten beschouwing gelaten in de MLR-analyses.

De variabele *Periode* is een samenvoeging van de kwartalen in vijf categorieën. Als de variabele *Kwartaalnummer* al in het model is opgenomen kan 'Periode' niets toevoegen. Wel is beoordeeld of 'Kwartaalnummer' te vervangen is door 'Periode'. Het blijkt dat voor alle leeftijdsgroepen 'Periode' een significant betere modelfit geeft dan helemaal geen informatie over de tijd. De variabele 'Kwartaalnummer' verklaart echter steeds veel meer dan 'Periode'. De effecten van veel kwartalen op de inconsistenties zijn gelijk, er is slechts af en toe een overgang. Die overgangen komen echter niet overeen met de vooraf gedefinieerde perioden. Daarom is in het vervolg alleen de variabele 'Kwartaalnummer' opgenomen en is 'Periode' buiten beschouwing gelaten.

Nadat de grootte van de globale effecten van iedere variabele hiermee in kaart gebracht is, is aan de hand van de geschatte parameters voor de categorieën van de variabelen beoordeeld wat de effecten op de inconsistenties zijn. De modellen bevatten een zeer groot aantal parameters en voor de analyses van ieder van de drie leeftijdsgroepen zijn alleen de conclusies samengevat. De tabellen met alle parameters zijn in bijlagen 5, 6 en 7 opgenomen. De relevantie van de parameters is niet goed te beoordelen op grond van de gebruikelijke significantietoets. Door het grote aantal records kunnen kleine effecten al significant zijn. Daarom is de omvang van de parameters beoordeeld voor de interpretatie. De parameters in een (multinomiaal) logit-model geven additieve effecten op de 'logit' weer of multiplicatieve effecten op de kansverhoudingen (zie de toelichting hieronder). Bij een effect op een kansverhouding groter dan 2,0 of lager dan 0,5 heeft deze twee sterren. De betreffende categorie van een verklarende variabele heeft dan een heel groot effect bij de MLR. Een effect op de kansverhouding tussen 1,25 en 2,0 of tussen 0,5 en 0,8 heeft een ster gekregen. Hier zijn de effecten groot. De effecten zijn alleen als ze groot of heel groot en significant zijn ($P < 0,01$) in de lijst van bevindingen opgenomen. De effecten zijn altijd ten opzichte van de voor de desbetreffende variabele gekozen referentiecategorie. Zie daarvoor ook bijlage 4.

Het begrip 'kansverhouding' kan in de context van deze analyse als volgt toegelicht worden. Stel p_1 is de kans op de inconsistentie 'EBB vast, Polis flex', p_2 is de kans op de inconsistentie 'EBB flex, Polis vast', en p_3 is de kans op overeenstemming: 'EBB en Polis gelijk'. Omdat 'EBB en Polis gelijk' gekozen is als referentiecategorie, zijn de

kansverhoudingen voor de beide inconsistenties respectievelijk $p1 / p3$ en $p2 / p3$. Deze kansverhoudingen zeggen hoeveel groter (of kleiner) de kans op een inconsistentie is dan de kans op overeenstemming. De parameters zeggen hoeveel effect een categorie van een verklarende variabele heeft op één van deze kansverhoudingen. Neem bijvoorbeeld de categorie 'EBB vast, Polis flex' bij de jongeren. Jongeren zijn onderverdeeld in twee leeftijdsklassen, namelijk '15 tot 19 jaar' en '19 tot 25 jaar'. De referentiecategorie bij variabele 'Leeftijdsklasse' is '19 tot 25 jaar'. Als de kansverhouding bij 'EBB vast, Polis flex' significant verschilt bij 15- tot 19-jarigen dan is de kansverhouding significant verschillend van de kansverhouding van de 19-25-jarigen. Er geldt dat deze kansverhouding een factor 0,874 kleiner is bij 15- tot 19-jarigen, zie de kolom 'Exp(B)' in bijlage 5. Dit effect is significant, want de overschrijdingskans is 0,000, zie kolom 'Sig.' in bijlage 5.

4.2 De resultaten van de MLR voor alle leeftijdsgroepen samen

De analyse van de effecten van de verklarende variabelen op de inconsistenties is opgesplitst in drie aparte analyses voor drie leeftijdscategorieën: '15 tot 25 jaar' (N= 84877), '25 tot 55 jaar' (N= 343056) en '55 jaar en ouder' (N=98103). Er wordt echter begonnen met een beschrijving van de omvang van de effecten voor alle drie de leeftijdscategorieën samen (N=526036). Daarna volgen de meer gedetailleerde analyses voor ieder van de drie leeftijdscategorieën afzonderlijk. Voor de analyse met alle drie de leeftijdscategorieën samen is het model geschat met de hoofdeffecten van alle verklarende variabelen. In tabel 4.2.1 zijn de likelihood-ratio toetsingsgrootheden (chi-kwadraatwaarde) voor het weglaten van ieder van de variabelen weergegeven. Voor iedere variabele blijkt dat het model *zonder* die variabele significant verschilt van het model *met* die variabele. Weglaten van één van de variabelen leidt dus in alle gevallen tot een model dat slechter bij de data past. De mate waarin dit het geval is wordt weergegeven door de waarde van de toetsingsgrootheid (chi-kwadraatwaarde). De variabelen in tabel 4.2.1 staan op volgorde van de grootte van de chi-kwadraatwaarde.

De eerste vijf variabelen hebben veel grotere effecten dan de overige variabelen. Het effect van Softwareleverancier is daarbij duidelijk het grootst, gevolgd door leeftijdsklasse. Bij cluster 3 van Softwareleverancier is de kansverhouding bij 'EBB vast, Polis flex' 4,7 keer zo groot als bij cluster 1. Bij cluster 4 is deze kansverhouding 2,6 keer zo groot als bij cluster 1. Bij cluster 2 is de kansverhouding bij 'EBB flex, Polis vast' 3,0 keer groter dan bij cluster 1. Bij 40 tot 65 jaar is de kansverhouding bij 'EBB vast, Polis flex' ongeveer twee keer kleiner dan bij 65 tot 75 jaar. De kansverhouding bij 'EBB vast, Polis flex' is het grootst bij 19 tot 25 jaar. De kansverhouding bij 'EBB flex, Polis vast' is het grootst bij 65 tot 75 jaar. Het totale effect van de SBI-categorieën is ook groot, maar dit is wel het gezamenlijke effect van 16 parameters, 8 voor de effecten van de SBI-categorieën op de eerste categorie van de doelvariabele ('EBB vast, Polis flex') en 8 voor de effecten van de SBI-categorieën op de tweede categorie van de doelvariabele ('EBB flex, Polis vast'). Het totale effect van 'Leeftijdsklasse' is het effect van 10 parameters en het effect van 'Contracturen' omvat 6 parameters.

4.2.1 Bijdrage aan de verklaring van inconsistenties tussen EBB en Polis in de waarneming van arbeidscontract excl. Bedrijfs-grootte, alle leeftijdsgroepen, 2010-2018

Variabele	Chi-kwadraat waarde	Aantal vrijheids-graden	Sign.
Softwarecluster	11026,8	6	0,000
Leeftijdsklasse	7615,0	10	0,000
SBI-groep	3726,6	16	0,000
Baanduur	2632,4	10	0,000
Contracturen	2495,5	6	0,000
Kwartaalnummer	768,0	68	0,000
Opleidingsniveau	121,7	6	0,000
Migratieachtergrond	120,7	12	0,000
Proxywaarneming	65,2	2	0,000
Waarneemmodus	42,3	4	0,000
Laatste respons	41,9	2	0,000
Geslacht	21,7	2	0,000

In deze analyse ontbreekt 'Bedrijfs-grootte'. Deze variabele is later toegevoegd aan het bestand voor de jaren 2016, 2017 en 2018 en is niet beschikbaar voor de eerdere verslagjaren. Om te beoordelen wat het belang is van 'Bedrijfs-grootte' voor het verklaren van verschillen tussen de EBB en de Polis is voor de jaren 2016, 2017 en 2018 deze analyse herhaald.

4.2.2 Bijdrage aan de verklaring van inconsistenties tussen EBB en Polis in de waarneming van arbeidscontract incl. Bedrijfs-grootte van het bedrijf, alle leeftijdsgroepen, 2016-2018

Variabele	Chi-kwadraat waarde zonder en met Bedrijfs-grootte		Aantal vrijheids-graden	Significantie zonder en met Bedrijfs-grootte	
Softwarecluster	4217,5	2970,3	6	0,000	0,000
Leeftijdsklasse	1745,9	1637,7	10	0,000	0,000
Bedrijfs-grootte		1410,2	4		0,000
SBI-groep	1363,6	833,7	16	0,000	0,000
Baanduur	1765,9	1584,0	10	0,000	0,000
Contracturen	772,8	710,7	6	0,000	0,000
Kwartaalnummer	86,6	83,8	68	0,000	0,000
Opleidingsniveau	40,2	35,9	6	0,000	0,000
Migratieachtergrond	46,9	48,0	12	0,000	0,000
Proxywaarneming	34,5	27,3	2	0,000	0,000
Waarneemmodus	35,4	36,0	4	0,000	0,000
Laatste respons	14,2	11,3	2	0,001	0,003
Geslacht	2,6	1,2	2	0,276	0,542

De variabelen in tabel 4.2.2 staan in volgorde van de grootte van de chi-kwadraatwaarden zoals die vastgesteld waren voor de data van 2010 t/m 2018 zonder de variabele 'Bedrijfsgrootte'. De conclusie is dat de eerste vijf variabelen (of zes met 'Bedrijfsgrootte') tot en met 'Contracturen', net als bij de eerdere analyse, de belangrijkste verklarende variabelen zijn. Het effect van 'Bedrijfsgrootte' behoort tot de vier belangrijkste bijdragen aan de verklaring van de inconsistenties tussen EBB en Polis. De parameterwaarden (hier niet getoond) laten zien dat de kans op verschillen tussen EBB en Polis kleiner wordt als de bedrijfsomvang toeneemt. Dit geldt zowel voor 'EBB flex, Polis vast' als voor 'EBB vast, Polis flex' en bovendien zowel voor alle leeftijdsgroepen samen als voor ieder van de leeftijdsgroepen apart. De effecten van 'Softwarecluster' en 'SBI-groep' worden minder als 'Bedrijfsgrootte' in het model wordt opgenomen maar blijven tot de belangrijkste verklarende variabelen behoren.⁵

4.3 De resultaten van de MLR voor 15- tot 25-jarigen

Bij de 15- tot 25-jarigen verklaart de variabele 'Softwarecluster' het meest (tabel 4.3.1). Bij 'cluster 3 en 4' is het effect op de kansverhouding (t.o.v. cluster 1) respectievelijk 1,4 en 1,3 bij 'EBB vast, Polis flex'. Bij 'cluster 2' is dit effect 3,3 voor 'EBB flex, Polis vast'. Dat is het grootste effect op de kansverhouding van alle categorieën van alle verklarende variabelen. De variabele 'SBI-groep' verklaart ook veel. De kans op inconsistenties is veel kleiner bij mensen die werken in de uitzendbranche (SBI=78): de kansverhouding is 2,5 keer kleiner dan bij de referentiecategorie. Met inconsistenties wordt zowel 'EBB vast, Polis flex' als 'EBB flex, Polis vast' bedoeld.

4.3.1 Bijdrage aan de verklaring van inconsistenties tussen EBB en Polis in de waarneming van arbeidscontract, 15- tot 25-jarigen

Variabele	Chi-kwadraat waarde	Aantal vrijheidsgraden	Sign.
Softwarecluster	2872,9	6	0,000
SBI-groep	1401,1	16	0,000
Contracturen	825,6	6	0,000
Kwartaalnummer	602,9	68	0,000
Baanduur	295,7	10	0,000
Waarneemmodus	49,2	4	0,000
Migratieachtergrond	42,4	12	0,000
Leeftijdsklasse	32,0	2	0,000
Opleidingsniveau	31,1	6	0,000
Proxywaarneming	11,5	2	0,003
Laatste respons	9,7	2	0,008
Geslacht	7,9	2	0,019

⁵ Deze analyse is herhaald voor de afzonderlijke leeftijdscategorieën. Die leverden steeds een soortgelijk resultaat op. Daarom wordt er niet over gerapporteerd.

De kans op 'EBB vast, Polis flex' is vooral groter als iemand meer dan 30 uur werkt. Dit is niet in overeenstemming met de verwachting. De kans op 'EBB flex, Polis vast' is wél kleiner (tabel 4.3.2, in bijlage 5 zijn de geschatte parameters opgenomen). Daarnaast geeft een langere baanduur een grotere kans op 'EBB vast, Polis flex'. Ook dit was onverwacht. Blijkbaar staat bij langere baanduren en meer contracturen het arbeidscontract in de Polis vaker ten onrechte als tijdelijk geregistreerd of dat EBB-respondenten met een langere baanduur of meer contracturen hun baan vaker als vast bestempelen bijvoorbeeld op grond van informele afspraken.

4.3.2 Bevindingen van de bijdrage van variabelen aan de verklaring van inconsistenties tussen EBB en Polis in de waarneming van arbeidscontract, 15- tot 25-jarigen

Variabele	Kans op EBB vast Polis flex	Kans op EBB flex Polis vast
Softwarecluster	Groter bij cluster 3 en 4. Iets kleiner bij cluster 2.	Veel groter bij cluster 2. Kleiner bij cluster 3 en 4.
SBI-groep	Groter bij detailhandel en overheid en onderwijs. Kleiner bij landbouw e.d., horeca, financiële instellingen en zorg. Veel kleiner bij uitzendbureaus.	Groter bij landbouw, bosbouw en visserij, en horeca. Kleiner bij zorg. Veel kleiner bij uitzendbureaus.
Contracturen	Groter bij meer contracturen. De toename is geleidelijk naarmate het aantal contracturen toeneemt	Kleiner bij meer contracturen. De afname is geleidelijk naarmate het aantal contracturen toeneemt
Kwartaalnummer	Groter bij kwartaal 2010-1 t/m 2011-4	Groter bij de meeste kwartalen in de periode 2010-1 t/m 2013-4
Baanduur	Groter bij langere baanduur. De toename is geleidelijk naarmate de baanduur toeneemt.	Idem
Waarneemmodus	Iets kleiner bij CAPI en CATI	Geen effect
Migratieachtergrond	Kleiner bij mensen met een Turkse migratieachtergrond	Kleiner bij mensen met een Surinaamse migratieachtergrond
Leeftijdsklasse	Iets groter bij leeftijd 19-25 jaar	Idem
Opleidingsniveau	Iets kleiner bij hoge opleiding	Iets groter bij hoge opleiding
Proxywaarneming	Geen effect	Iets groter bij proxy waarneming
Laatste respons	Geen effect	Iets groter bij laatste respons
Geslacht	Geen effect	Idem

De kans op 'EBB vast, Polis flex' is groter bij de kwartalen in 2010 en 2011 en is er dus een groot verschil tussen de metingen in de Ebb en de Polis. Een effect van een groot bedrijf dat de salarisadministratie anders ging invullen per 1 januari 2018 (vóór 2018 versus tijdens 2018) is niet terug te vinden in de data.

Voor migratieachtergrond geldt dat de effecten anders zijn dan verwacht. Er werd verwacht dat de kans op inconsistenties groter zou zijn bij een niet-Nederlandse migratieachtergrond, omdat de bekendheid met de Nederlandse taal en wet- en regelgeving kleiner zou zijn. Dit blijkt niet uit de cijfers: de kans op inconsistenties is juist kleiner bij mensen met een Turkse of Surinaamse migratieachtergrond.

4.4 De resultaten van de MLR voor 25- tot 55-jarigen

Tabel 4.4.1 geeft de chi-kwadraatwaarden voor de variabelen als voor het model geschat op de data in de leeftijdscategorie 25- tot 55-jarigen. De grootste effecten zijn binnen deze leeftijdsklasse, net als bij de analyse van alle leeftijdscategorieën, de effecten van Softwareleverancier, 'Baanduur', 'SBI-groep', 'Leeftijdsklasse' (hier de verdeling in 25 tot 40 en 40 tot 55 jaar) en 'Contracturen'. Ook binnen deze leeftijdsgroep is er dus nog een belangrijk leeftijdseffect, maar het relatieve belang van Leeftijd is natuurlijk lang niet meer zo groot als bij de analyse van alle leeftijdsgroepen samen.

4.4.1 Bijdrage aan de verklaring van inconsistenties tussen EBB en Polis in de waarneming van arbeidscontract, 25- tot 55-jarigen

Variabele	Chi-kwadraat waarde	Aantal vrijheidsgraden	Sign.
Softwarecluster	7217,6	6	0,000
Baanduur	4273,3	10	0,000
SBI-groep	2246,2	16	0,000
Leeftijdsklasse	1650,1	2	0,000
Contracturen	1517,6	6	0,000
Kwartaalnummer	395,1	68	0,000
Migratieachtergrond	108,3	12	0,000
Opleidingsniveau	77,4	6	0,000
Proxywaarneming	71,2	2	0,003
Waarneemmodus	32,9	4	0,000
Laatste respons	24,8	2	0,008
Geslacht	15,8	2	0,019

In bijlage 6 zijn de geschatte parameters voor dit model weergegeven. Bij 'cluster 3 en 4' van de softwareleveranciers zijn de effecten op de kansverhouding (ten opzichte van cluster 1) bij 'EBB vast, Polis flex' erg groot, namelijk respectievelijk 6,4 en 3,2. Er is voor 'EBB flex, Polis vast' een grotere kans voor 'cluster 2' met een effect op de kansverhouding van 2,6 en voor idem voor cluster 4 met een effect van 1,6. Bij 'Baanduur' blijken de baanduren korter dan 3 jaar zoals verwacht een grotere kans op beide inconsistenties te hebben. De effecten op de kansverhouding liggen bij 'EBB vast, Polis flex' rond de 2 en bij 'EBB flex, Polis vast' rond de 3.

De geselecteerde SBI-categorieën hebben soms een grotere (3 keer) en soms een kleinere (9 keer) kans op een van de inconsistenties dan de overige SBI's. De effecten lopen van 0,5 tot 1,7. Het effect van 'Leeftijdsklasse' laat een grotere kans voor beide inconsistenties zien bij de 25- tot 40-jarigen dan bij de 40- tot 55-jarigen (effecten rond 1,5). Minder dan 12 contracturen leidt bij 'EBB vast, Polis flex' tot meer inconsistenties dan de banen van 30 uur en meer, een effect van 1,4. Bij 'EBB flex, Polis vast' is het effect van contracturen groter. Het effect van de categorie '< 12 uur' is 4,2 en het effect van de categorie '< 20 uur' is 1,9.

4.4.2 Bevindingen van de bijdrage van variabelen aan de verklaring van inconsistenties tussen EBB en Polis in de waarneming van arbeidscontract, 25- tot 55-jarigen

Variabele	Kans op EBB vast Polis flex	Kans op EBB flex Polis vast
Softwarecluster	Veel groter bij cluster 3 en 4 dan bij cluster 1.	Veel groter bij cluster 2, groter bij cluster 4.
Baanduur	Groter bij baanduren korter dan 3 jaar.	Idem, maar een sterker effect.
SBI-groep	Groter bij detailhandel en horeca. Kleiner bij vervoer over water en luchtvaart, financiële instellingen, uitzendbureaus, overheid en onderwijs.	Groter bij horeca. Veel kleiner bij vervoer over water en luchtvaart. Kleiner bij financiële instellingen, uitzendbureaus, overheid en zorginstellingen.
Leeftijdsklasse	Groter bij de 25-40-jarigen	Groter bij de 25-40-jarigen
Contracturen	Groter bij <12 contracturen.	Veel groter bij <12 contracturen en groter bij <20 contracturen.
Kwartaalnummer	Iets groter bij de meeste kwartalen ten opzichte van de meest recente. Voor 4 verspreide kwartalen groter.	Groter (en twee keer veel groter) bij de kwartalen 2010-1 t/m 2017-1 dan bij de laatste 7 kwartalen.
Migratieachtergrond	Iets groter bij Overige niet-westerse landen.	Groter voor mensen met een Turkse migratieachtergrond, Overige niet-westerse landen en Overige westerse landen.
Opleidingsniveau	Iets kleiner bij hoge en midden opleidingsniveaus.	Groter bij opleiding onbekend.
Proxywaarneming	Iets kleiner bij geen proxy	Geen effect
Waarneemmodus	Iets kleiner bij CATI	Geen effect
Laatste respons	Iets kleiner bij geen laatste respons.	Geen effect
Geslacht	Iets kleiner bij mannen	Geen effect

Voor een viertal kwartalen is de kans op 'EBB vast, Polis flex' groter dan bij de meest recente twee kwartalen. Bij 'EBB flex, Polis vast' zijn bij veel meer kwartalen (26) grotere kansen dan bij de meest recente en bovendien zijn de effecten groter. Bij 'EBB flex, Polis vast' hebben 'Turken', 'Overige niet-westerse landen' en 'Overige westerse landen' een grotere kans dan mensen met een Nederlandse migratieachtergrond.

4.5 De resultaten van de MLR voor 55-jarigen en ouder

De laatste leeftijdscategorie die is onderzocht zijn de 55-jarigen en ouder. Uit tabel 4.5.1 blijkt dat net als bij de overige leeftijdscategorieën de vijf grootste effecten te zien zijn bij de variabelen: 'Softwarecluster', 'Contracturen', 'Baanduur', 'SBI-groep' en 'Leeftijdsklasse' (hier de verdeling in 55 tot 65 en 65 tot 75 jaar).

4.5.1 Bijdrage aan de verklaring van inconsistenties tussen EBB en Polis in de waarneming van arbeidscontract, 55-jarigen en ouder

Variabele	Chi-kwadraat waarde	Aantal vrijheidsgraden	Sign.
Softwarecluster	1820,4	6	0,000
Contracturen	1008,3	6	0,000
Baanduur	1007,3	10	0,000
SBI-groep	853,0	16	0,000
Leeftijdsklasse	256,4	2	0,000
Kwartaalnummer	127,2	68	0,000
Migratieachtergrond	26,0	12	0,000
Opleiding	14,4	6	0,000
Geslacht	11,2	2	0,001
Waarneemmodus	9,9	4	0,001
Proxywaarneming	7,5	2	0,012
Laatste respons	3,5	2	0,000

In bijlage 7 zijn de geschatte parameters voor dit model weergegeven. Bij 'cluster 3 en 4' van 'Softwarecluster' zijn de effecten op de kansverhouding (ten opzichte van 'cluster 1') bij 'EBB vast, Polis flex' erg groot, namelijk respectievelijk 8 en 3,6. Voor 'EBB flex, Polis vast' is een grotere kans bij bedrijven met software uit 'cluster 2' (met een effect op de kansverhouding van 2,5 in vergelijking met cluster 1). Minder contracturen leidt tot meer inconsistenties van beide soorten. Bij 'EBB vast, Polis flex' lopen de effecten van 1,3 voor '20 tot 30 uur' tot 1,9 voor 'minder dan 12 uur'. Bij 'EBB flex, Polis vast' zijn de effecten veel sterker en lopen van 1,8 tot 7. Bij 'Baanduur' blijken de baanduren korter dan 3 jaar een veel grotere kans op beide inconsistenties te hebben. De effecten op de kansverhouding liggen rond de 3. Voor de eerste categorie 'Minder dan 3 maanden' is de schatting van het effect voor beide inconsistenties echter onbetrouwbaarder waardoor het niet significant is bij een overschrijdingskans van 0,01. De geselecteerde SBI-groepen hebben soms een grotere (3 keer) en soms een kleinere (8 keer) kans op een van de inconsistenties dan de overige SBI's. De effecten lopen van 0,4 tot 1,8. Het effect van leeftijdsklasse laat een kleinere kans voor beide inconsistenties zien bij de categorie '55 tot 65 jaar' dan bij '65 jaar en ouder'. Het effect is bij 'EBB vast, Polis flex' 0,6 en met 0,3 nog veel sterker bij 'EBB flex, Polis vast'.

Voor drie kwartalen is de kans op 'EBB vast, Polis flex' groter dan bij de meest recente twee kwartalen. Bij 'EBB flex, Polis vast' zijn er in zeven kwartalen grotere kansen dan bij

de meest recente en bovendien zijn de effecten groter dan bij 'EBB vast, Polis flex'. Bij 'Migratieachtergrond' is er een grotere kans op beide inconsistenties voor de categorie 'Overige niet westerse landen'.

4.5.2 Bevindingen van de bijdrage van variabelen aan de verklaring van inconsistenties tussen EBB en Polis in de waarneming van arbeidscontract, 55-jarigen en ouder

Variabele	Kans op EBB vast Polis flex	Effect op EBB flex Polis vast
Softwarecluster	Veel groter bij cluster 3 en 4.	Veel groter bij cluster 2.
Contracturen	Groter als het aantal contracturen afneemt.	Veel groter als het aantal contracturen afneemt.
Baanduur	Baanduren van 3 maanden tot 3 jaar een veel grotere kans dan baanduren langer dan 3 jaar.	Idem
SBI-groep	Groter bij Detailhandel, Horeca. Kleiner bij Vervoer over water en door de lucht, Financiële instellingen, Uitzendbureaus en Zorginstellingen. Veel kleiner bij Overheid en onderwijs.	Groter bij Horeca. Kleiner bij Financiële instellingen, Uitzendbureaus en Zorginstellingen.
Leeftijdsklasse	Lager bij de categorie 55-64 dan bij ouder dan 64	Veel lager bij de categorie 55-64 dan bij ouder dan 64
Kwartaalnummer	Groter dan het meest recente kwartaal voor 3 verspreide eerdere kwartalen.	Groter bij 7 kwartalen. Sterkere effecten dan bij EBB vast Polis flex.
Migratieachtergrond	Groter bij Overig niet westers	Idem
Opleidingsniveau	Geen effect	Groter bij Opleiding onbekend.
Geslacht	Geen effect	Iets groter bij mannen
Waarneemmodus	Geen effect	Iets kleiner bij CATI
Proxywaarneming	Geen effect	Idem
Laatste respons	Geen effect	Idem

5. De schatting van de onnauwkeurigheden

5.1 De replicatie van de meetmodellen van 2007 en 2009

Dit onderdeel start met de replicatie van het meetmodel van Pavlopoulos en Vermunt (2015) en Pankowska et al. (2018). Dit wordt het oude model genoemd. Ten opzichte van deze studies is in de data een belangrijke wijziging aangebracht: de operationalisering van

de afhankelijke variabele. In de eerdere studies is gebruik gemaakt van de toenmalige standaard, waarbij personen die minder dan twaalf uur per week werkzaam waren niet tot de werkzame beroepsbevolking werden gerekend. In het huidige onderzoek is dit criterium verdwenen. Dit leidt er vooral toe dat er minder mensen in de categorie 'Overig' terecht komen dan voorheen. Voor de rest is zo precies mogelijk gerepliceerd, dat wil zeggen dat de modellen zijn geschat voor 25- tot 55-jarigen die in het eerste kwartaal van 2007, 2009, 2016, 2017 of 2018 voor het eerst aan de EBB meededen. De leeftijdsafbakening is conform die van de eerdere studies. Daarbij worden de respondenten vijftien maanden gevolgd. Sommigen hebben niet in alle vijf de EBB-peilingen gerepsondeerd maar zijn eerder uitgevallen. Bij de analyse over 2018 worden de respondenten twaalf maanden gevolgd, omdat nog niet alle benodigde data over 2019 beschikbaar waren. Het aantal personen waarop de modellen geschat worden is in deze vijf jaren, respectievelijk, 11605 (2007), 8486 (2009), 12843 (2016), 11910 (2017) en 14688 (2018).

Binnen de modellen die in de eerdere studies zijn geschat is voor de nieuwe data gezocht naar het zuinigste model dat de data adequaat beschrijft. Zuinigst betekent met zo weinig mogelijk te schatten parameters waarna wordt gekeken of de waargenomen data goed worden gereproduceerd door het model. Van de negen modellen die zijn geschat passen volgens een gangbare maat (de Bayesian Information Criterion, BIC) de twee meest uitgebreide modellen voor alle jaren het beste bij de data (zie model 8 en 9 in bijlage 8). Volgens deze twee modellen bevatten beide bronnen (Polis en EBB) onnauwkeurigheden. De foutkansen in de Polis zijn serieel gecorreleerd in de tijd en de toewijzing aan het werkelijke arbeidscontract is afhankelijk van leeftijd (drie klassen), land van herkomst (zeven klassen), geslacht (twee klassen) en opleidingsniveau (vier klassen). In model 8 is aangenomen dat de foutkansen in de EBB homogeen zijn. In model 9 mogen de foutkansen in de EBB afhangen van covariaten: leeftijd in drie klassen en al dan niet proxywaarneming. In 2007 en 2009 is model 8 en de overige drie jaren is model 9 volgens de BIC het best passende model. In het vervolg worden per jaar de uitkomsten voor het best passende model gepresenteerd. De uitkomsten voor model 8 en 9 verschillen overigens slechts minimaal van elkaar en leiden steeds tot dezelfde conclusies.

In tabel 5.1.1 zijn de onnauwkeurigheden weergegeven uit de replicaties. De resultaten zijn gewogen met het EBB-jaargewicht van peiling 1 van de respectievelijke jaren. Eerdere berekeningen lieten overigens zien dat voor de schatting van de onnauwkeurigheden de weging weinig verschil maakt. De weging heeft wel effect op de schatting van het aantal, aandeel en overgangskansen. Voor de Polis- en EBB-variabele zijn de geschatte onnauwkeurigheden weergegeven als kansen in percentages. Op de diagonalen zou overal 100 procent staan als er geen sprake was van onnauwkeurigheden. De percentages laten zich als volgt interpreteren. In 2007 heeft bijvoorbeeld van de mensen die een vast arbeidscontract hebben volgens de uitkomsten van het model, 99,6 procent volgens de EBB een vast arbeidscontract, 0,2 procent heeft volgens de EBB een flexibel arbeidscontract en 0,2 procent is volgens de EBB geen werknemer.

De uitkomsten laten zien dat het arbeidscontract voor mensen die volgens het model een vast arbeidscontract hebben of in de categorie 'Overig' vallen nauwkeuriger gemeten is in de EBB dan in de Polis. Voor mensen die volgens het model een flexibel arbeidscontract hebben doet de Polis het echter beter. In de EBB zegt een kwart tot een derde van de werknemers die volgens het model een flexibel arbeidscontract hebben dat ze een vast

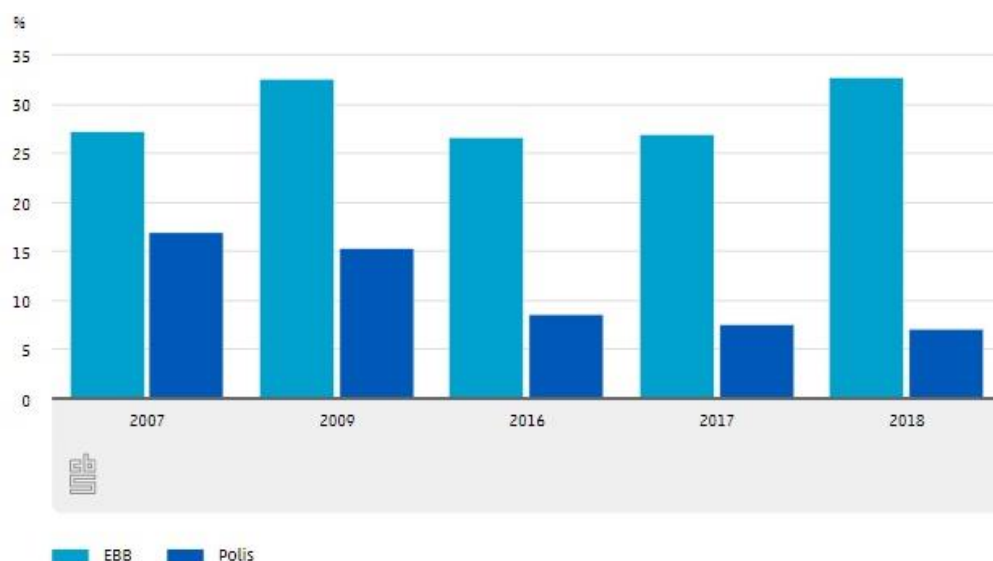
arbeidscontract hebben, terwijl dat in de Polis aanvankelijk in 2007 en 2009 ongeveer een zesde en vanaf 2016 nog geen 10 procent is.⁶

5.1.1 Onnauwkeurigheden in de EBB en de Polis volgens het oude model voor 25- tot 55-jarigen die in het eerste kwartaal van 2007, 2009, 2016, 2017 of 2018 voor het eerst respondeerden in de EBB

2007	Waargenomen in EBB				Waargenomen in Polis		
	Vast	Flex	Overig		Vast	Flex	Overig
Model	%						
Vast	99,6	0,2	0,2		90,5	7,1	2,3
Flex	27,3	69,2	3,6		17,0	73,5	9,5
Overig	0,1	0,3	99,5		3,0	2,1	94,9
2009	Waargenomen in EBB				Waargenomen in Polis		
	Vast	Flex	Overig		Vast	Flex	Overig
Model	%						
Vast	99,6	0,2	0,2		89,7	8,4	1,9
Flex	32,6	63,4	3,9		15,3	74,8	9,9
Overig	0,2	0,4	99,3		3,9	2,1	94,0
2016	Waargenomen in EBB				Waargenomen in Polis		
	Vast	Flex	Overig		Vast	Flex	Overig
Model	%						
Vast	99,5	0,3	0,2		91,0	7,4	1,6
Flex	26,6	69,8	3,6		8,6	86,6	4,8
Overig	0,4	0,7	98,9		2,0	2,1	95,9
2017	Waargenomen in EBB				Waargenomen in Polis		
	Vast	Flex	Overig		Vast	Flex	Overig
Model	%						
Vast	99,3	0,4	0,3		91,7	6,6	1,8
Flex	26,9	69,6	3,5		7,6	87,2	5,2
Overig	0,6	0,8	98,6		2,2	2,0	95,8
2018	Waargenomen in EBB				Waargenomen in Polis		
	Vast	Flex	Overig		Vast	Flex	Overig
Model	%						
Vast	99,1	0,5	0,4		93,8	4,5	1,7
Flex	32,8	64,2	3,1		7,1	89,3	3,7
Overig	0,7	1,3	98,0		2,0	2,5	95,6

⁶ Als de uitkomsten Pavlopoulos en Vermunt (2015) voor 2007, Pankowska *et al.* (2018) voor 2009 worden vergeleken met onze nieuwe uitkomsten, dan blijken de onnauwkeurigheden in de EBB bij de mensen met een flexibele arbeidsrelatie flink wat hoger te zijn. Dit kan verklaard worden door de wijziging die is aangebracht in de operationalisering van de afhankelijke variabele.

5.1.2 Percentage dat een flexibel contract heeft volgens het model en in de Ebb en de Polis een vast contract



De geschatte onnauwkeurigheid van de EBB meting is nogal wisselvallig (figuur 5.1.2). Zo is de onnauwkeurigheid in de meting van flexibel arbeidscontract in 2009 en 2018 het hoogst, en is deze onnauwkeurigheid in 2016 en 2017 een stuk lager. De nauwkeurigheid van de meting in de Polis is juist toegenomen. Bij de Polis is een relatief grote afname te zien van de mensen die volgens het model een flexibel arbeidscontract hebben, maar in de Polis staan geregistreerd met een vast arbeidscontract. Was het nog bijna 17 procent in 2009, in 2018 is dit gedaald naar 7 procent. Daarmee wordt bevestigd dat de kwaliteit van de Polis in de loop der tijd is toegenomen. Dit kan wellicht ook een verklaring zijn voor de uiteenlopende ontwikkeling van Polis en EBB. Om daar meer zicht op te krijgen is inzicht in de ontwikkeling in de onnauwkeurigheid per kwartaal nodig. Dat zou betekenen dat niet alleen de EBB-respondenten die in het eerste kwartaal voor het eerst aan de EBB meededen maar ook de instromers in de andere drie kwartalen onderzocht moeten worden. De resultaten worden gerapporteerd in paragraaf 5.6.

In de eerdere studies werd steeds gevonden dat als er een fout in de Polis gemaakt is, de kans relatief groot is dat deze in een volgende periode er nog steeds in zit. Dat wordt ook in 2016-2018 teruggevonden. In de eerste plaats omdat de modellen waarin de onnauwkeurigheden in de Polisdata serieel zijn gecorreleerd het beste bij de data passen.

5.1.3 Kansen op herhaling van de onnauwkeurigheden in de Polis volgens het oude model 2016, 2017, 2018 gemaakt in t-1

	Waargenomen in t								
	2016			2017			2018		
	Vast	Flex	Overig	Vast	Flex	Overig	Vast	Flex	Overig
Model									
Vast	-	98,0	98,0	-	98,5	98,7	-	99,5	98,5
Flex	96,0	-	89,0	95,8	-	91,1	96,5	-	88,3
Overig	95,6	88,7	-	96,5	89,8	-	96,1	89,1	-

In de tweede plaats kun je dat ook terugvinden in de foutkansen zoals die zijn weergegeven in tabel 5.1.3. Zo is bijvoorbeeld in 2016 als iemand in maand t-1 volgens het model een flexibel contract heeft en met een vast arbeidscontract staat geregistreerd in de Polis, de kans 96,0 procent is dat in maand t dezelfde (foute) waarde staat geregistreerd als het arbeidscontract volgens het model niet is veranderd.

5.1.4 Geschatte verdeling naar contractvorm volgens het oude model, de EBB en de Polis voor 2016, 2017 en 2018, 15-maands gemiddelde

	2016				2017				2018		
	Model	EBB	Polis		Model	EBB	Polis		Model	EBB	Polis
Contract	%										
Vast	52,0	57,1	49,4		51,8	56,9	49,5		51,4	57,9	50,2
Flex	17,9	11,9	20,0		18,7	12,5	20,3		19,7	12,6	20,6
Overig	30,2	31,0	30,6		29,5	30,6	30,2		28,9	29,5	29,2

In tabel 5.1.4 is de geschatte verdeling naar arbeidscontract weergegeven volgens het oude model, de EBB en de Polis voor 2016, 2017 en 2018. Daarbij is het gemiddelde over 15 maanden gegeven per cohort (voor 2018: 12 maanden). Het verschil in de schatting van het percentage flexibele arbeidscontracten tussen de EBB en de Polis wordt tussen 2016 en 2017 iets kleiner (van 8,1 naar 7,8 procent) en tussen 2017 en 2018 weer groter (naar 8,0 procent). De uitkomsten van het model blijven relatief dicht bij de verdeling zoals die in de Polis is geschat. Het percentage flexibele arbeidscontracten wordt door de EBB afgerond geschat op 12, 13 en 13, en door de Polis respectievelijk op 20, 20 en 21. Na correctie van beide bronnen wordt het percentage geschat op 18, 19 en 20. Dit is in overeenstemming met de eerder gevonden omvang van de onnauwkeurigheden in de EBB en de Polis.

Vervolgens zijn kwartaal-op-kwartaal overgangskansen geschat (tabel 5.1.5). De overgang van flexibel naar vast wordt in 2016, 2017 en 2018 respectievelijk geschat op 4,3, 4,9 en 4,4 procent. Dit is een lagere kans dan zowel in de Polis als in de EBB wordt geschat. De Polis schat deze overgangskans op 6,0, 6,6 en 5,8 procent en de EBB op 13,8, 15,5 en 16,7 procent. Rekening houdend met de toename van de onnauwkeurigheden in de EBB en de definitieverandering in de afhankelijke variabele is dit in overeenstemming met de eerdere bevindingen van Pavlopoulos en Vermunt (2015) en Pankowska et al. (2018). Beoordeeld op grond van de ontwikkeling is de schatting op basis van de Polis dicht bij de modelschatting dan die van de EBB. Hierin kan een verklaring gezien worden waarom de schattingen uit de Polis en de EBB verder uit elkaar zijn geraakt. Toch moet hieraan niet teveel waarde worden gehecht om twee redenen. De eerste reden is dat het model niet perfect is en ten tweede zijn de resultaten uitsluitend gebaseerd op de respondenten die in het eerste kwartaal van ieder jaar aan de EBB meedoen. In de volgende paragraaf wordt het model verbeterd.

5.1.5 Geschatte kwartaal-op-kwartaal overgangskansen volgens het oude model, de EBB en de Polis voor 2016, 2017 en 2018

2016	Vast-Vast	Vast-Flex	Vast-Overig	Flex-Vast	Flex-Flex	Flex-Overig	Overig-Vast	Overig-Flex	Overig-Overig
	%								
Model	98,8	0,6	0,6	4,3	92,1	3,7	0,2	3,9	95,8
Polis	97,3	1,7	1,1	6,0	88,9	5,1	0,7	4,8	94,5
EBB	97,9	1,0	1,1	13,8	79,8	6,4	1,1	4,5	94,3
2017	Vast-Vast	Vast-Flex	Vast-Overig	Flex-Vast	Flex-Flex	Flex-Overig	Overig-Vast	Overig-Flex	Overig-Overig
	%								
Model	98,9	0,7	0,5	4,9	91,4	3,7	0,1	4,1	95,8
Polis	97,4	1,7	0,9	6,6	88,3	5,1	0,6	4,8	94,5
EBB	97,8	1,1	1,2	15,5	77,8	6,7	1,1	4,4	94,5
2018	Vast-Vast	Vast-Flex	Vast-Overig	Flex-Vast	Flex-Flex	Flex-Overig	Overig-Vast	Overig-Flex	Overig-Overig
	%								
Model	98,6	0,9	0,5	4,4	92,5	3,1	0,3	4,0	95,8
Polis	97,1	1,8	1,0	5,8	89,5	4,7	0,7	5,2	94,0
EBB	97,7	1,3	1,0	16,7	77,3	6,0	1,1	5,3	93,7

5.2 De verbetering van het meetmodel voor de 25- tot 55-jarigen

In de vorige paragraaf is een replicatie uitgevoerd van het model zoals dat door Pavlopoulos en Vermunt (2015) en Pankowska et al. (2018) is geschat (het oude model), waarbij de metingen van het arbeidscontract in de Polis en de EBB zijn geactualiseerd en er drie jaargangen aan 2007 en 2009 zijn toegevoegd: 2016, 2017 en 2018. In dit hoofdstuk worden enkele verbeteringen doorgevoerd hierin voor de jaren 2016, 2017 en 2018. De mate van verbetering wordt weergegeven in termen van het Bayesian Information Criterion (BIC). Daarvoor geldt dat hoe lager de waarde daarvan is, hoe beter het model bij de gegevens past. De waarden daarvan worden alleen voor het verslagjaar 2016 gegeven, maar in 2017 en 2018 gelden soortgelijke verbeteringen. Daarnaast is ook het verschil in het aantal parameters gegeven. Dit is voor de interpretatie van de BIC niet per se noodzakelijk, maar geeft een indicatie hoeveel parameters aan het model zijn toegevoegd en daarmee van de complexiteit van het model.

In het oude model wordt verondersteld dat de waarnemingen in de achtereenvolgende metingen in de Polis gecorreleerd zijn als deze een fout bevatten. De onderbouwing daarvan is dat als er fouten bij de invoering van de gegevens in de Polis zijn gemaakt, deze vaak erin blijven omdat meestal niemand er belang bij heeft dat deze fouten worden hersteld. Pas wanneer er een wijziging in de administratie plaatsvindt wordt een dergelijke

fout opgemerkt en hersteld. Het is waarschijnlijk dat de waarnemingen in de Polis ook gecorreleerd zijn als deze juist waren. Dat is aan het model toegevoegd. Dit leverde in 2016 een substantiële verbetering op van het model. De BIC daalt van 84739 naar 80246 met een verlies van 3 vrijheidsgraden.

In het oude model is verondersteld dat de waarnemingen in de EBB in de loop der tijd niet gecorreleerd zijn. Het is echter waarschijnlijk dat als een werknemer het idee heeft een vast of flexibel contract te hebben, deze dat bij herhaling zal antwoorden in de EBB ook al klopt dat niet. Daarom is in het nieuwe model verondersteld dat de onnauwkeurigheden in de waarnemingen in de EBB ook gecorreleerd zijn, maar alleen wanneer er in de metingen sprake is van dezelfde baan. Als er van baan wordt gewisseld is het veel onwaarschijnlijker dat de onnauwkeurigheden gecorreleerd zijn. De waarnemingen in de EBB mogen gecorreleerd zijn als het werkelijk type arbeidscontract niet veranderd is ten opzichte van het vorige kwartaal, maar alleen als de respondent bij 'Zelfde werkkring als vorige peiling?' antwoordt (a) 'Zelfde werkgever' of (b) 'Niet van toepassing' in combinatie met type arbeidsrelatie vorig kwartaal 'Overig'. De aanname is dus dat er géén correlatie is als de respondent antwoordt: 'Andere werkgever'. Het antwoord 'Niet van toepassing' in combinatie met type arbeidsrelatie vorig kwartaal 'Vast' of 'Flex' komt in de data zo weinig voor dat een eventuele correlatie – als die er is – niet geschat kan worden. Dit levert een verbetering van het model op: de BIC daalt van 80246 naar 79316 met een verlies van 5 vrijheidsgraden.

Een volgende verbetering van het model is dat wordt verondersteld dat de onnauwkeurigheden in de waarnemingen in de Polis alleen gecorreleerd zijn als sprake is van dezelfde baan. Als van baan wordt veranderd worden werknemers weer opnieuw ingevoerd in de Polis en worden eventuele fouten niet herhaald. Bij de Polis is dit geoperationaliseerd als een contract tussen dezelfde werkgever en werknemer. Dit levert een substantiële verbetering van het model op: de BIC daalt van 79316 naar 55644 bij een verlies van 15 vrijheidsgraden.

In een HMM-model wordt verondersteld dat er geen onverklaarde variantie is in de eerste meting van de latente variabele, in dit geval het type arbeidscontract. Aan deze veronderstelling wordt meestal voldaan door een nieuwe latente variabele aan te maken die de eerste meting volledig verklaart. Door Pavlopoulos en Vermunt (2015) was een variabele met 3 categorieën toegevoegd aan het model. In dit geval is dit opgelost door een reeks van covariaten toe te voegen die van invloed is op de meting van het type arbeidscontract: 'Leeftijd' (lineair en in 2 klassen), 'Geslacht' (2 klassen), 'Opleidingsniveau' (4 klassen), 'SBI-groep' (9 klassen), 'Contracturen' (5 klassen), 'Baanduur' (lineair, het kwadraat en een dummyvariabele voor als de waarde ontbreekt), 'Softwarecluster' (5 klassen), 'Bedrijfsgrootte' (3 klassen). Daarmee wordt tegelijkertijd goed in beeld gebracht hoe de relaties zijn tussen deze covariaten en het type arbeidscontract. Met deze covariaten kan niet alle variantie worden verklaard in de eerste meting van het arbeidscontract, maar de toegevoegde latente variabele kan wel teruggebracht worden van drie naar twee klassen. 'Migratieachtergrond' is niet in het model opgenomen omdat deze geen variantie verklaart bovenop de variantie die door de andere covariaten wordt verklaard. In zijn geheel levert deze stap een substantiële verbetering van het model op: de BIC daalt van 55644 naar 40123 bij een verlies van 44 vrijheidsgraden.

De volgende verbetering van het model is dat covariaten van invloed mogen zijn op de overgangskansen in het type arbeidscontract tussen opeenvolgende maanden. Het blijkt dat de volgende covariaten het model verbeteren: 'Tijd' (lineair en kwadratisch), 'Leeftijd' (lineair), 'Geslacht' (2 klassen), 'Opleidingsniveau' (4 klassen), 'Zelfde baan of niet' (2 klassen), 'Softwarecluster' (5 klassen), 'SBI-groep' (9 klassen), 'Baanduur' (2 klassen), en 'Contracturen' (2 klassen), en de toegevoegde covariaat voor overgebleven heterogeniteit (2 klassen). Daarmee worden de standen per meetmoment en de overgangskansen zuiverder geschat. Dit levert een verbetering van het model op: de BIC daalt van 40123 naar 34831 met een verlies van 25 vrijheidsgraden. In totaal bevat het model nu 240 te schatten parameters.

Dit best passende model voor het verslagjaar 2016 is toegepast op de verslagjaren 2017 en 2018 (net als alle tussenstappen die steeds een verbetering opleveren ten opzichte van het simpeler model). Ook in 2017 en 2018 wordt de BIC substantieel verbeterd. In 2017 daalt deze van 79162 naar 33702 en in 2018 van 84757 naar 36244 met een verlies van 92 vrijheidsgraden. Vanwege deze grote verbetering van de BIC wordt dit model gebruikt voor de berekening van de onnauwkeurigheden in de meting, de modelschatting van de frequenties van de verschillende type arbeidscontracten en de overgangskansen daartussen in de tijd.

5.2.1 Geschatte onnauwkeurigheid van de metingen in de Polis en de EBB op basis van het nieuwe model, 25- tot 55-jarigen

2016	Waargenomen in EBB			Waargenomen in Polis		
	Vast	Flex	Overig	Vast	Flex	Overig
Model	%					
Vast	98,6	0,5	1,0	93,2	6,7	0,1
Flex	33,2	59,7	7,1	10,7	89,2	0,1
Overig	4,0	2,6	93,4	0,4	0,0	99,6
2017	Waargenomen in EBB			Waargenomen in Polis		
	Vast	Flex	Overig	Vast	Flex	Overig
Model	%					
Vast	98,2	0,5	1,4	92,9	7,0	0,1
Flex	29,6	63,3	7,1	8,9	91,1	0,0
Overig	4,2	3,0	92,8	0,1	0,1	99,8
2018	Waargenomen in EBB			Waargenomen in Polis		
	Vast	Flex	Overig	Vast	Flex	Overig
Model	%					
Vast	98,2	0,7	1,1	93,1	6,8	0,2
Flex	30,2	62,3	7,4	8,3	91,6	0,1
Overig	4,0	2,7	93,3	0,4	0,0	99,6

In tabel 5.2.1 zijn de schattingen weergegeven van de onnauwkeurigheden in de metingen van de Polis en de EBB volgens het nieuwe model in 2016-2018. Hieruit blijkt dat de nauwkeurigheid van de Polis nog iets hoger en die van de EBB iets lager wordt geschat dan in het oude model, maar veel maakt het niet uit. De onnauwkeurigheid is nog steeds

vooral groot bij de waarneming in de EBB voor hen die volgens het model een flexibel arbeidscontract hebben. Ook in de Polis is dit een probleem, maar is dit veel kleiner.

5.2.2 Geschatte verdeling naar contractvorm volgens het nieuwe model, de EBB en de Polis voor 2016, 2017 en 2018, 25- tot 55-jarigen

	2016				2017				2018		
	Model	EBB	Polis		Model	EBB	Polis		Model	EBB	Polis
Contract	%										
Vast	50,8	57,1	49,4		51,5	56,9	49,5		52,2	57,9	50,2
Flex	18,6	11,9	20,0		18,3	12,5	20,3		18,6	12,6	20,6
Overig	30,7	31,0	30,6		30,2	30,6	30,2		29,2	29,5	29,2

Vervolgens is de verdeling geschat naar de onderscheiden soorten arbeidscontracten in 2016, 2017 en 2018 (tabel 5.2.2). De modelschattingen verschillen slechts in geringe mate van die van het oude model. De schatting ligt dicht bij de geobserveerde percentages van de Polis. De geobserveerde percentages uit de EBB liggen hoger voor de vaste en lager voor de flexibele arbeidscontracten. Volgens het model stijgt het percentage vaste arbeidscontracten monotoon, terwijl dat voor de Polis tussen 2016 en 2017 langzaam stijgt en daarna sneller stijgt. Voor de EBB geldt dat het geschatte percentage vaste contracten wat springerig is: van 57,1, via 56,9 naar 57,9 procent.

Als laatste worden de overgangskansen beoordeeld en in het bijzonder de kans om in een kwartaal van een flexibele naar een vast arbeidscontract over te stappen (tabel 5.2.3). Volgens het nieuwe model wordt dit laatste percentage in 2016 geschat op 5,2, terwijl in de Polis en in de EBB dit hoger wordt geschat (respectievelijk 6,0 en 13,8 procent). Bij de EBB wordt dit gecompenseerd door ook een veel kleinere kans te schatten voor de overgang van een flexibele naar weer een flexibel arbeidscontract. Bedenk hierbij dat dit niet alleen gaat om werknemers die van baan veranderen, maar voor alle werknemers met een flexibel arbeidscontract. Opvallend is verder dat de overgangskans van een flexibel naar een vast arbeidscontract in de EBB tussen 2016 en 2018 stijgt, terwijl dit in het model en de Polis juist ongeveer gelijk blijft. De overige verschillen zijn relatief gering.

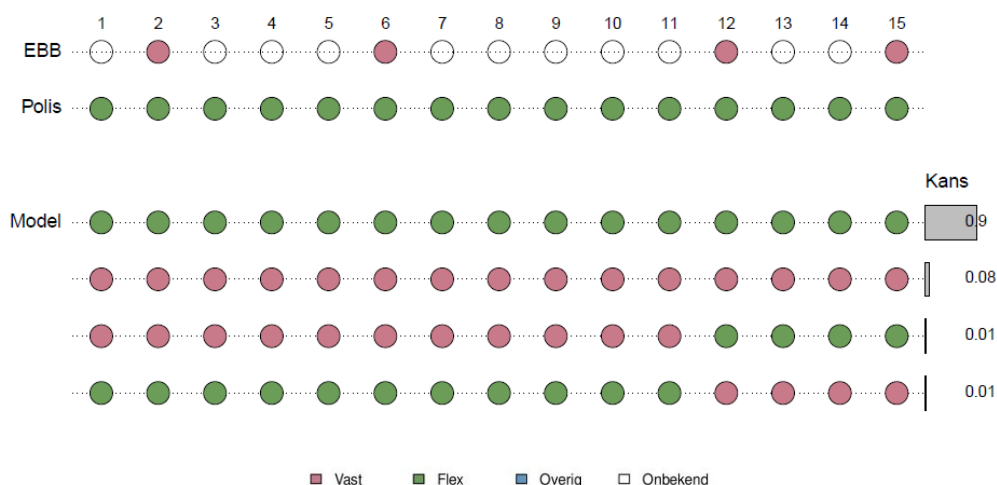
5.2.3 Geschatte overgangskansen volgens het nieuwe model, de EBB en de Polis voor 2016, 2017 en 2018, 25- tot 55-jarigen

2016	Vast-Vast	Vast-Flex	Vast-Overig	Flex-Vast	Flex-Flex	Flex-Overig	Overig-Vast	Overig-Flex	Overig-Overig
	%								
Model	97,8	1,3	0,9	5,2	88,9	5,8	0,4	5,0	94,6
Polis	97,3	1,7	1,1	6,0	88,9	5,1	0,7	4,8	94,5
EBB	97,9	1,0	1,1	13,8	79,8	6,4	1,1	4,5	94,3
2017	Vast-Vast	Vast-Flex	Vast-Overig	Flex-Vast	Flex-Flex	Flex-Overig	Overig-Vast	Overig-Flex	Overig-Overig
	%								
Model	97,7	1,5	0,8	5,6	88,7	5,7	0,4	5,1	94,6
Polis	97,4	1,7	0,9	6,6	88,3	5,1	0,6	4,8	94,5
EBB	97,8	1,1	1,2	15,5	77,8	6,7	1,1	4,4	94,5
2018	Vast-Vast	Vast-Flex	Vast-Overig	Flex-Vast	Flex-Flex	Flex-Overig	Overig-Vast	Overig-Flex	Overig-Overig
	%								
Model	97,7	1,5	0,9	5,5	89,2	5,3	0,6	5,3	94,1
Polis	97,1	1,8	1,0	5,8	89,5	4,7	0,7	5,2	94,0
EBB	97,7	1,3	1,0	16,7	77,3	6,0	1,1	5,3	93,7

5.3 De uitkomsten van het meetmodel aanschouwelijk gemaakt

Het model dat in deze studie gebruikt en verder ontwikkeld is, levert uitkomsten op waarvan niet direct duidelijk is hoe deze samenhangen met de data die als input hebben gediend. Om die relatie duidelijker te maken zijn de resultaten gevisualiseerd voor enkele concrete gevallen waarin het patroon van arbeidscontracten zoals die waargenomen worden in de Polis en de EBB door de tijd heen verschilt. Bij deze figuren zijn ook de kansen weergegeven die het model schat wat de werkelijke patronen zijn. Zo is in figuur 5.3.1 de kans op het eerste patroon 0,90, dat wil zeggen dat gegeven de waargenomen waarden in de EBB en de Polis de kans volgens het model 90 procent is dat dit het werkelijke patroon is. Het betreft een vereenvoudigde weergave: alleen de soort arbeidscontract is weergegeven. In werkelijkheid zijn in het model ook covariaten opgenomen die de schattingen in het model mede bepalen zoals leeftijd, softwarecluster, bedrijfsgrootte en baanduur.

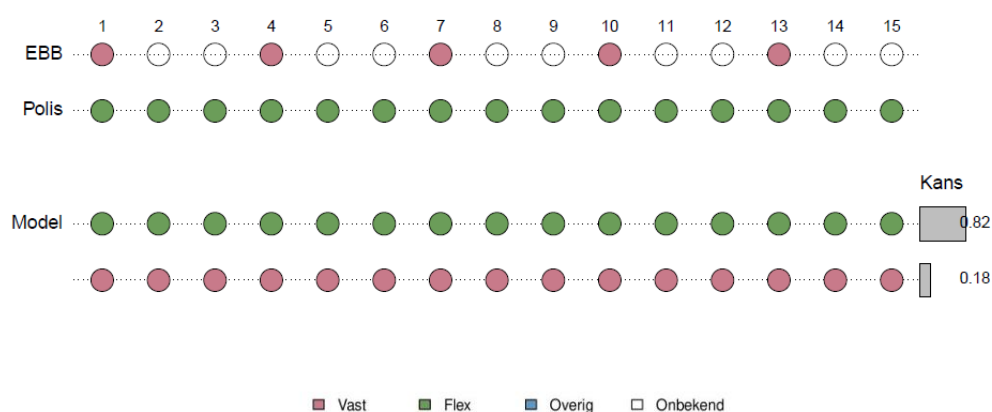
5.3.1 Voorbeeld 1



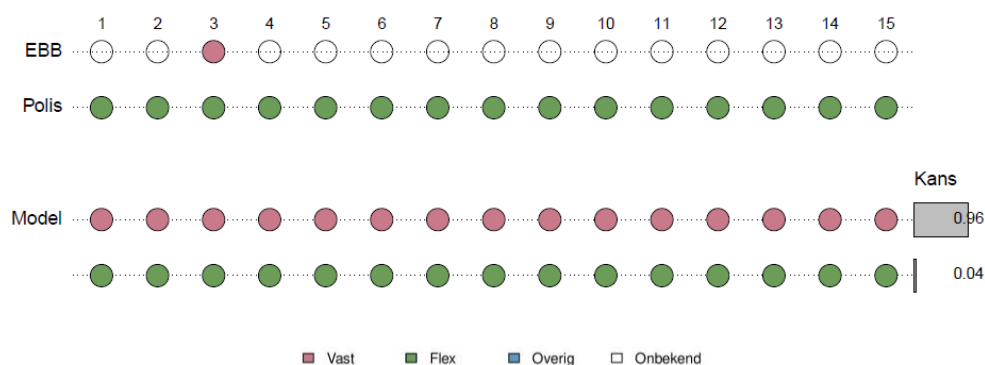
De bovenstaande figuur bevat de situatie dat in de EBB op alle peilmomenten een vast arbeidscontract wordt waargenomen, terwijl informatie uit de Polis juist aangeeft dat het een flexibel arbeidscontract zou betreffen. Het model heeft met behulp van de achtergrondkenmerken en de patronen berekend dat het met 90 procent zekerheid een flexibel arbeidscontract is en 8 procent dat het om een vast contract gaat. Maar andere mogelijkheden worden ook niet helemaal uitgesloten.

Hieronder is op het eerste gezicht precies eenzelfde situatie aanwezig als in het eerste voorbeeld. De achtergrondkenmerken verschillen echter en daardoor wordt berekend dat met 82 procent zekerheid dit een werknemer met een flexibel arbeidscontract is. Er is een kans van 18 procent dat deze baan voor de hele periode een vast contract betreft.

5.3.2 Voorbeeld 2

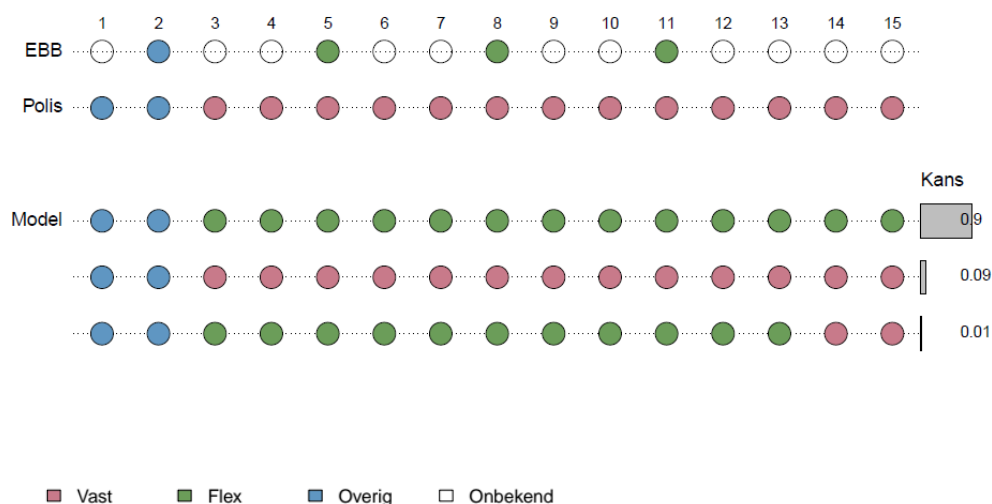


5.3.3 Voorbeeld 3



Ook in voorbeeld 3 is het de grootste kans dat het een vast arbeidscontract betreft. De genoemde situaties maken duidelijk dat de patronen in contractsoort volgens de EBB en de Polis niet de enige variabelen zijn die het model gebruikt om de werkelijke waarde te berekenen. De achtergrondvariabelen zijn tevens van groot belang voor de uiteindelijke vaststelling van de kansen.

5.3.4 Voorbeeld 4



EBB geeft in voorbeeld 4 aan dat er sprake is van een overgang van 'Overig' naar een flexibel contract, terwijl informatie uit de Polis aangeeft dat dit een overgang naar een vaste baan betreft. Het model heeft echter met behulp van de achtergrondkenmerken en de patronen berekend dat de baan met 90 procent zekerheid overgaat in een flexibele baan. Twee andere uitkomsten zijn ook mogelijk maar veel minder waarschijnlijk.

5.4 Het meetmodel voor de 55-jarigen en ouder

In paragraaf 5.2 is aangetoond dat het model van Pavlopoulos en Vermunt en Pankowska et al. verbeterd kon worden. De stappen die daarvoor zijn gezet voor de 25- tot 55-jarigen zijn herhaald voor de 55- tot 75-jarigen. Voor deze leeftijdsgroep is het aantal waarop de modellen geschat worden in 2016, 2017 en 2018 respectievelijk gelijk aan 7104, 6660 en

8468 personen. Iedere stap leverde weer een verbetering van het model op. De BIC verbetert in 2016 van 31873 met 146 parameters naar 11843 met 162 parameters. Voor het verslagjaar 2017 daalt de BIC van 29811 naar 10970 en voor 2018 van 33456 naar 12664. Dat betekent dat voor de 55- tot 75-jarigen een model gebruikt kan worden met dezelfde elementen als voor de 25- tot 55-jarigen. Wel is er een verschil tussen de achtergrondkenmerken die als covariaten in het definitieve model zijn opgenomen met een effect op de overgangskansen van het type arbeidscontract: ten opzichte van het model voor 25- tot 55-jarigen zijn de covariaten 'Softwarecluster' en 'SBI-groep' hier weggelaten, omdat het toevoegen van deze effecten volgens de BIC geen meerwaarde had. Dit verklaart waarom het aantal te schatten parameters in het model bij de twee leeftijdsgroepen verschilt. Op basis van dit model worden de geschatte onnauwkeurigheid van de metingen in de Polis en de EBB, de geschatte verdeling naar contractvorm en de geschatte overgangskansen weergegeven.

5.4.1 Geschatte onnauwkeurigheid van de metingen in de Polis en de EBB op basis van het nieuwe model, 55- tot 75-jarigen

2016	Waargenomen in EBB			Waargenomen in Polis		
	Vast	Flex	Overig	Vast	Flex	Overig
Model	%					
Vast	98,0	0,2	1,8	95,4	4,1	0,5
Flex	31,7	56,3	12,0	21,8	78,2	0,0
Overig	1,3	1,5	97,3	0,0	0,0	100,0
2017	Waargenomen in EBB			Waargenomen in Polis		
	Vast	Flex	Overig	Vast	Flex	Overig
Model	%					
Vast	97,7	0,1	2,2	96,1	3,4	0,4
Flex	28,2	59,5	12,2	15,8	83,6	0,5
Overig	1,3	1,0	97,7	0,0	0,0	100,0
2018	Waargenomen in EBB			Waargenomen in Polis		
	Vast	Flex	Overig	Vast	Flex	Overig
Model	%					
Vast	97,3	0,5	2,2	96,1	3,4	0,5
Flex	27,1	58,3	14,6	14,1	84,9	0,9
Overig	1,2	1,1	97,7	0,0	0,0	100,0

De geschatte onnauwkeurigheden in de meting van de soort arbeidscontract in de Polis en in de EBB hebben bij de 55- tot 75-jarigen hetzelfde patroon als bij de 25- tot 55-jarigen. De onnauwkeurigheden zijn het grootst bij de werknemers die volgens het model een flexibel arbeidscontract hebben. In de EBB antwoordt 27 tot 32 procent dat ze een vast contract hebben en rond de 13 procent dat ze behoren tot de categorie 'Overig'. Bij de Polis komt dezelfde onnauwkeurigheid voor. Van de werknemers die volgens het model een flexibel arbeidscontract hebben is 14 tot 22 procent geregistreerd met een vast arbeidscontract. Verder neemt zowel in de Polis als in de EBB de nauwkeurigheid tussen 2016 en 2018 iets toe. Hier moet wel bij gezegd worden dat de nauwkeurigheid van de

modelschattingen hier relatief laag is, omdat maar weinig ouderen een flexibel arbeidscontract hebben, zoals blijkt uit tabel 5.4.2 hieronder.

5.4.2 Geschatte verdeling naar contractvorm volgens het model, de EBB en de Polis voor 2016, 2017 en 2018, 55- tot 75-jarigen

	2016				2017				2018		
	Model	EBB	Polis		Model	EBB	Polis		Model	EBB	Polis
contractvorm	%				%				%		
Vast	24,3	26,6	24,3		24,5	27,5	24,4		25,0	28,3	24,8
Flex	5,1	3,7	5,0		5,0	3,7	5,0		5,7	3,8	5,7
Overig	70,6	69,8	70,7		70,4	68,9	70,6		69,3	67,8	69,5

De modelschattingen van de frequenties van de verschillende arbeidscontracten heel dicht bij die uit de Polis (tabel 5.4.2). In 2016-2018 vallen die schattingen nagenoeg samen. Ongeveer een kwart heeft een vast contract, een op de twintig een flexibel arbeidscontract en iets meer dan 70 procent is geen werknemer. De EBB-schattingen liggen overigens ook redelijk dichtbij de schattingen van het model en de Polis. De op basis van het model geschatte overgangskansen van een flexibele naar een vast arbeidscontract liggen, net als bij de 25- tot 55-jarigen, onder de schattingen op basis van de Polis en de EBB (tabel 5.4.3). Wel is ook hier de schatting van de Polis dichter bij de modelschatting dan die van de EBB.

5.4.3 Geschatte overgangskansen volgens het nieuwe model, de EBB en de Polis voor 2016, 2017 en 2018, 55- tot 75-jarigen

2016	Vast-Vast	Vast-Flex	Vast-Overig	Flex-Vast	Flex-Flex	Flex-Overig	Overig-Vast	Overig-Flex	Overig-Overig
	%								
Model	97,4	0,3	2,3	3,6	86,7	9,7	0,1	0,8	99,2
Polis	96,5	0,9	2,6	5,1	86,4	8,5	0,2	0,6	99,2
EBB	96,5	0,1	3,5	9,0	81,1	9,9	0,2	0,5	99,3
2017	Vast-Vast	Vast-Flex	Vast-Overig	Flex-Vast	Flex-Flex	Flex-Overig	Overig-Vast	Overig-Flex	Overig-Overig
	%								
Model	97,5	0,5	2,0	3,0	88,3	8,7	0,1	0,7	99,2
Polis	97,2	0,7	2,2	4,3	87,5	8,1	0,1	0,7	99,2
EBB	96,4	0,3	3,3	10,1	80,3	9,6	0,2	0,8	99,0
2018	Vast-Vast	Vast-Flex	Vast-Overig	Flex-Vast	Flex-Flex	Flex-Overig	Overig-Vast	Overig-Flex	Overig-Overig
	%								
Model	97,5	0,5	2,0	2,9	87,3	9,8	0,0	0,9	99,0
Polis	97,2	0,6	2,2	4,1	87,1	8,8	0,1	0,8	99,0
EBB	97,3	0,2	2,5	10,6	78,0	11,5	0,1	0,7	99,1

Samenvattend lijken de resultaten van de 55- tot 75-jarigen sterk op die voor de 25- tot 55-jarigen. Het model is op een paar kleinigheden na hetzelfde voor beide leeftijdscategorieën. Het patroon van de onnauwkeurigheden is ongeveer gelijk, terwijl ook bij de oudere leeftijdsgroep de schattingen van het model van de frequenties van de soort arbeidscontract de Polis- en de modelschattingen meer op elkaar lijken dan de schattingen vanuit de EBB. Het valt te overwegen om bij een vervolg van het onderzoek de modelanalyses van beide leeftijdsgroepen samen te voegen.

5.5 Het meetmodel voor de 15- tot 25-jarigen

Net als bij de 25- tot 55-jarigen en de 55-jarigen en ouder is het best passende model voor jongeren een model waarbij de foutkansen in de Polis serieel zijn gecorreleerd in de tijd en de waarnemingen in de EBB gecorreleerd mogen zijn als het werkelijke type arbeidscontract niet veranderd is ten opzichte van het vorige kwartaal. Wel zijn er, net als bij de 55- tot 75-jarigen, enkele verschillen in de selectie van covariaten ten opzichte van het model voor 25- tot 55-jarigen. Ten eerste is in het best passende model voor jongeren het werkelijke arbeidscontract in de eerste maand niet afhankelijk van het opleidingsniveau. Ten tweede zijn de overgangskansen tussen werkelijke arbeidscontracten in het best passende model voor jongeren niet afhankelijk van opleidingsniveau en 'SBI-groep'.

Bij de jongeren is het aantal personen waarop de modellen geschat zijn in 2016, 2017 en 2018 respectievelijk gelijk aan 4175, 4147 en 4810. Deze aantallen zijn duidelijk kleiner dan bij de andere leeftijdsgroepen. Het best passende model voor jongeren geeft een BIC van 23288 bij 160 te schatten parameters. Ter vergelijking: het beste model van Pavlopoulos en Vermunt (2015) toegepast op jongeren levert een BIC van 59192. Het best passende model geeft ten opzichte van het model van Pavlopoulos en Vermunt een verlies van 12 vrijheidsgraden. Voor de verslagjaren 2017 en 2018 levert het best passende model voor jongeren een BIC van respectievelijk 24205 en 23978, terwijl het model van Pavlopoulos en Vermunt voor die verslagjaren een BIC geven van 60031 en 56290.

In tabel 5.5.1 zijn voor 15- tot 25-jarigen die in het eerste kwartaal van 2016 in de EBB zijn ingestroomd de onnauwkeurigheden weergegeven op basis van het best passende model. De resultaten zijn gewogen met het EBB-jaargewicht van de eerste peiling van de respectievelijke jaren. Eerdere berekeningen lieten overigens zien dat voor de schatting van de onnauwkeurigheden de weging weinig verschil maakt. De weging heeft wel effect op de schatting van de frequenties en de overgangskansen. De tabel laat zien dat zowel in de EBB als in de Polis een aanzienlijk deel van de jongeren die volgens het model een vast arbeidscontract hebben, volgens één van beide bronnen een flexibel arbeidscontract hebben. In de EBB is dat bijna een kwart van de jongeren; in de Polis is dat bijna 45 procent. Het aandeel jongeren dat volgens het model een flexibel contract zou hebben maar volgens één van beide bronnen een vast contract heeft is veel lager: voor de EBB ongeveer 16 procent en voor de Polis ruim 4 procent. Voor jongeren die een flexibel contract hebben komt de Polis dus goed overeen met het model, zeker in vergelijking met de EBB, maar voor de (veel kleinere) groep jongeren met een vast contract is de schatting van het type arbeidscontract in de Polis minder nauwkeurig dan in de EBB.

5.5.1 Onnauwkeurigheden in de EBB en de Polis voor 15- tot 25-jarigen die in het eerste kwartaal van 2016, 2017 en 2018 voor het eerst respondeerden in de EBB

2016	waargenomen in EBB			waargenomen in Polis		
	Vast	Flex	Overig	Vast	Flex	Overig
Model	%					
Vast	75,3	22,8	1,9	55,5	44,5	0,0
Flex	16,1	70,4	13,5	4,3	95,7	0,0
Overig	4,1	12,2	83,7	0,0	0,2	99,8
2017	waargenomen in EBB			waargenomen in Polis		
	Vast	Flex	Overig	Vast	Flex	Overig
Model	%					
Vast	66,9	29,0	4,1	60,7	39,3	0,0
Flex	17,3	69,3	13,4	2,6	97,4	0,0
Overig	3,9	12,7	83,4	0,0	0,4	99,6
2018	waargenomen in EBB			waargenomen in Polis		
	Vast	Flex	Overig	Vast	Flex	Overig
Model	%					
Vast	74,3	23,0	2,7	59,5	40,3	0,2
Flex	17,3	70,0	12,7	3,8	96,2	0,0
Overig	3,5	14,7	81,9	0,0	0,0	100,0

Voor 2017 en 2018 is het beeld vergelijkbaar. Zowel in de Polis als in de EBB worden jongeren die volgens het model een vast contract hebben relatief vaak als flexibel geclassificeerd, terwijl in de EBB daarnaast jongeren met een latent flexibel contract vaak als vast worden geclassificeerd. Wel valt op dat de schattingen van de onnauwkeurigheden nogal fluctueren over de tijd. Het aandeel van de jongeren dat volgens het model een vast contract heeft maar volgens de EBB een flexibel stijgt van 2016 op 2017 van 22,8 procent naar 29,0 procent en daalt in 2018 weer tot 23,0 procent. Ook in de Polis is dit terug te zien: in 2016 bedroeg dit percentage 44,5, in 2018 was dit gedaald tot 40,3 procent. Hierbij moet bedacht worden dat het aandeel jongeren in de potentiële beroepsbevolking relatief laag is. In het eerste kwartaal 2016 bedroeg dit aandeel 16 procent. Het aantal waarnemingen voor jongeren in ons model is om die reden ook veel lager dan voor de 25- tot 55-jarigen en voor de groep van 55 jaar en ouder. Daarbij komt nog dat een vast contract relatief weinig voorkomt bij deze leeftijdscategorie, waardoor het aantal waarnemingen van jongeren met een vast contract (in beide bronnen) eveneens laag is, hetgeen gevolgen kan hebben voor de stabiliteit van de schattingen. Ook het feit dat er over de hele linie en in beide bronnen hoge kansen buiten de diagonaal voorkomen is een aanwijzing dat de modelschattingen bij deze groep relatief meer onzekerheid zullen bevatten dan bij de andere groepen.

Tabel 5.5.2 geeft de geschatte verdeling van vast, flex en overig volgens het model, de EBB en de Polis in 2016, 2017 en 2018. Bij de 25- tot 55-jarigen ligt de modelschatting dichter bij de geobserveerde percentages van de Polis dan bij de geobserveerde percentages van de EBB. De modelschattingen voor jongeren laten een ander beeld zien. Volgens het model had in 2016 17,4 procent van de jongeren een vast contract. Volgens de EBB lag dit op 20,0 procent en volgens de Polis op 11,6 procent. Het percentage jongeren met een vast contract neemt zowel volgens de EBB als volgens de Polis langzaam toe in 2017 en 2018. Volgens het model is er eveneens een langzame toename in 2017 maar neemt het percentage weer af in 2018. Het percentage jongeren met een flexibel contract neemt

5.5.2 Geschatte verdeling naar contractvorm volgens het model, de EBB en de Polis voor 2016, 2017 en 2018, 15- tot 25-jarigen die in het eerste kwartaal van 2016, 2017 en 2018 voor het eerst respondeerden in de EBB

	2016				2017				2018		
	Model	EBB	Polis		Model	EBB	Polis		Model	EBB	Polis
Contract	%										
Vast	17,4	20,0	11,6		19,3	20,7	12,9		18,8	22,2	13,0
Flex	44,3	37,7	50,2		45,3	39,4	51,8		48,5	39,6	54,2
Overig	38,3	42,3	38,3		35,5	40,0	35,3		32,7	38,2	32,8

5.5.3 Geschatte overgangskansen volgens het model, de EBB en de Polis voor 2016, 2017 en 2018, 15- tot 25-jarigen

2016	Vast-Vast	Vast-Flex	Vast-Overig	Flex-Vast	Flex-Flex	Flex-Overig	Overig-Vast	Overig-Flex	Overig-Overig
	%								
Model	91,2	5,3	3,6	2,4	85,0	12,6	1,2	18,7	80,1
Polis	85,2	9,9	4,9	3,5	85,4	11,1	1,5	18,3	80,2
EBB	89,6	5,6	4,8	8,7	79,3	12,0	1,9	16,6	81,5
2017	Vast-Vast	Vast-Flex	Vast-Overig	Flex-Vast	Flex-Flex	Flex-Overig	Overig-Vast	Overig-Flex	Overig-Overig
	%								
Model	88,5	7,5	3,9	3,9	84,2	11,8	1,7	20,2	78,0
Polis	86,5	9,2	4,3	3,9	85,5	10,6	1,5	20,4	78,1
EBB	91,1	4,6	4,3	10,0	79,1	11,0	2,0	16,7	81,3
2018	Vast-Vast	Vast-Flex	Vast-Overig	Flex-Vast	Flex-Flex	Flex-Overig	Overig-Vast	Overig-Flex	Overig-Overig
	%								
Model	90,1	7,1	2,8	1,5	88,1	10,4	1,1	22,3	76,6
Polis	87,6	8,8	3,6	3,8	86,7	9,4	1,7	21,6	76,7
EBB	92,1	4,1	3,8	11,5	79,4	9,1	2,2	19,6	78,2

zowel volgens het model als volgens de EBB en de Polis toe in 2017 en 2018. De toename van het aandeel jongeren dat als werknemer met een vast of flexibel contract werkzaam is weerspiegelt zich in een geschatte afname van het percentage jongeren in de categorie 'Overig'. Volgens de EBB is deze afname iets minder groot dan volgens het model en volgens de Polis.

Evenals bij de andere leeftijdscategorieën is de geschatte kans op een overgang van een flexibel naar een vast contract lager volgens het model dan voor elk van de bronnen (tabel 5.5.3). Volgens het model bedraagt de overgangskans van flex naar vast in 2016 2,4 procent, volgens de Polis 3,5 procent en volgens de EBB 8,7 procent. Met name de overgangskans op basis van de EBB wijkt sterk af van de overgangskans geschat door het model. Verder blijkt de overgangskans volgens de EBB tussen 2016 en 2018 aanzienlijk toe te nemen terwijl die volgens de Polis slechts licht toeneemt en volgens het model sterk daalt. De overgangskans van flex naar de categorie 'Overig' wordt voor alle verslagjaren door het model hoger ingeschat dan door zowel de Polis als de EBB. De overgangskans van vast naar flex ligt volgens de Polis veel hoger dan volgens het model en de EBB: in 2016 op 9,9 procent, volgens het model was dat 5,3 procent en volgens de EBB 5,6 procent. Voor deze overgangskans ligt de schatting volgens de EBB dus dicht bij de schatting van het model. In 2018 is de situatie omgekeerd: de overgang van vast naar flex volgens het model is toegenomen ten opzichte van 2016, voor de Polis is de kans juist afgenomen waardoor in 2018 de schatting van de Polis dicht bij die van het model ligt. Aangezien het niet zo waarschijnlijk is dat een werkgever een vast contract omzet in een flexibel contract, zal het bij de overgang van een vast naar een flexibel contract moeten gaan om een situatie waarbij sprake is van een baanwisseling.

De resultaten voor jongeren in de leeftijdscategorie '15 tot 25 jaar' wijken dus op een aantal punten duidelijk af van de resultaten voor 25- tot 55-jarigen. Ten eerste blijkt dat voor jongeren de kans veel groter is dat ze volgens het model een vast contract en volgens de EBB of Polis een flexibel contract hebben dan voor de 25- tot 55-jarigen. Daarnaast geldt voor jongeren, in tegenstelling tot voor de 25-54 jarigen, dat de schattingen volgens het model niet dicht bij de schattingen op basis van de Polis liggen. De afwijking van de modelschattingen is voor beide bronnen in dezelfde orde van grootte. Tenslotte blijkt dat voor jongeren bepaalde overgangskansen, zoals van vast naar flex en van flex naar overig op basis van het model juist hoger worden ingeschat dan op basis van de EBB. Bovendien lijken de modelschattingen minder stabiel. Eerder werd al opgemerkt dat er bij jongeren sprake is van relatief weinig waarnemingen, dat komt doordat de groep jongeren sowieso veel kleiner is dan de groep 25-54 jarigen en bovendien een vast contract bij jongeren weinig voorkomt waardoor het aantal waarnemingen van jongeren met een vast contract klein is. Dit vraagt om vervolgonderzoek. In de conclusies wordt daar nader op ingegaan.

5.6 De precisering van de ontwikkeling in de onnauwkeurigheden in 2017 en 2018

Om een beeld te krijgen van de ontwikkeling van de onnauwkeurigheden in 2017 en 2018 is de onderzoekspopulatie uitgebreid met personen die in de kwartalen 2, 3, en 4 aan de eerste peiling van de EBB hebben meegedaan. Daardoor zijn er meer waarnemingen (een hogere resolutie). Bovendien kunnen kwartaalcijfers samengesteld worden door deze instromers mee te nemen in de analyses die qua populatie beter overeenkomen met de

kwartaalcijfers die doorgaans op basis van de EBB gepubliceerd worden⁷. Ook zijn de data van 2019 toegevoegd zodat de instromers uit 2018 vijftien maanden gevolgd kunnen worden. Op basis van deze uitgebreidere populatie kan de ontwikkeling van de aandelen vaste en flexibele werknemers en de onnauwkeurigheden in beide bronnen beter in kaart gebracht worden. In dit hoofdstuk is het aandeel van personen die tot de categorie 'Overig' worden gerekend buiten beschouwing gelaten. De reden hiervoor is dat in de eerste plaats de verschillen tussen de Polis en de EBB voor wat betreft de aantallen en aandelen vaste en flexibele werknemers van belang zijn. De categorie 'Overig' is een restcategorie waarover in deze samenstelling (totaal van zelfstandigen, werklozen en inactieven) nooit apart wordt gepubliceerd. Bovendien wordt de Polis door het CBS ook niet als bron gebruikt om over deze groepen (met uitzondering van de zelfstandigen die tevens directeur groot aandeelhouder zijn) te publiceren.

5.6.1 Ontwikkeling aandeel werknemers met een flexibel contract (als percentage van totaal aantal werknemers)



Figuur 5.6.1 laat de ontwikkeling zien van het aandeel werknemers met een flexibel contract volgens het model, volgens de EBB en volgens de Polis-administratie.⁸ De standaardfouten hiervan zijn klein. Om een idee te geven: de standaardfouten van de modelschattingen voor de totale populatie werknemers zijn 0,10 tot en met 0,12 procent, voor de jongeren is dat 0,31 tot en met 0,34 procent. Voor de totale populatie werknemers neemt het geschatte aandeel zowel volgens het model als volgens de Polis-administratie toe, terwijl de EBB een lichte daling laat zien. De schattingen van het aandeel flexibele werknemers volgens de EBB wijkt in de beschouwde periode geleidelijk meer af van het

⁷ Een beperkte groep respondenten uit de EBB is niet terug te vinden in de Basisregistratie personen (BRP), deze groep kan niet gekoppeld worden en is daarom buiten beschouwing gebleven. De modelschattingen van het totaal aantal personen die deel uitmaken van de potentiële beroepsbevolking, dat wil zeggen het aantal werknemers (met vast en flexibel contract) plus de groep 'overig' (zelfstandigen en personen zonder betaald werk) is daarom wat lager dan de aantallen die op Statline worden gerapporteerd.

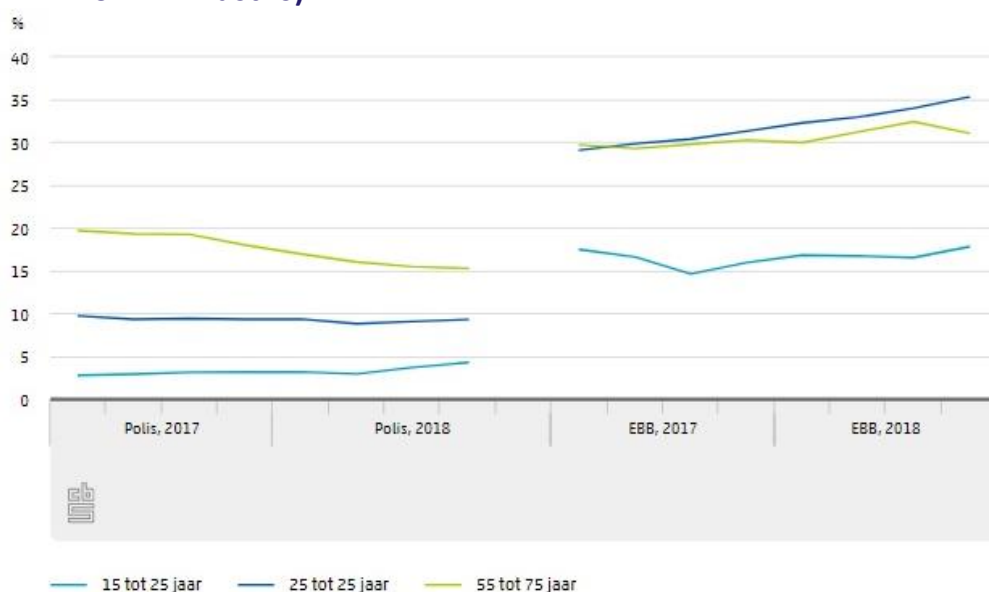
⁸ Het geschatte percentage werknemers met een flexibel contract volgens de EBB komt niet overeen met het percentage werknemers met een flexibel dienstverband zoals af te leiden valt uit Statline. De reden hiervoor is dat hier een iets andere afbakening van flexibele werknemers is gehanteerd. Zie hiervoor ook de bijlage 1.

geschatte aandeel volgens het model. Er zijn wat dit betreft wel duidelijke verschillen tussen leeftijdscategorieën. Bij ouderen stijgt gedurende de periode 2017-2018 het aandeel werknemers met een flexibel contract zowel volgens het model, de Polis als de EBB. Wel lijkt het aandeel flexibele werknemers onder deze groep volgens de EBB wat minder toe te nemen dan volgens de Polis. Opvallend is bovendien dat de ontwikkelingen volgens de Polis vrijwel geheel in lijn zijn met de geschatte ontwikkeling op basis van het model: deze lijnen vallen helemaal over elkaar heen. Voor jongeren onder de 25 jaar laat de EBB een daling zien van het aandeel flexibele werknemers, volgens de Polis is dit aandeel vrij stabiel en volgens het model stijgt het aandeel jongeren met een flexibele contract. Wel ligt het aandeel werknemers met een flexibele arbeidsrelatie volgens de EBB hier wat dichterbij de modelscores. Voor werknemers tussen de 25 en 55 jaar laten zowel het model als de Polis een stijgend aandeel flexibele werknemers zien, terwijl dit aandeel volgens de EBB gedurende 2017 en 2018 vrijwel stabiel blijft.

Ontwikkeling onnauwkeurigheden

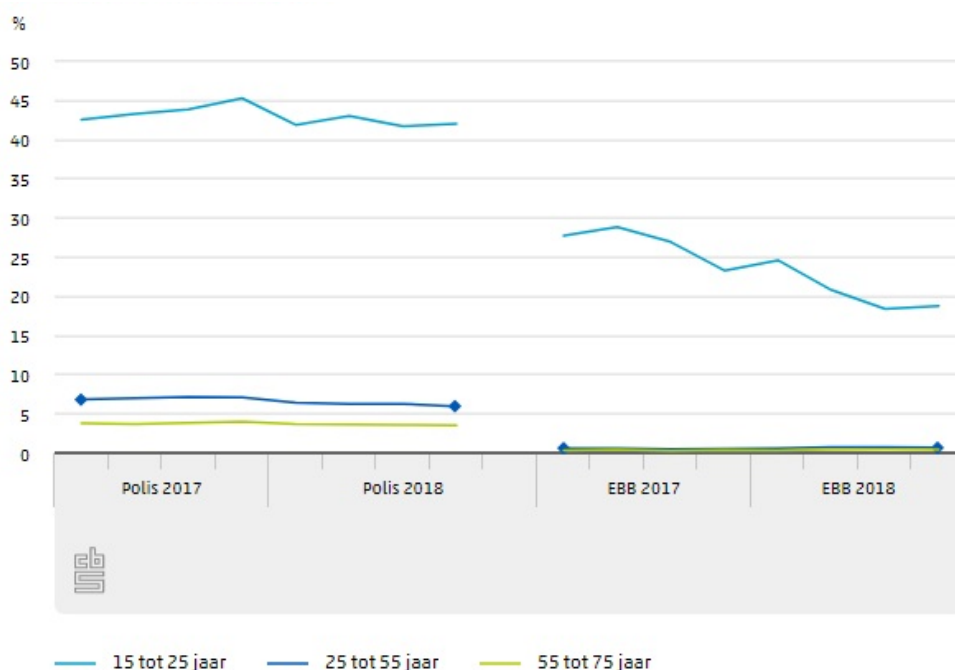
De onnauwkeurigheden in de EBB nemen toe, in het bijzonder voor de groep werknemers in de leeftijdscategorie '25 tot 55 jaar' (figuur 5.6.2). De kans dat een werknemer die volgens het model een flexibel contract heeft op basis van de EBB wordt geclassificeerd als een werknemer met een vast dienstverband neemt toe van 29 procent in het eerste kwartaal van 2017 tot 35 procent in het vierde kwartaal van 2018. In de Polis blijft deze kans vrij stabiel in de tijd en ligt rond de 9 procent. Voor de oudere werknemers lijken de onnauwkeurigheden in de EBB minder te zijn toegenomen. De kans dat iemand in de leeftijdscategorie '55 tot 75 jaar' die volgens het model een flexibel contract heeft volgens de EBB als vaste werknemer wordt geclassificeerd was in 2017 eerste kwartaal 29 procent en in 2018 vierde kwartaal 31 procent. In de Polis neemt dit voor de oudere werknemers af. Voor jongeren tenslotte, is de kans in de Polis op deze onnauwkeurigheid zeer laag maar vanaf 2018 wel wat toegenomen. In de EBB is dit type onnauwkeurigheid veel groter, maar wel vrij stabiel (rond de 17 procent).

5.6.2 Ontwikkeling van de geschatte onnauwkeurigheden (kans dat een flexibel contract volgens het model volgens de Polis of EBB vast is)



De kans dat een werknemer volgens het model een vaste baan heeft maar wordt geclassificeerd als een flexibele werknemer is voor werknemers ouder dan 25 jaar in de EBB klein en stabiel door de tijd: deze kans schommelt rond de 0,6 procent (figuur 5.6.3). In de Polis komt dit type onnauwkeurigheid voor werknemers tussen de 25 en 55 jaar vaker voor maar er is wel een dalende trend zichtbaar: in 2017 eerste kwartaal bijna 7 procent, in 2018 vierde kwartaal nog geen 6 procent. Voor werknemers van 55 jaar of ouder schommelt deze kans tussen de 3,5 en 4 procent. Voor jongeren is de kans dat ze volgens één van beide bronnen een flexibel contract hebben terwijl het volgens het model om een vast contract gaat veel hoger. In de Polis schommelt deze kans rond de 43 procent. In de EBB is deze kans lager dan in de Polis en bovendien daalt deze kans van bijna 28 procent in 2017 eerste kwartaal naar 19 procent in 2018 laatste kwartaal.

5.6.3 Ontwikkeling van de geschatte onnauwkeurigheden (kans dat een vast contract volgens het model volgens de Polis of EBB flexibel is)

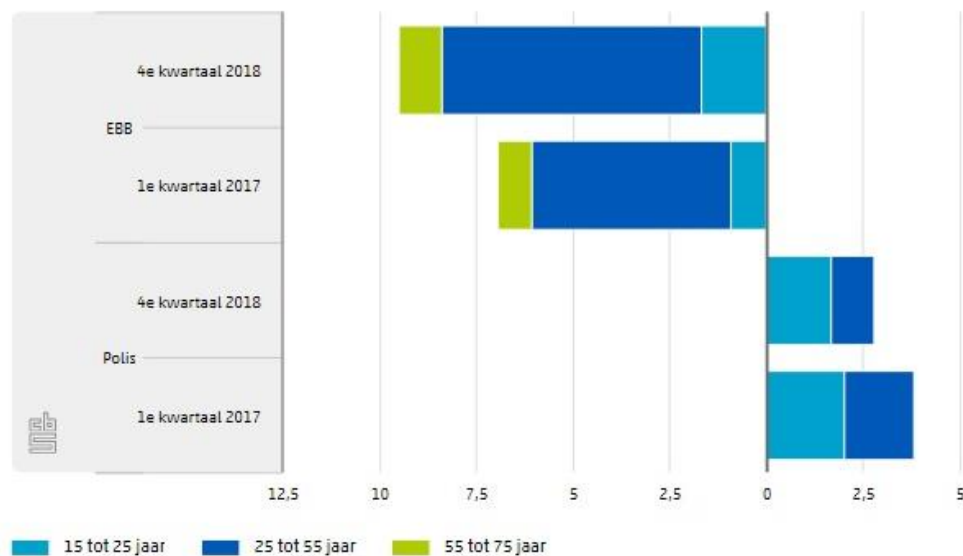


De decompositie van de verschillen naar leeftijdscategorie

De schattingen van het aandeel flexibele werknemers volgens de EBB en de Polis lopen het meest uiteen voor werknemers jonger dan 25 jaar. Aangezien jongeren volgens beide databronnen de grootste kans op een flexibel contract hebben, is het aandeel jongeren in de totale groep flexibele werknemers eveneens hoog, volgens de EBB 45 procent, volgens de Polis 40 procent. De kans dat een flexibel contract (volgens het model) in de Polis of in de EBB als vast wordt geclassificeerd is voor jongeren echter lager dan voor werknemers in de oudere leeftijdscategorieën. Om een beeld te krijgen in welke mate de verschillende leeftijdscategorieën bijdragen aan de over- of onderschatting van het aandeel flexibele werknemers volgens EBB of Polis is een decompositie gemaakt van de over- en onderschatting naar leeftijdscategorie. De 'onderschatting' van het aandeel flexibele werknemers op basis van de EBB ten opzicht van de modelschatting bedroeg in 2018 vierde kwartaal bijna 10 procentpunt, de 'overschatting' op basis van de Polis bedroeg nog

geen 3 procentpunt. Figuur 5.6.4 laat zien dat de 'onderschatting' op basis van de EBB grotendeels kan worden toegerekend aan werknemers tussen de 25 en 55 jaar oud. De bijdrage van de jongere werknemers is veel kleiner. Bij de Polis is de situatie omgekeerd: het grootste deel van de 'overschatting' kan worden toegerekend aan de jongste leeftijdscategorie. Voor zowel de Polis als de EBB geldt dat de bijdrage van de oudste werknemers gering is. De onderschatting van het aantal flexibele werknemers volgens de EBB is toegenomen ten opzichte van 2017 eerste kwartaal, de overschatting volgens de Polis is juist afgenomen. Deze ontwikkelingen kunnen noch bij de Polis noch bij de EBB duidelijk aan een specifieke leeftijdscategorie worden toegerekend.

5.6.4 Over- en onderschatting van het aandeel flexibele werknemers volgens Polis en EBB ten opzichte van de modelschatting naar leeftijdscategorie



6. Waar hangen de onnauwkeurigheden mee samen?

De onderscheiden covariaten worden in verband gebracht met de onnauwkeurigheden in de meting in de Polis en de EBB om te begrijpen welke samenhang er is. Op grond van die kennis kan ook mogelijk actie worden ondernomen die de onnauwkeurigheden in de meting van arbeidscontract in de Polis en de EBB zouden kunnen verminderen. Daarvoor wordt een tweetrapsprocedure ingezet. In de eerste stap worden de onnauwkeurigheden zo precies mogelijk geschat met behulp van een Hidden Markov Model. De uitkomsten daarvan zijn beschreven in de paragrafen 5.2, 5.4 en 5.5 voor respectievelijk de 25- tot 55-, de 55- tot 75- en de 15- tot 25-jarigen. In de tweede stap worden deze onnauwkeurigheden in verband gebracht met de variabelen die in de eerdere analyses in

hoofdstuk 4 van invloed bleken op de verschillen tussen metingen in de Polis en de EBB. Daarvoor worden multinomiale logistische regressieanalyses uitgevoerd. De uitkomsten dienen ook in lijn te zijn met de resultaten van de eerder uitgevoerde MLR voor de verschillende metingen in de Polis en de EBB. Als dat niet het geval blijkt is dat een argument tegen de gebruikte meetmodellen.

De precieze procedure is opgenomen in bijlage 2. Hiervoor worden de modeluitkomsten voor de jaren 2017 en 2018 gebruikt zoals gerapporteerd in paragraaf 5.6. De analyses worden afzonderlijk uitgevoerd voor de drie leeftijdscategorieën en worden geconcentreerd op de metingen die in de HMM-analyse de grootste onnauwkeurigheden te zien gaven. In bijlage 9 zijn de gedetailleerde resultaten weergegeven.

6.1 Waar hangen de onnauwkeurigheden in de meting van arbeidscontract in de EBB en de Polis mee samen voor de 25- tot 55-jarigen

Voor de 25- tot 55-jarigen is de grootste onnauwkeurigheid daar waar volgens het model werknemers een flexibel contract hebben en in de EBB 29 tot 36 procent antwoordt dat sprake is van een vast contract. De samenhang met het hebben van dezelfde werkgever volgens de EBB is daarbij het grootst, gevolgd door de economische activiteit en de grootte van het bedrijf, baanduur en contracturen. Omdat de ontwikkeling in 2017 en 2018 relevant is, wordt ook de invloed van kwartaalnummer beoordeeld.

Bij de werknemers die volgens de EBB dezelfde werkgever houden is de kans dat ze volgens het model een flexibel en volgens de EBB een vast arbeidscontract hebben veel groter dan als ze van werkgever zijn gewisseld of dat dit onbekend is. In sommige branches is de kans op deze onnauwkeurigheden kleiner, bijvoorbeeld in de 'Arbeidsbemiddeling, uitzendbureaus en personeelsbeheer' en 'Landbouw, bosbouw en visserij'. Daarentegen is de kans weer groter bij vervoerders die dat via het water of de lucht doen. Hoe kleiner het bedrijf, hoe langer iemand de baan al heeft en hoe groter het aantal contracturen, hoe groter de kans dat deze onnauwkeurigheid optreedt. Bovendien neemt de kans op deze onnauwkeurigheid langzamerhand toe in de loop van 2017 en 2018, zonder dat dit verklaard wordt door de andere genoemde covariaten.

Het op een na grootste verschil tussen de waarnemingen en de modelschatting is dat in de Polis een vast contract wordt waargenomen, maar dit volgens het model een flexibel contract is. Dat komt in ongeveer 10 procent van de gevallen voor. De belangrijkste variabelen die daarmee samenhangen zijn softwarecluster, baanduur, 'SBI-groep', bedrijfsgrootte en of men volgens de Polis dezelfde baan heeft als in de vorige meting. In vergelijking met softwarecluster 1, heeft softwarecluster 2 een grotere kans terwijl de overige clusters juist kleinere kansen opleveren op deze onnauwkeurigheid. Bij baanduur treedt het tegenovergestelde effect op als bij de kans op EBB vast en model flex: de kans op deze onnauwkeurigheid is juist bij langere baanduren kleiner. Bij bedrijfsgrootte is dat niet het geval: de onnauwkeurigheid is ook in de Polis groter bij de kleinere en middelgrote bedrijven. Ook voor de 'SBI-groep' geldt dit met uitzondering van het vervoer over water en luchtvaart: bij de EBB heeft dit een grotere en bij de Polis een kleinere kans. Voor wat betreft migratieachtergrond zijn er slechts twee significante effecten: bij werknemers met

een Antilliaanse achtergrond komt deze onnauwkeurigheid vaker en bij een Turkse achtergrond minder vaak voor.

Een ongeveer even belangrijke onnauwkeurigheid is waar het model schat dat het een flexibel arbeidscontract betreft, terwijl de EBB op de categorie 'Overig' uitkomt. In ongeveer 7 procent van de gevallen treedt dit op. De categorie 'Overig' betreft een verzameling van verschillende subcategorieën zoals bijvoorbeeld zelfstandigen, werklozen, arbeidsongeschikten en mensen die niet actief zijn op de arbeidsmarkt. De grootste invloed op deze onnauwkeurigheid heeft het aantal contracturen. Voor werknemers met een contract van minder dan 30 uur is de kans op deze onnauwkeurigheid maar liefst vier keer zo groot als voor de mensen met meer contracturen. Ook of men van werkgever is veranderd, bij kleinere bedrijven en bij mannen levert een grotere kans hierop op. Bij de detailhandel en zorginstellingen komt dit juist vaker voor.

De in omvang vierde onnauwkeurigheid is waar in de Polis een flexibel arbeidscontract wordt gemeten en het model schat dat het een vast contract is. Dat komt in 6 à 7 procent van de gevallen voor. Voor deze onnauwkeurigheid zijn vooral softwarecluster, bedrijfsgrootte, leeftijdsklasse, 'SBI-groep' en baanduur verantwoordelijk. De kans op de onnauwkeurigheid dat in de Polis een flexibel contract staat geregistreerd en het model een vast contract schat komt vaker voor bij de bedrijven die software gebruiken uit de clusters 3, 4 of 5. Dat is bijna het spiegelbeeld van de resultaten bij de onnauwkeurigheid dat de Polis een vast contract en het model een flexibel contract schat. Bij bedrijfsgrootte geldt ook hier dat de onnauwkeurigheid vaker voorkomt bij het midden- en kleinbedrijf dan bij de grote bedrijven. Voor het eerst is er ook een effect van leeftijd: deze onnauwkeurigheid komt vaker voor bij de 25- tot 40-jarigen dan bij de 40- tot 55-jarigen. Ook de 'SBI-groep' is van belang. In de detailhandel en bij de arbeidsbemiddeling, uitzendbureaus en personeelsbeheer komt het vaker voor, terwijl het bij de financiële instellingen juist minder vaak voorkomt.

Samenvattend. Voor de 25- tot 55-jarigen geldt dat de belangrijkste onnauwkeurigheid in de EBB in de meting van het arbeidscontract is dat in de EBB een vast arbeidscontract wordt waargenomen en dat het model een flexibel contract vermoedt. Deze onnauwkeurigheid komt vaak voor bij werknemers die dezelfde werkgever houden, die een baan hebben die al langer dan drie jaar duurt, in het midden- en kleinbedrijf en in sommige SBI-groepen. De belangrijkste onnauwkeurigheden in de Polis zijn dat vast wordt waargenomen en het model flexibel schat en andersom. Welke software wordt gebruikt is daarbij van het grootste belang. Hoe dit geïnterpreteerd moet worden is echter niet duidelijk omdat er weinig kennis is over de kenmerken van die softwarepakketten. Het is mogelijk dat het eigenlijk staat voor niet-gemeten variabelen die samenhangen met de softwarepakketten die gebruikt worden. Verder is duidelijk dat ook bij de onnauwkeurigheden in de Polis de grootte van het bedrijf een rol speelt: het komt veel vaker voor in het midden- en kleinbedrijf dan in de grotere bedrijven.

6.2 Waar hangen de onnauwkeurigheden in de meting van arbeidscontract in de EBB en de Polis mee samen voor de 55- tot 75-jarigen

Voor de ouderen zijn drie onnauwkeurigheden relevant. Deze komen grotendeels overeen met die van de 25- tot 55-jarigen. Het gaat dan om dat in de EBB of in de Polis een vast arbeidscontract wordt waargenomen en dat het model uitkomt op een flexibel arbeidscontract. De derde is dat de EBB 'Overig' scoort en het model schat dat het gaat om een flexibel arbeidscontract. Deze drie onnauwkeurigheden komen in respectievelijk 32, 22 en 12 procent van de gevallen voor.

De in omvang belangrijkste onnauwkeurigheid is die waarin de EBB een vast arbeidscontract waarneemt en het model een flexibel arbeidscontract schat. Daarvoor is de economische activiteit van het bedrijf een belangrijke verklarende factor. Vooral bij de arbeidsbemiddeling, uitzendbureaus en personeelsbeheer, vervoer over water en luchtvaart en onderwijs- en overheidsinstellingen komt deze onnauwkeurigheid minder vaak voor. Een tweede belangrijke voorspeller hiervan is of men dezelfde werkgever heeft als bij de vorige meting in de EBB. Is dat het geval dan is de kans dat in de EBB een vaste en volgens het model een flexibel arbeidscontract wordt gemeten veel groter. Ook bedrijfsgrootte, contracturen en baanduur zijn van belang: de onnauwkeurigheid komt vaker voor in het midden- en kleinbedrijf, bij werknemers met een contract van meer dan 30 uur en baanduren van meer dan 36 maanden. De kans op deze onnauwkeurigheid loopt in de twee jaar die zijn onderzocht ook bij de ouderen op, maar niet monotoon zoals bij de 25- tot 55-jarigen. De ouderen laten wat schokken hierin zien. Ook de migratieachtergrond heeft een kleine invloed: de kans is groter bij werknemers met een Turkse of Antilliaanse en kleiner bij werknemers met een Marokkaanse migratieachtergrond.

De tweede relevante onnauwkeurigheid bij de ouderen betreft dat in de Polis een vast contract wordt waargenomen, terwijl het model een flexibel contract schat. Ook deze onnauwkeurigheid hangt net als bij de 25- tot 55-jarigen samen met softwarecluster, 'SBI-groep', baanduur, bedrijfsgrootte en of een werknemer dezelfde baan heeft als bij de vorige meting. De interpretatie is echter niet altijd geheel dezelfde. De onnauwkeurigheid treedt vaker op bij softwarecluster 2 en minder vaak bij softwareclusters 3, 4 en 5. Ook de economische activiteit hangt samen met deze onnauwkeurigheid: deze komt veel minder vaak voor bij de arbeidsbemiddeling, uitzendbureaus en personeelsbeheer en minder vaak bij horeca en vervoer over water en luchtvaart. De kans dat de Polis een vast contract meet en het model een flexibel contract schat is groter bij banen die korter duren dan 36 maanden, in kleine bedrijven, en als men dezelfde baan heeft als bij de vorige meting. Het komt minder vaak voor bij werknemers met een Antilliaanse en vaker bij werknemers met een Surinaamse migratieachtergrond.

De laatste relevante onnauwkeurigheid bij de ouderen is dat de EBB 'Overig' scoort en het model schat dat het gaat om een flexibel arbeidscontract. De economische activiteit van het bedrijf hangt het hoogste samen met deze onnauwkeurigheid. Het komt vooral minder vaak voor bij de detailhandel, horeca en zorginstellingen en juist vaker bij overheids- en onderwijsinstellingen. Ook of men dezelfde werkgever heeft als bij de vorige meting in de EBB is van belang. Voor werknemers die van baan zijn veranderd geldt dat zij een veel grotere kans hebben op deze onnauwkeurigheid. Het komt ook vaker voor in kleine bedrijven, bij werknemers met de contract van minder dan 30 uur, bij hoger opgeleiden,

de 65- tot 75-jarigen en werknemers met een Surinaamse of Turkse migratieachtergrond. Het komt minder vaak voor bij werknemers met een baanduur van meer dan 36 maanden en werknemers met een Antilliaanse migratieachtergrond.

Samenvattend. Bij de ouderen komen bij benadering dezelfde onnauwkeurigheden voor als bij de 25- tot 55-jarigen. De belangrijkste onnauwkeurigheid in de meting van het arbeidscontract is dat in de EBB een vast arbeidscontract wordt waargenomen en dat het model een flexibel contract vermoedt. Daarna dezelfde maar dan voor de waarneming van een vaste relatie in de Polis en tot slot een EBB-meting 'Overig' die volgens het model flexibel zou moeten zijn. Ook de samenhang tussen de covariaten vertonen grote overeenkomsten met die bij de 25- tot 55-jarigen. Een rode draad is dat er meer fouten voorkomen in kleine bedrijven en dat de door het bedrijf gebruikte software hoog samenhangt met de onnauwkeurigheden in de meting van de Polis.

6.3 Waar hangen de onnauwkeurigheden in de meting van arbeidscontract in de EBB en de Polis mee samen voor de 15- tot 25-jarigen

Voor jongeren is de grootste onnauwkeurigheid bij hen die volgens het model een vast contract en volgens de Polis een flexibel arbeidscontract hebben. Deze onnauwkeurigheid varieert voor de verslagjaren 2016 tot en met 2018 van 42 tot 45 procent. Deze kans hangt samen met het softwarepakket dat is gebruikt voor de aangifte. Softwarepakketten ingedeeld in de clusters 3, 4 en 5 laten een verhoogde kans zien ten opzichte van pakketten in cluster 1, softwarepakketten in cluster 2 een iets verlaagde kans. Verder komt de onnauwkeurigheid vaker voor bij een baanduur korter dan 36 maanden en vaker als er ten opzichte van de vorige maand sprake is geweest van een baanwisseling. Ook de bedrijfstak lijkt er toe te doen, bij uitzendbureaus, detailhandel en bij de overheid en onderwijs komt deze onnauwkeurigheid vaker voor en bij landbouw juist minder vaak.

Ook bij de EBB komt het relatief vaak voor dat een flexibel contract wordt gemeten terwijl het model een vast contract schat: variërend van 18 tot 28 procent. Net als bij de Polis komt bij de EBB dit type onnauwkeurigheid vaker voor als er sprake is van een kortere baanduur (minder dan 36 maanden) en als er sprake is van een baanwisseling. Ook komt deze onnauwkeurigheid relatief vaak voor als er in de vorige peiling geen sprake was van een werkgever (niet werkzaam of zelfstandig). Bij werkenden die via een uitzendbureau werkzaam zijn komt deze onnauwkeurigheid eveneens vaker voor. Opvallend is dat bij de EBB, in tegenstelling tot de Polis, deze onnauwkeurigheid juist vaker bij de landbouw voorkomt. Ook in de horeca komt deze onnauwkeurigheid vaker voor. Andere factoren die van belang zijn: het aantal contracturen, bij een contractduur van minder dan 30 uur per week, de bedrijfsgrootte (vaker bij midden- en kleinbedrijf dan bij grootbedrijf) en leeftijd (minder vaak bij degenen die 19 jaar of ouder zijn). Tenslotte neemt dit type onnauwkeurigheid bij de EBB in de tijd af.

Het andere type onnauwkeurigheid, namelijk dat er volgens het model sprake is van een flexibel dienstverband maar volgens de bron van een vast contract komt bij jongeren minder vaak voor. Bij de Polis is deze onnauwkeurigheid laag (3 à 4 procent), bij de EBB aanzienlijk hoger (15 à 18 procent). Wel is er bij de Polis sprake van een toename van deze kans tussen 2017 en 2018. Bij de Polis hangt deze kans wederom samen met het

softwarecluster. Bij softwarepakketten die tot cluster 2 worden gerekend is er sprake van een verhoogde kans (ten opzichte van cluster 1), bij de softwarepakketten die tot cluster 3, 4 en 5 worden gerekend een verlaagde kans. Verder lijken ook bedrijfstak en bedrijfsgrootte wederom een rol te spelen. Dit type onnauwkeurigheid komt vaker voor bij het klein- en middenbedrijf dan bij het grootbedrijf. Sectoren waar deze onnauwkeurigheid naar verhouding minder vaak voorkomt zijn de uitzendbranche, zorginstellingen en de detailhandel. Baankenmerken zijn eveneens van belang: bij kleine banen van minder dan 30 contracturen per week komt het vaker voor, bij banen van korter dan 36 maanden minder vaak. Tenslotte doet voor deze onnauwkeurigheid ook herkomst er toe: bij personen met een Marokkaanse achtergrond komt het minder vaak voor dan bij personen met een Nederlandse achtergrond.

De kans dat volgens het model sprake is van een flexibel en volgens de EBB van een vast contract komt net als bij de Polis vaker voor bij het kleinbedrijf en minder vaak bij het grootbedrijf. Het komt niet vaker voor bij het middenbedrijf in tegenstelling tot bij de Polis. Een andere overeenkomst is dat deze onnauwkeurigheid bij de EBB, evenals bij de Polis, minder vaak voorkomt bij uitzendbureaus. Ook bij de landbouw is deze onnauwkeurigheid in de EBB zeldzamer. Bij de overheid en het onderwijs komt dit type onnauwkeurigheid juist vaker voor. Verder komt dit type onnauwkeurigheid in de EBB minder vaak voor als er sprake was van een wisseling van werkgever ten opzichte van het voorgaande kwartaal of als er voorheen geen werkgever was. Opvallend is verder dat kleine banen van minder dan 30 uur minder vaak een dergelijke onnauwkeurigheid geven dan grotere banen. Tenslotte spelen ook persoonskenmerken een rol: de kans dat er volgens de EBB sprake is van een vast en volgens het model van een flexibel contract is groter voor jongeren van 19 tot 25 jaar dan voor jongeren van 15 tot 18 jaar. Bij personen met een Nederlandse achtergrond komt dit type onnauwkeurigheid minder vaak voor dan bij personen met een migratieachtergrond, met name (voormalige) Nederlandse Antillen, Marokko en Turkije.

7. Conclusies en discussie

In dit project worden de oorzaken onderzocht van de uiteenlopende uitkomsten van de Polis en de EBB over het aantal werknemers met een flexibele en vast arbeidscontract. Het onderzoek richt zich op het beantwoorden van de volgende vragen:

- Hoe groot is de onnauwkeurigheid in de meting van de vaste en flexibele arbeidscontracten van werknemers in de Polisadministratie (Polis) en in de Enquête beroepsbevolking (EBB)?
- Hoe heeft de onnauwkeurigheid in de meting van de vaste en flexibele arbeidscontracten zich ontwikkeld in 2017 en 2018 en kan daarmee het verder uiteenlopen van de schattingen in de Polis en de EBB worden verklaard?
- Waardoor worden de (ontwikkelingen in de) onnauwkeurigheden verklaard?

7.1 De ontwikkeling van de meetmodellen

De meting van de vaste en flexibele arbeidscontracten in de Polis en de EBB verschilt, ook nadat de populaties van beide bronnen zijn gelijkgeschakeld en is gecorrigeerd voor het gebruik van verschillende eenheden (banen vs. werknemers). Door gebruik te maken van Hidden Markov Modellen (HMM) is geschat hoe groot de onnauwkeurigheden in de Polis en de EBB zijn. Uitgangspunt van de modellen was het model dat in eerste instantie is ontwikkeld door Pavlopoulos en Vermunt (2015) en daarna is gerepliceerd door Pankowska et al. (2018) voor de 25- tot 55-jarigen. Dat model is verbeterd door toe te staan dat de onnauwkeurigheden in de meting van de Polis en de EBB serieel gecorreleerd zijn voor zover sprake is van dezelfde baan. Daarnaast zijn covariaten toegevoegd aan het model: een aantal variabelen die de verschillende kansen op een vast of flexibel arbeidscontract beïnvloeden. Deze covariaten zijn geselecteerd op basis van een Multinomiale Logistische Regressieanalyse (MLR), waarin de consistentie in de meting tussen EBB en Polis de afhankelijke variabele was. Alleen variabelen die een belangrijke invloed hebben op die inconsistenties zijn als covariaat in de HMM-analyse opgenomen. Dat heeft geleid tot een model dat goed bij de gegevens past. Deze procedure is herhaald voor de 55- tot 75- en de 15- tot 25-jarigen. Voor beide leeftijdsgroepen geldt dat bijna hetzelfde model als best passend model wordt gevonden als bij de 25- tot 55-jarigen. Alleen de effecten van enkele covariaten blijken niet relevant te zijn.

De toetsing of de uitkomsten op basis van het model vertrouwd kunnen worden kent een technische en een inhoudelijke kant. Voor de technische toetsing wordt de mate waarin het model in staat is de oorspronkelijke kruistabel te reproduceren (passendheid) en de stabiliteit van het model beoordeeld. Voor de mate van passendheid van het model is de BIC gebruikt. De uitbreiding van het model zoals in dit onderzoek voorgesteld ten opzichte van de eerdere modellen van Pavlopoulos en Vermunt (2015) en Pankowska et al. (2018) leidt voor de 25- tot 55-jarigen tot een flinke reductie van de BIC. Daarnaast is het voor het verslagjaar 2016 ontwikkelde model toegepast op de verslagjaren 2017 en 2018. Dat leverde ook een goede passendheid op. Vervolgens blijkt op een kleinigheid na het model ook goed te passen bij de gegevens voor de 55- tot 75-jarigen voor dezelfde drie verslagjaren. Door deze stabiliteit zowel wat betreft verslagjaar als leeftijdscategorie neemt het vertrouwen in het model toe. Het model voor de 15- tot 25-jarigen is echter minder stabiel door het relatief geringe aantal waarnemingen. Daardoor variëren de resultaten meer dan bij de andere twee leeftijdscategorieën. Gezien het feit dat jongeren een belangrijk deel van de flexibele arbeidscontracten bezitten is het belangrijk om dit in het achterhoofd te houden. Als een model voor de drie leeftijdsgroepen ongeveer hetzelfde is, wil dat nog niet zeggen dat de uitkomsten die het model oplevert ook hetzelfde zijn. De resultaten van de jongeren wijken flink af van de twee andere leeftijdsgroepen.

Bij de inhoudelijke toetsing gaat het om de vraag of de gevonden modelresultaten van de HMM-analyses aansluiten bij de verwachtingen die zijn besproken in hoofdstuk 3. Dit hoeft niet één op één het geval te zijn, omdat de verwachtingen ook niet altijd heel goed onderbouwd zijn. Enkele verwachtingen komen niet uit, vooral niet bij de jongeren, maar hiervoor zijn goede alternatieve verklaringen mogelijk. De samenhang tussen de onnauwkeurigheden zoals die met het model zijn berekend en de covariaten bieden een verdere toetsingsmogelijkheid. Als deze uitkomsten goed aansluiten bij die van de eerdere MLR-analyses neemt het vertrouwen in het model verder toe. Ook deze resultaten zijn

bemoedigend. Weliswaar zijn de afhankelijke variabelen anders gekozen maar de uitkomsten van beide analyses zijn voor een belangrijk deel met elkaar in overeenstemming. Er zijn ook uitzonderingen: leeftijd is een belangrijke variabele in de MLR-analyses, maar is minder relevant als covariaat in hoofdstuk 6.

Ook al sluiten de uitkomsten van de HMM-analyses in hoofdstuk 5 goed aan bij de inhoudelijke verwachtingen en kan het model ook de technische toetsing doorstaan, toch is het goed een slag om de arm te houden over de precieze betekenis van deze onnauwkeurigheden en de latente klassen waarop deze onnauwkeurigheden zijn gebaseerd. Er valt niet met zekerheid vast te stellen of de latente klassen goed overeenkomen met de 'werkelijke contractvariabele'. De betekenis van de meting in de EBB is enigszins anders dan die in de Polis, omdat in de EBB ook informele afspraken mee kunnen tellen in de bepaling of een arbeidsovereenkomst vast is of niet. Weliswaar komt dit waarschijnlijk niet zo vaak voor in Nederland, zeker is dat echter niet. Daarom zal door de latente klassen als echte contractvariabele op te vatten waarschijnlijk een zekere mate van vertekening moeten worden geaccepteerd. Onduidelijk is hoe groot die vertekening precies is. De meer definitieve bevestiging van de uitkomsten van de HMM-modellen kunnen wellicht bij een experimentele toetsing in een veldwerkonderzoek worden gevonden.

7.2 De grootte van de onnauwkeurigheden in de meting van arbeidscontract van werknemers in de EBB en in de Polis

Het model laat zien dat er onnauwkeurigheden zijn in de Polis en in de EBB. De omvang is beoordeeld aan de hand van het verschil tussen schattingen uit de oorspronkelijke bronnen en die volgens het model voor de verslagjaren 2016, 2017 en 2018. Daarbij is gekeken naar de meting op één bepaald moment en naar de overgangen tussen verschillende tijdstippen. Voor de 25-jarigen en ouder blijkt dat de Polis meestal nauwkeuriger waarneemt dan de EBB. Bij de jongeren is het beeld anders. Daar is de onnauwkeurigheid in de Polis groter dan die in de EBB.

Bij de 25-jarigen en ouder is de grootste onnauwkeurigheid in de EBB die waarin de EBB een vast contract waarneemt en het model schat dat het gaat om een flexibel contract. In 27 tot 33 procent van de gevallen, afhankelijk van het verslagjaar en de leeftijdscategorie, vind je deze onnauwkeurigheid in de EBB terug. Dat leidt ertoe dat de EBB sterker afwijkt van de modelschattingen dan de Polis, zowel in de schatting van het percentage dat een flexibel arbeidscontract heeft als in de schatting van het percentage dat in drie maanden de overgang van een flexibel naar een vast contract maakt. Voor wat betreft het laatste wordt het eerder gevonden resultaat van Pavlopoulos en Vermunt (2015) en Pankowska et al. (2018) nogmaals bevestigd dat het model een lager percentage overgangen schat dan de afzonderlijke bronnen.

Verder bevat de EBB nog een substantiële onnauwkeurigheid voor de 25-jarigen en ouder. Dan gaat het om ongeveer 7 procent bij de 25- tot 55-jarigen waarin het model schat dat het gaat om een flexibel contract, terwijl in de EBB de categorie 'Overig' wordt gestoord. Bij de 55-plussers is deze onnauwkeurigheid nog wat groter: 12 tot 15 procent. Het is wat lastig om deze onnauwkeurigheden te interpreteren vanwege het feit dat de categorie 'Overig' een restcategorie is.

Ook de Polis bevat onnauwkeurigheden voor de 25-jarigen en ouder, zij het dat deze onnauwkeurigheden veel kleiner zijn dan die in de EBB. Ook hier is de grootste onnauwkeurigheid die waarin de Polis een vast contract waarneemt en het model een flexibel contract schat. Bij de 25- tot 55-jarigen is dat de laatst onderzochte drie jaren ongeveer 11, 9 en 8 procent, en bij de ouderen 22, 16 en 14 procent. Op basis van deze modelresultaten kun je concluderen dat deze onnauwkeurigheid in de Polis de laatste drie onderzochte verslagjaren kleiner wordt. Verder is er nog een onnauwkeurigheid van belang: daar waar de Polis een flexibel contract waarneemt terwijl het model schat dat het gaat om een vast contract. Voor de 25- tot 55-jarigen is de kans hierop twee keer zo groot als voor de 55-plussers (respectievelijk ongeveer 7 en 3,5 procent).

Bij de jongeren zijn patronen anders. Bij hen is de kans dat de Polis een flexibel arbeidscontract waarneemt terwijl het model een vast contract schat erg groot, groter dan bij de EBB. Bovendien is die kans ook veel groter dan bij de oudere leeftijdscategorieën. Daarnaast liggen voor jongeren de schattingen volgens het model niet dicht bij de schattingen op basis van de Polis, dit in tegenstelling tot bij de 25- tot 55-jarigen. De afwijking van de modelschattingen is voor beide bronnen in dezelfde orde van grootte. Tenslotte worden voor jongeren bepaalde overgangskansen op basis van het model juist hoger ingeschat dan op basis van de EBB, zoals van vast naar flex en van flex naar overig.

7.3 De ontwikkeling van de onnauwkeurigheden en de schattingen uit de Polis en de EBB

De schattingen vanuit de Polis en de EBB lopen tussen 2017 en 2018 langzamerhand verder uit elkaar. De modelschatting kan uiteengelegd worden in onder- en overschatting in de EBB en de Polis naar leeftijdscategorie en kwartaal. Dit is uitgerekend voor de acht kwartalen van 2017 en 2018. Daaruit kun je aflezen of de ontwikkeling kan worden verklaard door ontwikkelingen in de onnauwkeurigheden in de meting van de Polis en de EBB.

De ontwikkelingen in onnauwkeurigheden in de Polis en de EBB verklaren voor een belangrijk deel het verder uiteenlopen van de schattingen van het aandeel flexibele en vaste arbeidscontracten in de Polis en de EBB. Voor de totale populatie werknemers neemt het geschatte aandeel flexibele contracten zowel volgens het model als volgens de Polis toe, terwijl de EBB een lichte daling laat zien. De schattingen van het aandeel flexibele werknemers volgens de EBB gaat door de tijd dus meer afwijken van het geschatte aandeel volgens het model. In het bijzonder is de mate van onnauwkeurigheid in de EBB bij de 25- tot 55-jarigen daaraan debet: de onnauwkeurigheid dat er een vast contract bij de EBB wordt gemeten waar het model schat dat het gaat om een flexibel contract. De afname van de onnauwkeurigheden in de meting van de Polis kan worden verklaard door de betere werking van de loonaangifteketen over de tijd. Hiervoor zijn diverse aanwijzingen te vinden. Dit is in overeenstemming met de resultaten van de reeks 2007, 2009, 2016, 2017 en 2018.

Toch is dit maar een kant van het verhaal. Voor jongeren is de kans dat ze volgens één van beide bronnen een flexibel contract hebben terwijl het volgens het model om een vast contract gaat veel hoger. In de Polis schommelt deze kans rond de 43 procent. In de EBB is

deze kans lager dan in de Polis en bovendien daalt deze kans van bijna 28 procent in 2017 eerste kwartaal naar 19 procent in 2018 laatste kwartaal. Het is echter wederom de vraag hoeveel waarde je hieraan moet hechten, gezien het feit dat de resultaten van jongeren zo weinig stabiel zijn. Dat kan een gevolg zijn van het gering aantal waarnemingen. Dit vraagt om nader onderzoek.

7.4 Mogelijke verklaring van de onnauwkeurigheden in de Polis en de EBB

In hoofdstuk 3 zijn mogelijke verklaringen gegeven van de onnauwkeurigheden in de metingen van het soort arbeidscontract in de EBB en de Polis in de vorm van verwachtingen. In deze paragraaf wordt nagegaan of deze verwachtingen door de resultaten worden ondersteund. Daarvoor is naar de uitkomsten van twee analyses gekeken: de MLR-analyses (hoofdstuk 4) en de geschatte invloed van de covariaten op de onnauwkeurigheden zoals vastgesteld in de HMM-analyse (hoofdstuk 6).

De *vraagstelling in de EBB* is tot 2013 enkele keren *gewijzigd*. Daarnaast zijn er veranderingen geweest in de waarneming van de Polis door *verandering van wetgeving of administratieve praktijk*. Er zijn echter geen duidelijke cesuren in de mate van inconsistenties en de omvang van onnauwkeurigheden op de desbetreffende momenten aangetroffen. Wel hebben de verschillende kwartalen in 2017 en 2018 invloed op de onnauwkeurigheden in de meting, maar dat valt niet samen met de wijzigingen in de EBB of de Polis.

De *vraagstelling in de EBB is niet duidelijk genoeg*. Daarom zouden opleidingsniveau, migratieachtergrond, leeftijd, aantal contracturen en baanduren invloed hebben. Leeftijd, aantal contracturen en baanduren hebben sterke effecten, de effecten van de overige variabelen zijn veel kleiner. De effecten van leeftijd, aantal contracturen en baanduren zijn echter slechts deels in de voorspelde richting. Leeftijd heeft grotendeels effecten in de voorspelde richting. De baanduren bij jongeren hebben echter onverwachte effecten: des te langer de baanduur, des te groter de kans op inconsistenties en onnauwkeurigheid in de meting. Een mogelijke verklaring bij EBB vast, en volgens het model flex: bij een langere baanduur heeft een jonge EBB-respondent vaker ten onrechte het idee dat de baan vast is of zijn er informele afspraken tussen werkgever en werknemer hierover. De effecten van baanduren zijn bij 25-jarigen en ouder in de voorspelde richting: naarmate de baanduur toeneemt, neemt de kans op inconsistenties en onnauwkeurigheden in de meting in de Polis en de EBB af. Voor de jongste leeftijdscategorie geldt dat echter niet. De inconsistenties zijn groter bij de 19- tot 25-jarigen dan de nog jongere categorie. Wellicht is het voor de jongste categorie wel duidelijk dat ze een flexibel contract hebben, maar als de (bij)banen serieuzer worden dat minder duidelijk wordt zoals bij de op een na jongste categorie. 'Proxywaarneming' blijkt in tegenstelling tot de verwachting geen effect te hebben op de onnauwkeurigheden in de meting.

Er werd verwacht dat EBB-respondenten die na de eerste golf besluiten te *stoppen*, al minder gemotiveerd waren om nauwkeurige antwoorden te geven. Dit wordt niet verworpen door de uitkomsten van de MLR-analyses, maar de effecten zijn wel erg klein. In de HMM-analyses speelt deze variabele geen rol meer.

Bij de EBB is de *waarneemmodus* in de loop der tijd *gewijzigd*. Het is bekend dat de waarneemmodus een effect kan hebben op de meting van variabelen. Dit wordt door de uitkomsten van de analyses niet ondersteund: de waarneemmodus heeft vrijwel geen belangrijk effect op de inconsistenties tussen Polis en EBB. Daarom is het effect op de onnauwkeurigheden van de meetfouten niet in de HMM-analyses opgenomen.

Voor werkgevers is mogelijk de *betekenis van de arbeidsrelatie* in de Polis *onduidelijk* en de gevolgen van verkeerd invullen tot 2020 gering. Dit geldt natuurlijk in de eerste plaats voor die branches waarin veel tijdelijke contracten voorkomen. Daarom werd verwacht dat in branches waarin veel tijdelijke contracten voorkomen, deze fouten vaker worden gemaakt. Dit wordt deels ondersteund door de uitkomsten van de analyses. De bedrijven met economische activiteiten waar veel tijdelijke arbeidscontracten voorkomen laten meer inconsistenties zien in de meting van de Polis en de EBB. De invloed op de onnauwkeurigheden in de meting van de EBB en de Polis is echter klein.

De vraag of de kwaliteit waarmee de soort arbeidscontract in de Polis wordt ingevuld beïnvloed wordt door het ontbreken van een belang voor de werkgever dit goed te doen kan niet definitief beantwoord worden. Wel worden eventuele onnauwkeurige registraties in de loop der tijd 'doorgegeven' en niet altijd direct verbeterd. Meer indirecte bewijsvoering wordt gevonden in data van latere datum. Vanaf 1 januari 2020 zie je in de Polis een trendbreuk in de verhouding tussen vaste en flexibele contracten als gevolg van de invoering van de WAB. Het niveau van de vaste contracten is sinds 1 januari 2020 iets omhoog gegaan. Dit duidt er op dat de kwaliteit van het invullen met ingang van die datum is toegenomen.

Bedrijven leveren de informatie aan via standaard *softwarepakketten*.

Onnauwkeurigheden in de meting in de Polis hangen in belangrijke mate samen met de loonadministratiesoftware en de grootte van het bedrijf. Dit is goed verklaarbaar vanuit de verwachting dat vooral bij kleinere bedrijven relatief minder aandacht wordt besteed aan een gedegen administratie van de loongegevens. Ook spelen mogelijk het gebruik van bepaalde default settings of verschillen in de afhandeling van de loongegevens door de verschillende softwarepakketten een rol. Dat bij langere baanduren de onnauwkeurigheden in de Polis afnemen kan ook worden verklaard: waarschijnlijk worden loonadministraties van tijd tot tijd opgeschoond waardoor meetfouten geleidelijk en na een langere periode uit de administraties verdwijnen. Dat dit niet in alle gevallen gebeurt, of niet snel gebeurt, blijkt uit de hoge correlaties tussen de onnauwkeurigheden in de Polis-metingen over de tijd. Overigens hebben baanduren en contracturen nog zelfstandige invloed op de onnauwkeurigheden in de metingen van de Polis en EBB nadat rekening gehouden wordt met verschillen tussen de gebruikte softwarepakketten.

Ook *bedrijfsgrootte* heeft een zelfstandige invloed. In kleine en middelgrote ondernemingen komen vaker onnauwkeurigheden voor dan in de grote bedrijven. Dat was ook de verwachting die vooraf is geformuleerd. Bij kleine bedrijven is het een taak die of wordt uitbesteed aan een gespecialiseerd kantoor of door iemand van het bedrijf 'erbij wordt gedaan'. In het laatste geval zou dat een negatieve invloed op de kwaliteit hebben. Dit kan echter alleen een verklaring leveren voor de onnauwkeurigheden in de Polis, terwijl bedrijfsgrootte ook van invloed is op de onnauwkeurigheden in de EBB. De kans dat het model schat dat het gaat om een flexibel en de EBB om een vast contract is groter bij kleine dan bij grote bedrijven. Wellicht kan dit verklaard worden doordat er in de kleine

bedrijven vaker mondelinge afspraken worden gemaakt dat werknemers een vast contract hebben waardoor er een verschil ontstaat met de feitelijke registratie.

De *conjunctuur* en *werkgelegenheidsstructuur* kunnen ook invloed hebben op de perceptie van vaste en flexibele arbeidscontracten. De verwachting was dat mensen bij een positieve verwachting van het economisch klimaat voor het komende jaar eerder van mening zijn dat ze een vaste baan hebben, ook als dit formeel niet zo is en er ook geen informele afspraken over gemaakt zijn. Een positieve verwachting voor het economisch klimaat zou dan een significante toename te zien geven van de kans dat in de EBB een vast arbeidscontract wordt waargenomen terwijl het model een flexibel contract schat. Uit onze resultaten blijkt dat de kans hierop in de loop van 2017 en 2018 toeneemt, los van de invloed van andere covariaten uit dit onderzoek. Aangezien er in deze twee jaren een flinke economische groei was en de arbeidsmarkt steeds krappere werd, wordt deze verwachting door onze resultaten ondersteund.

7.5 Aanbevelingen

De EBB en de Polis verschillen sterk in de schatting van het aantal en aandeel vaste en flexibele werknemers. Dat wordt in ons onderzoek nogmaals bevestigd. De oorzaken van de onnauwkeurigheden zijn niet goed boven tafel te krijgen met alleen kwantitatief onderzoek. Er is aanvullend kwalitatief onderzoek nodig zowel onder personen als onder bedrijven. Doel van het aanvullende onderzoek is om inzicht te verkrijgen in de mechanismen die zorgen voor het ontstaan van onnauwkeurigheden in de metingen van het type arbeidscontract in de EBB en in de Polis.

Vanaf 1 januari 2020 is de Wet Arbeidsmarkt in Balans in werking getreden. Als gevolg van deze wet zijn er voor de werkgever financiële consequenties aan het soort contract van de werknemer verbonden. Het spreekt vanzelf dat dit invloed heeft op de invulling van dit onderdeel in de salarisadministratie. Daarnaast wordt in 2021 ook de vraagstelling op onder andere het gebied van contractsoort bij de EBB gewijzigd. Daarom zou vervolgonderzoek zich moeten richten op de nieuwe situatie.

Bij de EBB wordt verder geadviseerd om te verklaren hoe respondenten komen tot een antwoord op de vragen naar de soort arbeidscontract. Begrijpt de respondent de vraag niet goed? Verstaat de respondent wat anders onder een vast contract dan een contract voor onbepaalde tijd? Is het type contract niet goed geregistreerd? Voorgesteld wordt om een personensteekproef op basis van de Polis te trekken gestratificeerd naar variabelen waarvan bekend is dat die samenhangen met een grotere kans op onnauwkeurigheden (bijvoorbeeld jongeren, personen werkzaam in bepaalde sectoren, kleine banen, kleine bedrijven). Dit kan bijvoorbeeld in de vorm van enkele focusgroepen of door het houden van cognitieve interviews. Voor dergelijk onderzoek zou eerst informatie kunnen worden verzameld over de arbeidsrelatie op basis van de vraagstelling die nu wordt gehanteerd in de EBB, die vervolgens kan worden 'geconfronteerd' met de in de Polis geregistreerde informatie. Door gericht in te gaan op de vraagstelling en het verschil in meting, kunnen naar verwachting betere verklaringen gevonden worden over de oorzaken van de verschillen. Daarmee kan worden nagegaan of en voor welke respondenten de vraagstelling onduidelijk is, op welke punten dat het geval is en voor welke respondenten de registratie niet klopt.

Bij de Polis zou het onderzoek zich moeten richten op de mensen in de bedrijven die verantwoordelijk zijn voor het aanleveren van de informatie voor de Polis en de door hen gebruikte softwarepakketten. Ook voor deze groep kunnen soortgelijke vragen worden geformuleerd als hierboven aan de respondenten in focusgroepen of in cognitieve interviews gesteld. Een bijzondere positie neemt hier de software in die wordt gebruikt voor deze aangiftes. Gezien de sterke effecten van softwarecluster op fouten in de Polis moet dit in het vervolgonderzoek worden meegenomen. Ook dit onderzoek zou een kwalitatief karakter moeten hebben.

De rol van de categorie 'Overig' vraagt ook om vervolgonderzoek. De Polis wordt door het CBS niet als bron gebruikt om schattingen te maken van het aantal en aandeel personen zonder betaald werk en bevat ook maar een deel van de zelfstandigen. De groep 'Overig' in de Polis betreft dus EBB-respondenten die ofwel niet in de Polis voorkomen ofwel Directeur- Grootaandeelhouder zijn. Hoe moeten de geschatte onnauwkeurigheden met betrekking tot de categorie 'Overig' geïnterpreteerd worden? Resultaten voor deze categorie hebben invloed op de schattingen van aandelen en aantallen vaste en flexibele werknemers. Dit speelt waarschijnlijk het meest voor jongeren omdat die veel vaker een periode zonder werk doormaken (met name als het een bijbaan naast de studie betreft).

Een vervolgonderzoek zou ook moeten ingaan op verschil in onderliggende concepten wat betreft de onderscheiden arbeidsrelaties en de meting van de niet-werknemers. Flexibele arbeidscontracten bestaan uit tijdelijke contracten, uitzendcontracten en oproepcontracten. Het ligt voor de hand dat bij de verschillende soorten flexibele arbeidscontracten ook andere oorzaken wat betreft de onnauwkeurigheden een rol spelen. Daarbij speelt ook nog mee dat oproepkrachten en uitzendkrachten een contract voor onbepaalde tijd kunnen hebben. Er kunnen dus in beide bronnen op meerdere dimensies een fout worden gemaakt. Vervolgonderzoek zou hier ook aandacht aan moeten besteden.

Referenties

Bakker, B.F.M. & P. Daas, 2012. Some Methodological Issues of Register Based Research. *Statistica Neerlandica*, jrg. 66, pp. 2-7

Bakker, B.F.M., J. van Rooijen & L. van Toor, 2014. The system of social statistical datasets of Statistics Netherlands: an integral approach to the production of register-based social statistics. *Statistical Journal of the International Association of Official Statistic*, jrg. 30, pp. 411-424 DOI 10.3233/SJI-140803

Bakk, Z., D.L. Oberski & J.K. Vermunt, 2016. Relating Latent Class Membership to Continuous Distal Outcomes: Improving the LTB Approach and a Modified Three-Step Implementation. *Structural Equation Modeling: A Multidisciplinary Journal*, jrg. 23, pp. 278-289

- Baum, L.E., T. Petrie, G. Soules, & N. Weiss, 1970. A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *The Annals of Mathematical Statistics*, jrg. 41, pp. 164-171
- Berg, van den, H., J. Erkens, L. Hoekema, L. Kösters, H. Muller, J. Nieuweboer, N. Pouwels-Urlings, M. Walthouwer, & J. van der Zwan, 2016. Werknemers en zelfstandigen: overeenkomsten en verschillen tussen CBS-cijfers. *Sociaal-economische Trends*, jrg. 2016, pp. 1-23
- Boeschoten, L., D.L. Oberski & T. de Waal, 2017. Estimating Classification Errors under Edit Restrictions in Composite Survey-Register Data using Multiple Imputation Latent Class Modelling (MILC). *Journal of Official Statistics*, jrg. 33, pp. 921–962
- Bruggink, J., J. Michiels & L. van Toor, 2008. *Overschatting baandynamiek in het SSB*. (Heerlen / Voorburg: CBS-nota)
- Dempster, A.P., N.M. Laird & D.B. Rubin, 1977. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B*, jrg. 39, pp. 1-38
- Gringhuis, G., 2011. *Kwaliteit van de variabelen arbeidscontract en contractsoort in de Polisadministratie*. (Den Haag: CBS-nota)
- Gringhuis, G., J. Michiels, & B. Buelens, 2012. *Flexibele arbeidscontracten op basis van Polis en EBB*. (Heerlen: CBS-nota)
- Klausch, Th., B. Schouten & J. Hox, 2015. Evaluating bias of sequential mixed-mode designs against benchmark surveys. *Sociological Methods & Research*, jrg. 46, pp. 456-489
- Little, R.J., & D.B Rubin, 2002. *Statistical analysis with missing data*. (New Jersey: Wiley)
- Mars, G., 2011. *Cijfers over flexibele arbeidscontracten – confrontatie van bronnen en definities*. (Heerlen: CBS-nota)
- Meertens, V., S. Krieg & M. Cremers, 2018. *Resultaten evaluatie aangepaste vraagstellingen in de EBB & POLIS* (Heerlen: CBS-nota BPM-2018-198-09-VMTS-MCAM)
- Pankowska, P., B. Bakker, D.L. Oberski, & D. Pavlopoulos, D., 2018. Reconciliation of inconsistent data sources by correction for measurement error: The feasibility of parameter re-use. *Statistical Journal of the IAOS*, jrg. 34, pp. 317-329
- Pavlopoulos, D., & J.K. Vermunt, 2015. Measuring temporary employment. Do survey or register data tell the truth? *Survey Methodology*, jrg. 41, pp. 197–214
- Schwarz, G.E., 1978. Estimating the dimension of a model, *Annals of Statistics*, jrg. 6, pp. 461–464
- Vermunt, J.K., B. Tran, & J. Magidson, 2008. Latent class models in longitudinal research. S. Menard (red.), *Handbook of Longitudinal Research: Design, Measurement, and Analysis* (Burlington, MA: Elsevier) pp. 373-385

Bijlage 1. De variabele soort arbeidscontract

Werknemers of zelfstandigen?

Dit onderzoek richt zich primair op verschillen in de schatting van het aantal flexibele en vaste arbeidscontracten van werknemers. Er is een groep werkenden die in de ene statistiek als werknemer wordt getypeerd en in de andere statistiek als zelfstandige, namelijk de DGA's (van den Berg et al., 2016). Banen van DGA's worden in de statistiek Werkgelegenheid en lonen (en in de Arbeidsrekeningen) – op basis van de Polis – tot de werknemersbanen gerekend. In de statistieken over de beroepsbevolking op basis van de EBB en in de Inkomensstatistiek worden DGA's tot zelfstandigen gerekend. Dit verschil betreft geen onnauwkeurigheid maar een verschil in definitie⁹. In dit onderzoek wordt de doelpopulatie afgebakend als de groep werkenden die volgens beide statistieken tot de categorie 'Werknemers' wordt gerekend. DGA's worden in dit onderzoek dus niet tot de werknemers gerekend.

Vast of flex?

Om te bepalen of er sprake is van een vast dan wel een flexibel arbeidscontract zijn er twee dimensies van belang: de looptijd van het contract en de contracturen¹⁰. Een vast arbeidscontract is gedefinieerd als een arbeidscontract waarbij sprake is van een contract voor onbepaalde tijd en een vast aantal uren per week. Een flexibel arbeidscontract is een contract voor bepaalde tijd of van wisselende uren.

Operationalisatie in de EBB

In statistieken op basis van de EBB worden de volgende groepen werknemers tot de werknemers met een flexibel arbeidscontract gerekend:

- Werknemers met een vast arbeidscontract zonder vaste uren
- Werknemers met een tijdelijk contract die uitzicht hebben op een vast arbeidscontract
- Werknemers met een tijdelijk contract langer dan een jaar
- Werknemers met een tijdelijk contract korter dan een jaar
- Uitzendkrachten
- Oproepkrachten
- Werknemers met een tijdelijk contract zonder vaste uren

Alleen de werknemers die zowel een arbeidscontract voor onbepaalde tijd als vaste uren hebben worden tot de werknemers met een vast arbeidscontract gerekend. Verder geldt dat oproepkrachten en uitzendkrachten altijd tot de flexibele werknemers worden gerekend, ook als zij een contract voor onbepaalde tijd hebben. Voor oproepkrachten is dat omdat er bij dit type contract sprake is van wisselende uren. Bij uitzendkrachten is de overweging dat ook als er sprake is van een contract voor onbepaalde tijd, het wegvallen van de inlener een grond voor ontslag kan zijn.

⁹ Zie ook <https://www.cbs.nl/nl-nl/achtergrond/2016/46/cbs-statistieken-over-zelfstandigen>

¹⁰ Zie ook <https://www.cbs.nl/nl-nl/dossier/dossier-flexwerk/hoofdcategorieen/wat-zijn-flexwerkers->

Operationalisering in de Polis

In de statistieken op basis van de Polis (Statistiek werkgelegenheid en lonen) worden de volgende banen tot de flexibele banen gerekend¹¹:

- Reguliere banen met een contract voor bepaalde tijd
- WSW-banen met een contract voor bepaalde tijd
- Banen van stagiairs (zowel bepaalde als onbepaalde tijd)
- Uitzendbanen (bepaalde en onbepaalde tijd)
- Oproepbanen (bepaalde en onbepaalde tijd)

Reguliere banen met een contract voor onbepaalde tijd, banen van WSW'ers voor onbepaalde tijd en banen van DGA's worden tot de vaste banen gerekend.

Operationalisering in dit onderzoek

De categorieën (flexibele) werknemers die in de EBB en de Polis kunnen worden onderscheiden sluiten niet helemaal aan. De groep werknemers met een vast arbeidscontract zonder vaste uren is niet te onderscheiden in de Polis. Omgekeerd kunnen in de EBB de DGA's niet worden onderscheiden omdat DGA's meestal aangeven zelfstandige te zijn. Voor de vergelijking tussen Polis en de EBB worden vast, flex en overig daarom als volgt geclassificeerd:

EBB

Vast:

- Werknemers met een vast arbeidscontract met vaste uren
- Werknemers met een vast arbeidscontract zonder vaste uren

Flexibel:

- Werknemers met een tijdelijk contract die uitzicht hebben op een vast arbeidscontract
- Werknemers met een tijdelijk contract langer dan een jaar
- Werknemers met een tijdelijk contract korter dan een jaar
- Uitzendkrachten
- Oproepkrachten
- Werknemers met een tijdelijk contract zonder vaste uren

Overig:

- Personen zonder betaald werk
- Zelfstandigen

Polis

Vast:

- Werknemers met een reguliere baan voor onbepaalde tijd
- Werknemers met een baan in het kader van de WSW voor onbepaalde tijd

Flexibel

- Werknemers met een reguliere baan voor bepaalde tijd
- Werknemers met een WSW-baan voor bepaalde tijd
- stagiairs
- Werknemers met een uitzendbaan (bepaalde en onbepaald)
- Werknemers met een oproepbaan (bepaald en onbepaald)

¹¹ Zie ook <https://www.cbs.nl/nl-nl/nieuws/2018/17/1-op-3-werkt-10-jaar-of-langer-bij-dezelfde-werkgever>

Overig

- Werknemers met een baan als DGA
- Personen zonder baan als werknemer

Bijlage 2. De procedure voor het schatten van de invloed van covariaten op onnauwkeurigheden in de meting in Polis en EBB

Een van de doelen van het onderzoek naar onnauwkeurigheid in de meting in de Polis en in de EBB is het vinden van groepen in de populatie waarbij onnauwkeurigheden in een of in beide bronnen relatief vaak voorkomen. Het is daarom wenselijk om een analyse te doen naar de invloed van achtergrondkenmerken (covariaten) op de geschatte onnauwkeurigheden. Echter, het direct opnemen in een Hidden Markov Model (HMM) van effecten van covariaten op (onnauwkeurigheid in de meting in) de waargenomen arbeidscontracten heeft een theoretisch en een praktisch nadeel:

- Het meetmodel (het deel van het latente klasse-model dat de relatie tussen de latente arbeidscontracten en de waargenomen arbeidscontracten beschrijft) bepaalt expliciet hoe de latente variabele gedefinieerd is (Bakk et al., 2016). Het direct opnemen van covariaten in het meetmodel kan daarom als gevolg hebben dat de betekenis van de latente variabele in het model verandert: in plaats van 'datgene dat de metingen in Polis en EBB gemeenschappelijk hebben' wordt de definitie dan 'datgene dat de metingen in Polis en EBB gemeenschappelijk hebben en dat niet verklaard wordt door de covariaten'.
- In ons geval wordt een HMM met effecten van (veel) covariaten in het meetmodel zeer complex, zowel wat het aantal te schatten parameters betreft als de grootte van de kruistabellen die Latent Gold intern berekent om het model te schatten. De rekentijd die Latent Gold nodig heeft om dergelijke modellen te schatten wordt daardoor onacceptabel lang.

In dit onderzoek is een aanpak ontwikkeld geïnspireerd op de MILC-methode van Boeschoten et al. (2017):

- Schat met Latent Gold een HMM waarin covariaten invloed mogen hebben op de latente tijdreeks (initiële kansen en overgangskansen) maar niet op het meetmodel.
- Op basis van de output van Latent Gold uit de vorige stap worden latente tijdreeksen van vijftien maanden $X^{(1)}, \dots, X^{(15)}$ gesimuleerd per persoon die voldoen aan de met behulp van het model geschatte posteriorkansen $P(X^{(1)} = k)$ en $P(X^{(t)} = l | X^{(t-1)} = k)$, met $k = 1, 2, 3$ en $l = 1, 2, 3$ (1 = vast, 2 = flex, 3 = overig). Genereer, zeg, $S = 100$ onafhankelijke datasets met zulke tijdreeksen. Noteer verder het waargenomen arbeidscontract in maand t uit EBB en Polis als, respectievelijk, $Y_{Ebb}^{(t)}$ en $Y_{Polis}^{(t)}$.
- Doe voor elke dataset $s = 1, \dots, S$ uit de vorige stap het volgende.
 - a. Gebruik alle waarnemingen met $X_s^{(t)} = 1$ (vast) om multinomiale logistische regressiemodellen (MLR) te schatten voor de kansen op onnauwkeurigheid in respectievelijk de EBB en de Polis $P(Y_{Ebb}^{(t)} = y | X^{(t)} = 1)$ en $P(Y_{Polis}^{(t)} = y | X^{(t)} = 1)$

- als functie van covariaten. Neem in beide gevallen $y = 1$ (d.w.z. geen onnauwkeurigheid) als referentiecategorie. Noteer de gevonden regressiecoëfficiënten als $\hat{\beta}_{1s} = (\hat{\beta}_{1s1}, \dots, \hat{\beta}_{1sR})'$ en de bijbehorende geschatte varianties van deze coëfficiënten als $\hat{\sigma}_{1s}^2 = (\hat{\sigma}_{1s1}^2, \dots, \hat{\sigma}_{1sR}^2)'$.
- b. Doe hetzelfde voor alle waarnemingen met $X_s^{(t)} = 2$ (flex) en de kansen $P(Y_{Ebb}^{(t)} = y | X^{(t)} = 2)$ en $P(Y_{Polis}^{(t)} = y | X^{(t)} = 2)$, nu met $y = 2$ als referentiecategorie. Noteer de gevonden regressiecoëfficiënten als $\hat{\beta}_{2s}$ en de geschatte varianties als $\hat{\sigma}_{2s}^2$.
- c. Doe hetzelfde voor alle waarnemingen met $X_s^{(t)} = 3$ (overig) en de kansen $P(Y_{Ebb}^{(t)} = y | X^{(t)} = 3)$ en $P(Y_{Polis}^{(t)} = y | X^{(t)} = 3)$, nu met $y = 3$ als referentiecategorie. Noteer de gevonden regressiecoëfficiënten als $\hat{\beta}_{3s}$ en de geschatte varianties als $\hat{\sigma}_{3s}^2$.
- Bereken voor elk effect uit de MLR-analyses de gemiddelde coëfficiënt uit de $S = 100$ simulatieronden als einduitkomst: $\hat{\beta}_1 = \frac{1}{S} \sum_{s=1}^S \hat{\beta}_{1s}$, $\hat{\beta}_2 = \frac{1}{S} \sum_{s=1}^S \hat{\beta}_{2s}$ en $\hat{\beta}_3 = \frac{1}{S} \sum_{s=1}^S \hat{\beta}_{3s}$. Een indicatie voor de variantie van een geschat effect uit $\hat{\beta}_k$ ($k = 1, 2, 3$), zeg $\hat{\beta}_{kr}$, kan worden afgeleid met behulp van Rubins regel voor multiële imputatie (Little & Rubin, 2002):

$$\begin{aligned} \widehat{\text{var}}(\hat{\beta}_{kr}) &= \hat{V}_{binnen} + \left(1 + \frac{1}{S}\right) \hat{V}_{tussen} \\ &= \frac{1}{S} \sum_{s=1}^S \hat{\sigma}_{ksr}^2 + \left(1 + \frac{1}{S}\right) \frac{1}{S-1} \sum_{s=1}^S (\hat{\beta}_{ksr} - \hat{\beta}_{kr})^2 \end{aligned}$$

Afgezien van het herhaaldelijk toepassen op gesimuleerde data zijn deze MLR-analyses vergelijkbaar met de MLR-analyses op verschillen tussen EBB en Polis die zijn beschreven in hoofdstuk 4. Nu zijn de doelvariabelen echter: verschillen tussen EBB en modelschatting en verschillen tussen Polis en modelschatting. Uit de gevonden effecten kunnen conclusies worden getrokken over groepen in de populatie waarbij relatief veel of weinig onnauwkeurigheden voorkomen in EBB en/of Polis, op dezelfde manier als eerder conclusies getrokken werden over groepen waarbij relatief veel of weinig verschillen voorkomen tussen EBB en Polis.

De varianties berekend in stap 4 zijn alleen indicatief en waarschijnlijk eerder iets te laag dan iets te hoog, om een aantal redenen:

- Hoewel Rubins regel wordt toegepast is er geen sprake van echte multiële imputatie: de onzekerheid in de geschatte parameters van het HMM is verwaarloosd. Om hiermee rekening te houden zou het HMM meerdere keren geschat kunnen worden op gebootstrapte versies van de oorspronkelijke dataset. Dit was in dit onderzoek echter geen optie gezien de beperkte tijd die beschikbaar was om het onderzoek af te ronden.
 - In de MLR-analyses is geen rekening gehouden met het feit dat de waarnemingen niet onafhankelijk zijn, omdat er steeds meerdere waarnemingen gaan over dezelfde persoon (15 bij de Polis, maximaal 5 bij de EBB). De geschatte varianties in $\hat{\sigma}_{1s}^2$, $\hat{\sigma}_{2s}^2$ en $\hat{\sigma}_{3s}^2$ zullen daarom te laag zijn. In principe zou dit kunnen worden opgelost door het MLR-model uit te breiden met een multilevel-component. Ook dit was gezien de beperkte tijd in dit onderzoek geen optie; bovendien is geen standaard software beschikbaar waarmee een multilevel-multinomiale logistische regressie geschat kan worden.
 - Er is geen rekening gehouden met het steekproefontwerp van de EBB.
- Het laatste punt is ook van toepassing op de eerdere MLR-analyses naar verschillen tussen Ebb en Polis uit hoofdstuk 4. Net als in die eerdere analyses hoeft het ontbreken van

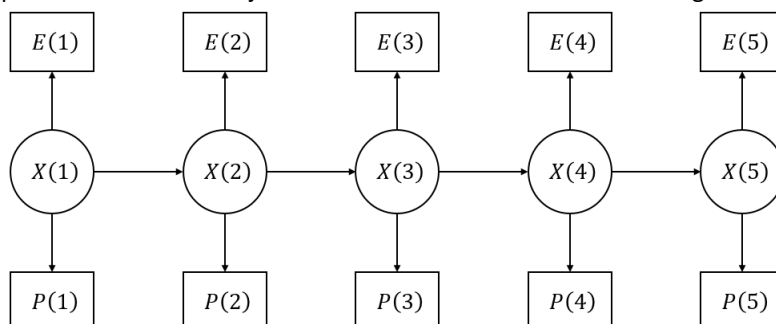
exacte varianties geen groot bezwaar te zijn; er wordt immers niet statistisch getoetst welke effecten significant zijn. Wel wordt een beeld verkregen welke effecten de grootste zijn.

Boeschoten et al. (2017) hebben laten zien dat de MILC-aanpak goede resultaten geeft voor een aantal typen statistieken, mits de voorspelkracht van het latente klasse-model groot genoeg is. Als maat voor de voorspelkracht hebben ze gekeken naar de zogenaamde entropie- R^2 , die voor een dataset met n waarnemingen en een latente variabele X met K klassen is gegeven door:

$$R_{ent}^2 = 1 - \frac{-\sum_{i=1}^n \sum_{k=1}^K P(X_i = k) \log P(X_i = k)}{n \log K},$$

waarbij $P(X_i = k)$ een door het model voorspelde posteriorkans is. Per definitie geldt: $0 \leq R_{ent}^2 \leq 1$, waarbij een waarde dicht bij 1 wijst op een hogere voorspelkracht. Boeschoten et al. (2017) vonden in een simulatiestudie dat de MILC-aanpak goed werkt als $R_{ent}^2 \geq 0,9$ en voor sommige toepassingen zelfs al bij $R_{ent}^2 \geq 0,6$. De auteurs hebben daarbij echter niet specifiek gekeken naar een toepassing op het modelleren van onnauwkeurigheidskansen. Om te testen of de bovenstaande aanpak ook werkt in die context is deze toegepast op een klein voorbeeld met synthetische data, waar de werkelijke effecten bekend zijn.

Het model in dit voorbeeld is een vereenvoudiging van het HMM voor arbeidscontract in EBB en Polis. Er zijn slechts 5 perioden en alle onnauwkeurigheden in de meting zijn onafhankelijk in de tijd. Voor het gemak wordt aangenomen dat 'EBB' en 'Polis' in elke periode beschikbaar zijn. Schematisch ziet het model er als volgt uit:



Er zijn drie latente klassen die voor het gemak worden aangeduid als: 1 = vast, 2 = flex, 3 = overig. Voor een dataset van $n = 1000$ personen zijn eerst latente variabelen gegenereerd uit een verdeling met kansen $P(X^{(1)} = 1) = 0,4$ en $P(X^{(1)} = 2) = P(X^{(1)} = 3) = 0,3$. De overgangskansen zijn gegeven door $P(X^{(t)} = k | X^{(t-1)} = k) = 0,9$ en $P(X^{(t)} = l | X^{(t-1)} = k) = 0,05$ voor $l \neq k$. Verder is een (tijdsafhankelijke) covariaat met twee categorieën gegenereerd, waarbij 30 procent van de personen valt in de eerste categorie (zeg 'A') en 70 procent in de tweede categorie (zeg 'B').

Vervolgens zijn hierbij waargenomen tijdreeksen gegenereerd uit twee bronnen die voor het gemak worden aangeduid als 'EBB' en 'Polis'. De onnauwkeurigheidskansen zijn in dit voorbeeld bij beide bronnen hetzelfde en verschillen per categorie van de covariaat. Voor categorie 'A' van de covariaat zijn de onnauwkeurigheidskansen in beide bronnen gegeven door de matrix (rij = werkelijk, kolom = waargenomen):

$$\begin{pmatrix} 0,90 & 0,05 & 0,05 \\ 0,10 & 0,80 & 0,10 \\ 0,05 & 0,10 & 0,85 \end{pmatrix}.$$

Voor categorie 'B' van de covariaat is de matrix:

$$\begin{pmatrix} 0,80 & 0,10 & 0,10 \\ 0,05 & 0,90 & 0,05 \\ 0,05 & 0,10 & 0,85 \end{pmatrix}.$$

Dus: in categorie 'B' komt meer onnauwkeurigheid in de meting voor bij mensen die volgens het model een vaste baan hebben [$X^{(t)} = 1$] en in categorie 'A' komt juist meer onnauwkeurigheid in de meting voor bij mensen die volgens het model een flexbaan hebben [$X^{(t)} = 2$]. De kans op onnauwkeurigheid in de meting voor mensen in de categorie 'Overig' [$X^{(t)} = 3$] hangt niet af van de covariaat.

Het HMM is geschat met Latent Gold. De door het model voorspelde marginale verdeling en overgangskansen is te vergelijken met de werkelijke verdeling en overgangskansen die in dit geval direct kunnen worden berekend uit de gegenereerde dataset. Voor de marginale verdeling is het resultaat (gemiddeld over de hele periode):

	werkelijk	'EBB'	'Polis'	model
categorie	%			
1 (vast)	38,1	34,6	34,6	37,9
2 (flex)	31,5	33,8	33,9	31,7
3 (overig)	30,4	31,6	31,5	30,4

Voor de overgangskansen is het resultaat (gemiddeld over de hele periode):

	(t-1) \ t	1 (vast)	2 (flex)	3 (overig)
		%		
werkelijk	1 (vast)	90,5	4,8	4,8
	2 (flex)	5,0	89,2	5,7
	3 (overig)	5,8	4,0	90,2
'EBB'	1 (vast)	71,0	14,2	14,8
	2 (flex)	14,6	66,7	18,7
	3 (overig)	15,5	17,6	66,8
'Polis'	1 (vast)	68,1	17,2	14,7
	2 (flex)	15,8	67,8	16,4
	3 (overig)	17,3	16,2	66,5
model	1 (vast)	90,5	4,4	5,1
	2 (flex)	5,0	89,4	5,6
	3 (overig)	5,7	4,0	90,3

Vervolgens zijn 100 voorspelde tijdreeksen gesimuleerd uit het geschatte model en is de bovenstaande procedure toegepast om MLR-modellen te schatten voor onnauwkeurigheidskansen met de covariaat als predictor. Ook hier kunnen de gevonden effecten vergeleken worden met het resultaat van het schatten van dezelfde MLR-modellen direct op de oorspronkelijke dataset, gebruikmakend van de bekende latente variabele. De resultaten van beide analyses voor onnauwkeurigheidskansen voor mensen met volgens het model een vast arbeidscontract [$X^{(t)} = 1$] staan hieronder.

		effect	vast (1)→ flex (2)		vast (1)→ overig (3)	
			coef.	std.fout	coef.	std.fout
werkelijk	'EBB'	intercept	-3,07	0,21	-2,84	0,19
		covariaat='B'	0,93	0,23	0,65	0,21
	'Polis'	intercept	-2,71	0,18	-2,68	0,18
		covariaat='B'	0,61	0,20	0,66	0,20
simulatieaanpak	'EBB'	intercept	-3,47	0,45	-3,34	0,42
		covariaat='B'	0,63	0,24	0,56	0,23
	'Polis'	intercept	-3,08	0,40	-3,21	0,39
		covariaat='B'	0,47	0,22	0,60	0,21

De aanpak op gesimuleerde tijdreeksen leidt hier tot de juiste conclusie: categorie 'B' van de covariaat heeft in alle gevallen een positieve coëfficiënt, dus er komt in beide bronnen meer onnauwkeurigheid in de meting van vast naar flex en van vast naar overig voor bij categorie 'B' van de covariaat dan bij categorie 'A'. (Ter herinnering: de interpretatie van coëfficiënten uit een MLR-model als effecten op kansverhoudingen is toegelicht aan het eind van paragraaf 4.1.) Gezien de standaardfouten liggen de gevonden effecten redelijk in de buurt van de effecten die zouden worden geschat op de werkelijke data.

Voor de mensen met volgens het model een flexbaan [$X^{(t)} = 2$] vinden we:

		effect	flex (2)→ vast (1)		flex (2)→ overig (3)	
			coef.	std.fout	coef.	std.fout
werkelijk	'EBB'	intercept	-2,32	0,18	-1,98	0,15
		covariaat='B'	-0,84	0,24	-0,73	0,20
	'Polis'	intercept	-1,93	0,15	-2,14	0,17
		covariaat='B'	-1,03	0,21	-0,88	0,22
simulatieaanpak	'EBB'	intercept	-1,67	0,44	-1,57	0,43
		covariaat='B'	-0,69	0,24	-0,50	0,23
	'Polis'	intercept	-1,32	0,38	-1,43	0,40
		covariaat='B'	-0,73	0,21	-0,76	0,22

Ook hier zou op basis van de gesimuleerde aanpak de juiste conclusie worden getrokken: categorie 'B' van de covariaat heeft steeds een negatieve coëfficiënt, dus er komen in beide bronnen minder onnauwkeurigheden in de meting van flex naar vast en van flex naar overig voor bij categorie 'B' dan bij categorie 'A'.

Ten slotte zijn de resultaten voor de mensen in de categorie 'Overig' [$X^{(t)} = 3$]:

		effect	overig (3)→vast (1)		overig (3)→flex (2)	
			coef.	std.fout	coef.	std.fout
werkelijk	'EBB'	intercept	-3,13	0,24	-2,33	0,17
		covariaat='B'	-0,06	0,29	0,28	0,19
	'Polis'	intercept	-3,00	0,23	-2,15	0,15
		covariaat='B'	-0,09	0,27	-0,04	0,19
simulatieaanpak	'EBB'	intercept	-3,09	0,46	-2,66	0,42
		covariaat='B'	-0,07	0,26	0,33	0,22
	'Polis'	intercept	-2,74	0,42	-2,07	0,40
		covariaat='B'	-0,16	0,23	-0,06	0,22

Ook in dit geval liggen de gevonden effecten bij de simulatieaanpak dicht bij de gevonden effecten op de werkelijke data. Gezien de (indicatieve) standaardfouten zou geconcludeerd worden dat de covariaat bij deze categorie waarschijnlijk geen effect heeft op de onnauwkeurigheid in de meting in beide bronnen.

Samengevat lijkt de aanpak in dit voorbeeld goed te werken. De entropie- R^2 is in dit voorbeeld overigens gelijk aan $R_{ent}^2 = 0,91$. Voor de modellen geschat op de werkelijke data vinden we, gemiddeld over de hele periode 2017-2018, de volgende waarden voor de entropie- R^2 : $R_{ent}^2 = 0,88$ ('15 tot 25 jaar'), $R_{ent}^2 = 0,95$ ('25 tot 55 jaar') en $R_{ent}^2 = 0,98$ ('55 tot 75 jaar').

Bijlage 3. De variabele softwarecluster

De data

De gegevens over de gebruikte software omvatten zeer veel verschillende softwarepakketten per jaar ook omdat het vaak verschillende versies van dezelfde softwarepakketten betreft. Een gerelateerde variabele met veel minder categorieën is de variabele softwareleverancier waarvan er 103 zijn. Deze laatste variabele, in het vervolg aangeduid als 'Software' wordt in deze analyse gebruikt.

Bij de analyses met de gekoppelde gegevens van Polis en EBB over de jaren 2010 t/m 2018 bevatte het databestand 526036 banen. Aan 519119 hiervan (98,7 procent) kon de variabele 'Software' worden gekoppeld, voor de overige 6917 is de software categorie 'Onbekend'. Sommige software wordt weinig toegepast. Dit leidt tot slecht gevulde categorieën van de variabele 'Software', met onbetrouwbare parameterschattingen tot gevolg. Daarom zijn de slecht gevulde categorieën van software samengevoegd tot een categorie 'Overig'. Omdat de analyses apart uitgevoerd worden voor drie leeftijdscategorieën zijn alle categorieën met minder dan 20 banen in één van de leeftijdscategorieën samengevoegd. Hierdoor is het aantal categorieën van 'Software' tot 57 teruggebracht.

Voor een aantal van deze 57 categorieën van 'Software' komt één van de categorieën van de doelvariabele, namelijk 'EBB flex, Polis vast', helemaal niet voor. De parameters van het logistische regressiemodel kunnen dan niet geschat worden. Daarom zijn deze categorieën van 'Software' buiten de analyse gehouden. Het resulterende aantal categorieën is daarmee 47 geworden en het aantal records is met 1,6 procent afgenomen tot 517432.

Toevoegen van de verklarende variabele software

Net zoals bij de eerdere analyses is het multinomiale regressiemodel toegepast voor ieder van de drie onderscheiden leeftijdscategorieën en voor alle leeftijdscategorieën samen. In onderstaande tabel staan toetsingsgrootheden voor modellen voor alle leeftijdscategorieën samen. Voor iedere variabele is de likelihood-ratio toetsingsgrootheid voor het weglaten van die variabele weergegeven (chi-square in de SPSS-output). Deze grootheid is een maat voor het verschil in fit van het model zonder en het model met de betrokken variabele. Grote waarden geven aan dat dit verschil groot is en de variabele veel effect heeft. De toetsingsgrootheden in de kolom 'Zonder software' zijn de waarden van het eerder toegepaste MLR-model, dus zonder de verklarende variabele 'Software'. Als de variabele 'Software' wordt toegevoegd toe blijkt dat deze variabele het grootste effect heeft van alle variabelen. Daarbij moet worden bedacht dat deze variabele ook de meeste parameters (92) introduceert. Per parameter beschouwd zijn de effecten van 'Leeftijdsklasse', 'SBI-groep', 'Contracturen' en 'Baanduur' groter. In onderstaande tabel

zijn de bijdragen opgenomen van variabelen aan de verklaring van inconsistenties tussen EBB en Polis in de waarneming van arbeidscontract voor alle leeftijdsgroepen.

Variabele	Aantal parameters	Zonder Software	Met Software in 47 categorieën	Met Software in 4 clusters
Software	92		15237,9	
Software clusters	6			11026,8
Leeftijdsklasse	10	8850,9	7343,6	7615,0
SBI-groep	16	4858,9	2538,1	3726,6
Contracturen	6	2601,0	2474,5	2495,5
Baanduur	10	2695,9	2549,4	2632,4
Kwartaalnummer	68	928,5	898,1	768,0
Opleiding	6	236,9	116,6	121,7
Migratieachtergrond	12	102,4	115,0	120,7
Proxywaarneming	2	79,2	61,3	65,2
Laatste respons	2	50,1	39,6	41,9
Geslacht	2	12,3	32,2	21,7
Waarneemmodus	4	36,0	41,4	42,3

Het is niet waarschijnlijk dat alle 92 parameters van 'Software' sterk verschillende effecten hebben. Een samenvoeging van parameters, of categorieën van 'Software', kan dan worden overwogen. Met behulp van clusteranalyse is een indikking van de variabele 'Software' tot 4 clusters van softwarepakketten verkregen (de volgende paragraaf geeft meer details over deze analyse). De laatste kolom geeft de toetsingsgrootheden voor deze tot 4 categorieën ingedikte variabele. Het effect van 'Software' neemt slechts af van 15000 naar 11000 terwijl het aantal parameters afneemt van 92 naar 6. De effecten van de andere verklarende variabelen blijven in grote mate gelijk, met als uitzondering het kleinere effect van 'SBI-groep' bij de uitgebreide 'Software'-variabele met 47 categorieën dan bij de ingedikte variabele met 4 categorieën (of helemaal geen 'Software'-variabele). Een plausibele verklaring voor deze uitzondering is dat de uitgebreide variabele 'Software' voor een aantal categorieën grotendeels samenvalt met 'SBI-groep', het betreft dan software die vrijwel uitsluitend in bepaalde SBI's wordt toegepast. De uitgebreide 'Software'-variabele neemt dan deels het effect van 'SBI-groep' over, waardoor dit effect vermindert wordt. Het gebruik van de ingedikte variabele in 4 clusters doet dit effect weer gedeeltelijk teniet. Eenzelfde effect zien we, in mindere mate, bij 'Opleiding'.

Het clusteren van de categorieën van 'Software'

De variabele 'software' heeft 47 categorieën en daarmee is het aantal parameters erg groot. Er zijn 47 parameters voor de effecten op de categorie 'EBB vast, Polis flex' van de doelvariabele (46 zonder de referentiecategorie) en evenzoveel voor de categorie 'EBB flex, Polis vast'. Omdat de analyses apart uitgevoerd worden voor drie leeftijdscategorieën, zijn er voor ieder van de 47 softwarecategorieën 6 parameters, twee per leeftijdscategorie. Onderzocht is of de categorieën van 'Software' ingedeeld kunnen worden in groepen waarbinnen de waarden op de 6 parameters niet heel veel van elkaar verschillen. Hiervoor is een clusteranalyse toegepast met als objecten de 47 categorieën van software en als variabelen de 6 parameterwaarden. Er zijn clusterings met verschillende aantallen

clusters en volgens verschillende methoden (Average Linkage en methode van Ward) bekeken. Een op het oog goede clusterindeling is een indeling in 4 clusters volgens de methode van Ward. Bij deze methode wordt de totale binnen-cluster variantie geminimaliseerd wat leidt tot homogene clusters.

In onderstaande tabel staan de chi-kwadraatwaarden voor verschillende clusterings van de variabele 'Software'. De waarden voor 47 categorieën en 4 clusters zijn identiek aan die in de vorige tabel, de andere waarden laten zien dat als minder dan vier clusters worden gekozen, de verklaring sterk afneemt terwijl meer clusters dan 4 niet heel veel toevoegt.

Software clusters	Aantal parameters	Chi-kwadraat
Software in 47 categorieën	92	15237,9
8 software clusters	14	12395,4
7 software clusters	12	11581,9
6 software clusters	10	11266,9
5 software clusters	8	11072,1
4 software clusters	6	11026,8
3 software clusters	4	7874,6
2 software clusters	2	3504,3

De 4 clusters kennen de volgende aantallen software categorieën en banen:

Cluster	Aantal software categorieën	Aantal banen	Percentage banen
1	30	720923	80,4
2	12	127232	14,2
3	2	10549	1,2
4	3	28858	3,3

Er is één groot cluster waarin 30 van de 47 categorieën zijn opgenomen, waaronder de referentiecategorie en de categorie 'Overig'. Dit cluster bevat ook 80 procent van de banen. De overige clusters zijn veel kleiner en bevatten software categorieën met meer extreme parameterwaarden. In het vervolg is deze clustering in 4 categorieën gebruikt als de nieuwe variabele 'Softwarecluster', waarbij de eerste categorie de referentiecategorie is.

In onderstaande tabel zijn de parameterschattingen voor de 4 categorieën van 'Softwarecluster' opgenomen. De getoonde parameters zijn de multiplicatieve parameters en dus gelijk aan 1 voor de referentiecategorieën.

	15 tot 25	jaar	25 tot 55	jaar	55 jaar en	ouder
Software-cluster	EBB vast Polis flex	EBB flex Polis vast	EBB vast Polis flex	EBB flex Polis vast	EBB vast Polis flex	EBB flex Polis vast
1	1	1	1	1	1	1
2	0,91	3,25	1,23	2,65	1,16	2,52
3	1,43	0,57	6,45	1,16	8,00	1,28
4	1,30	0,69	3,20	1,62	3,57	1,23

De software in cluster 2 laat een toename zien van de categorie EBB-flex Polis vast voor alle leeftijdscategorieën.

De software in clusters 3 én 4 laten beide een forse toename zien van de kans op 'EBB vast, Polis flex' voor de leeftijdscategorieën 25-54 en 55 en ouder, en deze toename is veel sterker voor cluster 3 dan voor cluster 4. Daarnaast laten deze beide clusters een lagere kans zien voor 'EBB flex, Polis vast' voor '15 tot 25 jaar'.

Analyse met software-clusters per leeftijdscategorie.

In onderstaande tabel zijn weer de toetsingsgrootheden per variabele weergegeven, maar nu voor ieder van de drie leeftijdscategorieën apart. 'Softwarecluster' heeft in alle drie de leeftijdscategorieën het grootste effect van alle predictoren. Een gedetailleerdere beschrijving van deze modellen en hun parameters staat in hoofdstuk 4 van deze eindrapportage.

	15-25 jaar	25-55 jaar	55 jaar en ouder
Variabele	Chi-kwadraat		
Softwarecluster	2872,9	7217,6	1820,4
Contracturen	825,6	1517,6	1008,3
SBI-groep	1401,1	2246,2	853,0
Baanduur	295,7	4273,3	833,2
Leeftijdsklasse	32,0	1650,1	256,4
Kwartaalnummer	602,9	395,1	127,2
Opleidingsniveau	31,1	77,4	14,4
Migratieachtergrond	42,2	108,3	26,0
Geslacht	7,9	15,8	11,2
Waarneemmodus	49,2	32,9	9,9
Laatste respons	9,7	24,8	3,5
Proxywaarneming	11,5	71,2	7,5

Bijlage 4. De overige variabelen

In deze bijlage zijn de overige variabelen beschreven, is aangegeven welke categorieën deze bevatten en welke de referentiecategorie is in de MLR-analyses uit hoofdstuk 4 en (waar van toepassing) de meetmodellen uit hoofdstuk 5 en 6.

Variabele *Kwartaalnummer* die aangeeft in welk kwartaal de betreffende EBB-meting heeft plaatsgevonden. Daarmee wordt de tijd als eerste in de analyse opgenomen in de vorm van dummy-variabelen. Doordat de jaren 2006-2009 op de valreep zijn verwijderd lopen de kwartaalnummers van 17 (2010-1) t/m 52 (2018-4). Zowel kwartaalnummer 51 als 52 zijn de referentiecategorie. Dit komt doordat de variabele 'Laatste respons' bij de eerste 4 EBB-peilingen altijd dezelfde waarde heeft bij 2018-4, namelijk -1 (onbekend).

Variabele *Periode* die het tijdvak 2010-2018 opdeelt in zes perioden, op basis van vijf overgangen (mogelijke breekpunten) als gevolg van een wijziging in de EBB, Polis of wetgeving:

- (1) 1-1-2006 tot 1-1-2010 invoeringstraject van de Polis
- (2) Per 1-1-2010 Invoering van Computer-Assisted Telephone Interviewing (CATI) bij 1e peiling in de EBB, EBB-vraagstelling iedere peiling opnieuw gevraagd m.u.v. de categorie 'tijdelijk met uitzicht op vast', wijziging in banenafbakening
- (3) Per 1-1-2012 In de EBB wordt ook voor de categorie 'tijdelijk met uitzicht op vast' opnieuw gevraagd
- (4) 1-1- 2015 Aanpassing 12 uur naar 1 uur grens in de EBB
- (5) 1-1-2017 Met ingang van deze datum is in de Loonaangifte een extra code voor payrollling toegevoegd. Vanaf dat moment is het pas mogelijk om payrollers apart te onderscheiden. Uit eerdere analyses is gebleken dat het grootste deel voorheen als uitzendkracht geregistreerd werd
- (6) Vanaf 1 januari 2018. Bij een grote bedrijf werden tot deze datum vrijwel alle werknemers met een flexibel arbeidscontract geregistreerd, daarna volgens het werkelijke arbeidscontract.

Leeftijdscategorie. Dit is een verfijning van de drie leeftijdscategorieën 15- 24 jaar, 25- 54 jaar en 55+. De categorie 15-18 jaar is toegevoegd, omdat tijdens een proefanalyse voor 2016 onder jongeren is gebleken dat bij mensen van 15-18 jaar de kans op 'EBB vast en Polis Flex' kleiner is. Dit komt mogelijk doordat de EBB-vragen dan zijn beantwoord door de ouders (proxy waarneming). Het heeft de categorieën: (1) 15-18 jaar (2) 19-24 jaar (referentiecategorie) (3) 25-39 jaar (4) 40-54 jaar (referentiecategorie) (5) 55-64 jaar (6) 65-75 jaar (referentiecategorie). In dit geval zijn drie referentiecategorieën aangewezen, voor de 15- tot 25-jarigen, de 25- tot 55-jarigen en de 55-jarigen en ouder.

Geslacht. Deze heeft de categorieën (1) mannen (2) vrouwen (referentiecategorie)

Opleidingsniveau. Deze heeft de categorieën: (2) Middel (3) Hoog (4) Onbekend en (10) Laag (referentiecategorie).

Migratieachtergrond, met als categorieën: (2) Marokko (3) Turkije (4) Suriname (5) (Voormalige) Nederlandse Antillen (6) Overige niet-westerse landen (7) Overige westerse landen (8) Nederland (referentiecategorie).

Proxywaarneming, met als categorieën: (1) Nee (2) Ja (referentiecategorie).

Waarneemmodus, met als categorieën: (1) CAPI (2) CATI (6) CAWI (referentiecategorie).

Variabele *Laatste respons* die aangeeft of voor de desbetreffende persoon voor de laatste keer een EBB-meting heeft plaats gevonden. Als dit het geval is, heeft de betreffende persoon bij de tweede golf al niet meer gerespondeerd. De categorieën zijn: (-1) Onbekend (0) Nee (1) Ja (referentiecategorie).

Baanduur van de hoofdbaan. Deze variabele geeft aan hoe lang de respondent de hoofdbaan al heeft, met de volgende categorieën: (1) < 3 maanden (2) 3 maanden - < 6 maanden (3) 6 maanden - < 1 jaar (4) 1 jaar - < 2 jaar (5) 2 jaar = < 3 jaar (6) ≥ 3 jaar (referentiecategorie).

Contracturen van de hoofdbaan, met als categorieën: (1) < 12 uur (2) 12 - < 20 uur (3) 20 - < 30 uur (4) 30 uur en meer (referentiecategorie).

SBI-groep van de werkgever. Er wordt onderscheid gemaakt tussen SBI-groepen op het niveau van 2 cijfers. Er zijn SBI-groepen met veel jongeren, (clusters van) 2 digits SBI-groepen die opvielen bij MLR-analyses en een groep overig. De categorieën zijn:

- (1) 01-03 (Landbouw, bosbouw en visserij)
- (2) 47 (Detailhandel)
- (3) 50-51 (Vervoer over water en luchtvaart)
- (4) 55-56 (Horeca)
- (5) 64-66 (Financiële instellingen)
- (6) 78 (Arbeidsbemiddeling, uitzendbureaus en personeelsbeheer)
- (7) 84-85 (Overheids- en onderwijsinstellingen)
- (8) 86-88 (Zorginstellingen)
- (0) Overig (referentiecategorie)

Personen zonder werkgever zijn ingedeeld bij de referentiecategorie.

Bedrijfsgrootte van de werkgever, gebaseerd op het aantal werkzame personen bij het bedrijf, met als categorieën: (1) 10 tot 100 werkzame personen (middenbedrijf), (2) 100 of meer werkzame personen (grootbedrijf), (3) minder dan 10 werkzame personen (kleinbedrijf, referentiecategorie). Personen zonder werkgever zijn ingedeeld bij de referentiecategorie.

Bijlage 5. De MLR voor 15- tot 25-jarigen

	Ebb vast Polis flex					
	B	SE	df	Sig.	Exp(B)	
Intercept	-0,754	0,061	1	0,000		
Leeftijdsklasse=15 tot 19 jaar	-0,135	0,026	1	0,000	0,874	
Leeftijdsklasse=19 tot 25 jaar	0b		0			
SBI=Landbouw, bosbouw en visserij	-0,241	0,067	1	0,000	0,786	*
SBI=Detailhandel	0,200	0,026	1	0,000	1,221	
SBI=Vervoer over water en luchtvaart	0,122	0,177	1	0,490	1,130	
SBI=Horeca	-0,108	0,031	1	0,001	0,897	
SBI=Financiële instellingen	-0,272	0,108	1	0,012	0,762	
SBI=Arbeidsbemiddeling, uitzendbureaus en personeelsbeheer	-0,844	0,040	1	0,000	0,430	**
SBI=Overheids- en onderwijsinstellingen	0,213	0,043	1	0,000	1,238	
SBI=Zorginstellingen	-0,150	0,033	1	0,000	0,860	
Overige SBI	0b		0			
Contracturen=<12 uur	-0,336	0,024	1	0,000	0,715	*
Contracturen=12 tot 20 uur	-0,302	0,030	1	0,000	0,739	*
Contracturen=20 tot 30 uur	-0,168	0,030	1	0,000	0,846	
Contracturen=30 uur en meer	0b		0			
Laatste respons=onbekend	0,037	0,071	1	0,600	1,038	
Laatste respons=nee	-0,025	0,020	1	0,225	0,976	
Laatste respons=ja	0b		0			
migratieachtergrond=Marokko	-0,205	0,085	1	0,015	0,814	
migratieachtergrond=Turkije	-0,263	0,077	1	0,001	0,769	*
migratieachtergrond=Suriname	-0,137	0,073	1	0,059	0,872	
migratieachtergrond=Voorm. Nederlandse Antillen	-0,068	0,101	1	0,502	0,935	
migratieachtergrond=Overige niet-westerse landen	-0,056	0,053	1	0,294	0,946	
migratieachtergrond=Overige westerse landen	-0,045	0,039	1	0,253	0,956	
migratieachtergrond=Nederland	0b		0			
opleidingsniveau=middel	-0,028	0,022	1	0,208	0,973	
opleidingsniveau=hoog	-0,123	0,034	1	0,000	0,884	
opleidingsniveau=onbekend	0,031	0,139	1	0,824	1,031	
opleidingsniveau=laag	0b		0			
Baanduur=<3 maanden	-0,582	0,079	1	0,000	0,559	*
Baanduur=3 tot 6 maanden	-0,376	0,046	1	0,000	0,687	*
Baanduur=6 tot 12 maanden	-0,314	0,034	1	0,000	0,730	*

	Ebb vast Polis flex (vervolg)					
	B	SE	df	Sig.	Exp(B)	
Baanduur=12 tot 24 maanden	-0,171	0,025	1	0,000	0,843	
Baanduur=24 tot 36 maanden	-0,018	0,025	1	0,479	0,982	
Baanduur=36 maanden en langer	0b		0			
Proxywaarneming=nee	-0,029	0,020	1	0,157	0,972	
Proxywaarneming=ja	0b		0			
Waarneemmodus= CAPI	-0,127	0,022	1	0,000	0,880	
Waarneemmodus= CATI	-0,175	0,029	1	0,000	0,839	
Waarneemmodus= CAWI	0b		0			
geslacht=man	0,045	0,019	1	0,017	1,046	
geslacht=vrouw	0b		0			
Kwartaal 2010-1	0,344	0,065	1	0,000	1,411	*
Kwartaal 2010-2	0,391	0,066	1	0,000	1,478	*
Kwartaal 2010-3	0,388	0,076	1	0,000	1,474	*
Kwartaal 2010-4	0,323	0,076	1	0,000	1,382	*
Kwartaal 2011-1	0,348	0,075	1	0,000	1,416	*
Kwartaal 2011-2	0,460	0,073	1	0,000	1,584	*
Kwartaal 2011-3	0,242	0,075	1	0,001	1,274	*
Kwartaal 2011-4	0,297	0,076	1	0,000	1,345	*
Kwartaal 2012-1	0,196	0,078	1	0,012	1,216	
Kwartaal 2012-2	0,102	0,063	1	0,107	1,108	
Kwartaal 2012-3	0,086	0,064	1	0,178	1,090	
Kwartaal 2012-4	0,101	0,072	1	0,160	1,106	
Kwartaal 2013-1	0,158	0,072	1	0,028	1,171	
Kwartaal 2013-2	0,111	0,073	1	0,131	1,117	
Kwartaal 2013-3	-0,161	0,077	1	0,037	0,852	
Kwartaal 2013-4	0,012	0,075	1	0,872	1,012	
Kwartaal 2014-1	-0,030	0,076	1	0,694	0,971	
Kwartaal 2014-2	-0,075	0,077	1	0,329	0,928	
Kwartaal 2014-3	-0,140	0,078	1	0,073	0,869	
Kwartaal 2014-4	0,004	0,075	1	0,961	1,004	
Kwartaal 2015-1	0,009	0,074	1	0,901	1,009	
Kwartaal 2015-2	-0,063	0,076	1	0,408	0,939	
Kwartaal 2015-3	-0,149	0,076	1	0,052	0,862	
Kwartaal 2015-4	-0,069	0,074	1	0,351	0,933	
Kwartaal 2016-1	-0,051	0,075	1	0,495	0,950	
Kwartaal 2016-2	-0,192	0,075	1	0,011	0,825	
Kwartaal 2016-3	-0,184	0,078	1	0,018	0,832	
Kwartaal 2016-4	-0,074	0,075	1	0,324	0,929	
Kwartaal 2017-1	-0,124	0,075	1	0,098	0,883	
Kwartaal 2017-2	-0,246	0,077	1	0,001	0,782	*
Kwartaal 2017-3	-0,156	0,074	1	0,036	0,856	
Kwartaal 2017-4	-0,021	0,073	1	0,770	0,979	
Kwartaal 2018-1	-0,106	0,071	1	0,136	0,899	
Kwartaal 2018-2	0,076	0,070	1	0,279	1,079	

	Ebb vast Polis flex (slot)					
	B	SE	df	Sig.	Exp(B)	
Kwartaal 2018-3 en 2018-4	0b		0			
Softwarecluster=2,00	-0,092	0,022	1	0,000	0,912	
Softwarecluster=3,00	0,361	0,049	1	0,000	1,435	*
Softwarecluster=4,00	0,262	0,032	1	0,000	1,300	*
Softwarecluster=overig	0b		0			

	Ebb flex Polis vast					
	B	SE	df	Sig.	Exp(B)	
Intercept	-2,818	0,096	1	0,000		
Leeftijdsklasse=15 tot 19 jaar	-0,100	0,034	1	0,003	0,905	
Leeftijdsklasse=19 tot 25 jaar	0b		0			
SBI=Landbouw, bosbouw en visserij	0,496	0,066	1	0,000	1,643	*
SBI=Detailhandel	-0,071	0,035	1	0,045	0,932	
SBI=Vervoer over water en luchtvaart	0,511	0,238	1	0,032	1,667	
SBI=Horeca	0,225	0,038	1	0,000	1,253	*
SBI=Financiële instellingen	-0,287	0,163	1	0,078	0,751	
SBI=Arbeidsbemiddeling, uitzendbureaus en personeelsbeheer	-1,020	0,062	1	0,000	0,361	**
SBI=Overheids- en onderwijsinstellingen	-0,187	0,079	1	0,017	0,829	
SBI=Zorginstellingen	-0,405	0,050	1	0,000	0,667	*
Overige SBI	0b		0			
Contracturen=<12 uur	0,748	0,036	1	0,000	2,113	**
Contracturen=12 tot 20 uur	0,719	0,042	1	0,000	2,052	**
Contracturen=20 tot 30 uur	0,452	0,046	1	0,000	1,571	*
Contracturen=30 uur en meer	0b		0			
Laatste respons=onbekend	-0,076	0,116	1	0,513	0,927	
Laatste respons=nee	-0,086	0,028	1	0,002	0,917	
Laatste respons=ja	0b		0			
migratieachtergrond=Marokko	0,086	0,109	1	0,429	1,090	
migratieachtergrond=Turkije	-0,187	0,112	1	0,095	0,829	
migratieachtergrond=Suriname	-0,334	0,112	1	0,003	0,716	*
migratieachtergrond=Voorm. Nederlandse Antillen	0,036	0,137	1	0,793	1,037	
migratieachtergrond=Overige niet-westerse landen	0,155	0,068	1	0,023	1,168	
migratieachtergrond=Overige westerse landen	-0,091	0,056	1	0,104	0,913	
migratieachtergrond=Nederland	0b		0			
opleidingsniveau=middel	0,034	0,031	1	0,268	1,035	
opleidingsniveau=hoog	0,150	0,049	1	0,002	1,162	
opleidingsniveau=onbekend	-0,397	0,260	1	0,126	0,672	
opleidingsniveau=laag	0b		0			
Baanduur=<3 maanden	-0,360	0,098	1	0,000	0,698	*

	Ebb flex Polis vast (vervolg)					
	B	SE	df	Sig.	Exp(B)	
Baanduur=6 tot 12 maanden	-0,354	0,048	1	0,000	0,702	*
Baanduur=12 tot 24 maanden	-0,258	0,035	1	0,000	0,773	*
Baanduur=24 tot 36 maanden	-0,148	0,036	1	0,000	0,862	
Baanduur=36 maanden en langer	0b		0			
Proxywaarneming=nee	-0,096	0,029	1	0,001	0,909	
Proxywaarneming=ja	0b		0			
Waarneemmodus= CAPI	0,034	0,032	1	0,282	1,035	
Waarneemmodus= CATI	0,019	0,039	1	0,634	1,019	
Waarneemmodus= CAWI	0b		0			
geslacht=man	0,049	0,026	1	0,060	1,050	
geslacht=vrouw	0b		0			
Kwartaal 2010-1	0,337	0,100	1	0,001	1,401	*
Kwartaal 2010-2	0,362	0,100	1	0,000	1,437	*
Kwartaal 2010-3	0,293	0,114	1	0,011	1,340	
Kwartaal 2010-4	0,399	0,111	1	0,000	1,490	*
Kwartaal 2011-1	0,455	0,111	1	0,000	1,576	*
Kwartaal 2011-2	0,589	0,108	1	0,000	1,802	*
Kwartaal 2011-3	0,364	0,111	1	0,001	1,440	*
Kwartaal 2011-4	0,563	0,110	1	0,000	1,756	*
Kwartaal 2012-1	0,475	0,112	1	0,000	1,608	*
Kwartaal 2012-2	0,369	0,096	1	0,000	1,447	*
Kwartaal 2012-3	0,492	0,095	1	0,000	1,636	*
Kwartaal 2012-4	0,450	0,106	1	0,000	1,569	*
Kwartaal 2013-1	0,361	0,108	1	0,001	1,434	*
Kwartaal 2013-2	0,539	0,106	1	0,000	1,714	*
Kwartaal 2013-3	0,251	0,111	1	0,023	1,286	
Kwartaal 2013-4	0,379	0,110	1	0,001	1,460	*
Kwartaal 2014-1	0,250	0,113	1	0,027	1,284	
Kwartaal 2014-2	0,272	0,111	1	0,014	1,313	
Kwartaal 2014-3	0,395	0,109	1	0,000	1,484	*
Kwartaal 2014-4	0,162	0,114	1	0,153	1,176	
Kwartaal 2015-1	0,166	0,113	1	0,142	1,180	
Kwartaal 2015-2	0,157	0,112	1	0,162	1,170	
Kwartaal 2015-3	0,103	0,114	1	0,369	1,108	
Kwartaal 2015-4	-0,262	0,121	1	0,030	0,770	
Kwartaal 2016-1	-0,095	0,120	1	0,429	0,909	
Kwartaal 2016-2	0,052	0,113	1	0,645	1,053	
Kwartaal 2016-3	-0,006	0,119	1	0,959	0,994	
Kwartaal 2016-4	-0,105	0,119	1	0,379	0,901	
Kwartaal 2017-1	0,080	0,115	1	0,489	1,083	
Kwartaal 2017-2	0,020	0,117	1	0,865	1,020	
Kwartaal 2017-3	-0,163	0,120	1	0,175	0,849	

	Ebb flex Polis vast (slot)					
	B	SE	df	Sig.	Exp(B)	
Kwartaal 2017-4	-0,030	0,118	1	0,799	0,970	
Kwartaal 2018-1	-0,114	0,114	1	0,318	0,892	
Kwartaal 2018-2	-0,019	0,115	1	0,868	0,981	
Kwartaal 2018-3 en 2018-4	0b		0			
Softwarecluster=2,00	1,179	0,027	1	0,000	3,253	**
Softwarecluster=3,00	-0,564	0,108	1	0,000	0,569	*
Softwarecluster=4,00	-0,370	0,061	1	0,000	0,691	*
Softwarecluster=overig	0b		0			

Bijlage 6. De MLR voor 25- tot 55-jarigen

	Ebb vast Polis flex					
	B	SE	df	Sig.	Exp(B)	
Intercept	-2,532	0,041	1	0,000		
Leeftijdsklasse=25 tot 40 jaar	0,462	0,012	1	0,000	1,587	*
Leeftijdsklasse=40 tot 55 jaar	0b		0			
SBI=Landbouw, bosbouw en visserij	0,152	0,059	1	0,011	1,164	
SBI=Detailhandel	0,562	0,021	1	0,000	1,753	*
SBI=Vervoer over water en luchtvaart	-0,558	0,099	1	0,000	0,572	*
SBI=Horeca	0,479	0,032	1	0,000	1,614	*
SBI=Financiële instellingen	-0,558	0,037	1	0,000	0,572	*
SBI=Arbeidsbemiddeling, uitzendbureaus en personeelsbeheer	-0,500	0,029	1	0,000	0,607	*
SBI=Overheids- en onderwijsinstellingen	-0,363	0,021	1	0,000	0,696	*
SBI=Zorginstellingen	-0,059	0,017	1	0,001	0,943	
Overige SBI	0b		0			
Contracturen=<12 uur	0,334	0,026	1	0,000	1,396	*
Contracturen=12 tot 20 uur	0,114	0,021	1	0,000	1,120	
Contracturen=20 tot 30 uur	-0,050	0,017	1	0,003	0,951	
Contracturen=30 uur en meer	0b		0			
Laatste respons=onbekend	0,067	0,047	1	0,159	1,069	
Laatste respons=nee	-0,066	0,013	1	0,000	0,936	
Laatste respons=ja	0b		0			
migratieachtergrond=Marokko	-0,138	0,056	1	0,014	0,871	
migratieachtergrond=Turkije	-0,051	0,047	1	0,276	0,950	
migratieachtergrond=Suriname	-0,073	0,042	1	0,080	0,929	
migratieachtergrond=Voorm. Nederlandse Antillen	-0,021	0,063	1	0,737	0,979	
migratieachtergrond=Overige niet-westerse landen	0,034	0,032	1	0,289	1,034	
migratieachtergrond=Overige westerse landen	0,087	0,021	1	0,000	1,091	
migratieachtergrond=Nederland	0b		0			
opleidingsniveau=middel	-0,105	0,016	1	0,000	0,901	
opleidingsniveau=hoog	-0,070	0,017	1	0,000	0,932	
opleidingsniveau=onbekend	-0,096	0,054	1	0,076	0,909	
opleidingsniveau=laag	0b		0			
Baanduur=<3 maanden	0,302	0,105	1	0,004	1,353	*
Baanduur=3 tot 6 maanden	0,533	0,057	1	0,000	1,704	*
Baanduur=6 tot 12 maanden	0,703	0,033	1	0,000	2,019	**

	Ebb vast Polis flex (vervolg)					
	B	SE	df	Sig.	Exp(B)	
Baanduur=12 tot 24 maanden	0,842	0,021	1	0,000	2,322	**
Baanduur=24 tot 36 maanden	0,825	0,021	1	0,000	2,283	**
Baanduur=36 maanden en langer	0b		0			
Proxywaarneming=nee	-0,104	0,012	1	0,000	0,902	
Proxywaarneming=ja	0b		0			
Waarneemmodus= CAPI	0,007	0,016	1	0,659	1,007	
Waarneemmodus= CATI	-0,064	0,017	1	0,000	0,938	
Waarneemmodus= CAWI	0b		0			
geslacht=man	-0,054	0,014	1	0,000	0,947	
geslacht=vrouw	0b		0			
Kwartaal 2010-1	0,209	0,045	1	0,000	1,232	
Kwartaal 2010-2	0,147	0,046	1	0,001	1,159	
Kwartaal 2010-3	0,210	0,051	1	0,000	1,234	
Kwartaal 2010-4	0,244	0,051	1	0,000	1,276	*
Kwartaal 2011-1	0,176	0,052	1	0,001	1,192	
Kwartaal 2011-2	0,209	0,051	1	0,000	1,233	
Kwartaal 2011-3	0,208	0,052	1	0,000	1,231	
Kwartaal 2011-4	0,280	0,051	1	0,000	1,323	*
Kwartaal 2012-1	0,222	0,052	1	0,000	1,249	
Kwartaal 2012-2	0,231	0,043	1	0,000	1,259	*
Kwartaal 2012-3	0,195	0,043	1	0,000	1,215	
Kwartaal 2012-4	0,163	0,048	1	0,001	1,178	
Kwartaal 2013-1	0,192	0,048	1	0,000	1,212	
Kwartaal 2013-2	0,209	0,049	1	0,000	1,232	
Kwartaal 2013-3	0,132	0,050	1	0,008	1,141	
Kwartaal 2013-4	0,200	0,050	1	0,000	1,221	
Kwartaal 2014-1	0,152	0,050	1	0,002	1,165	
Kwartaal 2014-2	0,053	0,052	1	0,311	1,054	
Kwartaal 2014-3	0,198	0,050	1	0,000	1,219	
Kwartaal 2014-4	0,130	0,051	1	0,010	1,138	
Kwartaal 2015-1	0,193	0,049	1	0,000	1,213	
Kwartaal 2015-2	0,191	0,050	1	0,000	1,210	
Kwartaal 2015-3	0,183	0,050	1	0,000	1,200	
Kwartaal 2015-4	0,211	0,049	1	0,000	1,235	
Kwartaal 2016-1	0,226	0,049	1	0,000	1,254	*
Kwartaal 2016-2	0,200	0,049	1	0,000	1,221	
Kwartaal 2016-3	0,112	0,051	1	0,028	1,118	
Kwartaal 2016-4	0,167	0,050	1	0,001	1,182	
Kwartaal 2017-1	0,065	0,050	1	0,197	1,067	
Kwartaal 2017-2	0,046	0,050	1	0,361	1,047	
Kwartaal 2017-3	0,139	0,049	1	0,005	1,149	
Kwartaal 2017-4	0,065	0,048	1	0,183	1,067	
Kwartaal 2018-1	-0,037	0,047	1	0,434	0,964	
Kwartaal 2018-2	0,006	0,048	1	0,904	1,006	

	Ebb vast Polis flex (slot)					
	B	SE	df	Sig.	Exp(B)	
Kwartaal 2018-3 en 2018-4	0b		0			
Softwarecluster=2,00	0,208	0,014	1	0,000	1,231	
Softwarecluster=3,00	1,863	0,030	1	0,000	6,445	**
Softwarecluster=4,00	1,163	0,021	1	0,000	3,199	**
Softwarecluster=overig	0b		0			

	Ebb flex Polis vast					
	B	SE	df	Sig.	Exp(B)	
Intercept	-4,841	0,087	1	0,000		
Leeftijdsklasse=25 tot 40 jaar	0,349	0,023	1	0,000	1,418	*
Leeftijdsklasse=40 tot 55 jaar	0b		0			
SBI=Landbouw, bosbouw en visserij	-0,159	0,126	1	0,206	0,853	
SBI=Detailhandel	0,153	0,044	1	0,001	1,165	
SBI=Vervoer over water en luchtvaart	-0,724	0,232	1	0,002	0,485	**
SBI=Horeca	0,482	0,053	1	0,000	1,620	*
SBI=Financiële instellingen	-0,606	0,079	1	0,000	0,545	*
SBI=Arbeidsbemiddeling, uitzendbureaus en personeelsbeheer	-0,432	0,049	1	0,000	0,649	*
SBI=Overheids- en onderwijsinstellingen	-0,357	0,044	1	0,000	0,700	*
SBI=Zorginstellingen	-0,255	0,033	1	0,000	0,775	*
Overige SBI	0b		0			
Contracturen=<12 uur	1,442	0,037	1	0,000	4,229	**
Contracturen=12 tot 20 uur	0,656	0,038	1	0,000	1,926	*
Contracturen=20 tot 30 uur	0,207	0,034	1	0,000	1,230	
Contracturen=30 uur en meer	0b		0			
Laatste respons=onbekend	0,013	0,105	1	0,903	1,013	
Laatste respons=nee	-0,032	0,026	1	0,219	0,968	
Laatste respons=ja	0b		0			
migratieachtergrond=Marokko	0,102	0,100	1	0,305	1,108	
migratieachtergrond=Turkije	0,230	0,082	1	0,005	1,258	*
migratieachtergrond=Suriname	0,125	0,077	1	0,104	1,133	
migratieachtergrond=Voorm. Nederlandse Antillen	0,144	0,112	1	0,198	1,155	
migratieachtergrond=Overige niet-westerse landen	0,306	0,053	1	0,000	1,358	*
migratieachtergrond=Overige westerse landen	0,277	0,038	1	0,000	1,319	*
migratieachtergrond=Nederland	0b		0			
opleidingsniveau=middel	-0,105	0,031	1	0,001	0,900	
opleidingsniveau=hoog	-0,061	0,033	1	0,067	0,941	
opleidingsniveau=onbekend	0,383	0,085	1	0,000	1,466	*
opleidingsniveau=laag	0b		0			
Baanduur=<3 maanden	1,170	0,136	1	0,000	3,223	**
Baanduur=3 tot 6 maanden	0,921	0,094	1	0,000	2,513	**

	Ebb flex Polis vast (vervolg)					
	B	SE	df	Sig.	Exp(B)	
Baanduur=6 tot 12 maanden	1,322	0,052	1	0,000	3,750	**
Baanduur=12 tot 24 maanden	1,319	0,036	1	0,000	3,741	**
Baanduur=24 tot 36 maanden	1,215	0,037	1	0,000	3,371	**
Baanduur=36 maanden en langer	0b		0			
Proxywaarneming=nee	0,002	0,024	1	0,921	1,002	
Proxywaarneming=ja	0b		0			
Waarneemmodus= CAPI	0,070	0,030	1	0,022	1,072	
Waarneemmodus= CATI	-0,031	0,034	1	0,357	0,970	
Waarneemmodus= CAWI	0b		0			
geslacht=man	-0,043	0,028	1	0,126	0,958	
geslacht=vrouw	0b		0			
Kwartaal 2010-1	0,580	0,092	1	0,000	1,786	*
Kwartaal 2010-2	0,424	0,095	1	0,000	1,528	*
Kwartaal 2010-3	0,491	0,105	1	0,000	1,633	*
Kwartaal 2010-4	0,530	0,104	1	0,000	1,699	*
Kwartaal 2011-1	0,595	0,103	1	0,000	1,813	*
Kwartaal 2011-2	0,531	0,104	1	0,000	1,701	*
Kwartaal 2011-3	0,468	0,107	1	0,000	1,596	*
Kwartaal 2011-4	0,620	0,103	1	0,000	1,859	*
Kwartaal 2012-1	0,704	0,103	1	0,000	2,021	**
Kwartaal 2012-2	0,664	0,088	1	0,000	1,943	*
Kwartaal 2012-3	0,570	0,089	1	0,000	1,768	*
Kwartaal 2012-4	0,657	0,096	1	0,000	1,929	*
Kwartaal 2013-1	0,571	0,099	1	0,000	1,769	*
Kwartaal 2013-2	0,671	0,098	1	0,000	1,956	*
Kwartaal 2013-3	0,621	0,099	1	0,000	1,860	*
Kwartaal 2013-4	0,660	0,099	1	0,000	1,936	*
Kwartaal 2014-1	0,618	0,101	1	0,000	1,856	*
Kwartaal 2014-2	0,703	0,100	1	0,000	2,021	**
Kwartaal 2014-3	0,567	0,103	1	0,000	1,763	*
Kwartaal 2014-4	0,508	0,103	1	0,000	1,662	*
Kwartaal 2015-1	0,494	0,103	1	0,000	1,639	*
Kwartaal 2015-2	0,431	0,105	1	0,000	1,539	*
Kwartaal 2015-3	0,426	0,104	1	0,000	1,531	*
Kwartaal 2015-4	0,421	0,104	1	0,000	1,523	*
Kwartaal 2016-1	0,416	0,105	1	0,000	1,516	*
Kwartaal 2016-2	0,230	0,108	1	0,033	1,258	
Kwartaal 2016-3	0,320	0,107	1	0,003	1,377	*
Kwartaal 2016-4	0,273	0,108	1	0,011	1,314	
Kwartaal 2017-1	0,257	0,107	1	0,016	1,293	
Kwartaal 2017-2	0,077	0,111	1	0,486	1,080	
Kwartaal 2017-3	0,094	0,109	1	0,387	1,099	

	Ebb flex Polis vast (slot)					
	B	Std. Error	df	Sig.	Exp(B)	
Kwartaal 2017-4	0,154	0,104	1	0,139	1,167	
Kwartaal 2018-1	0,130	0,101	1	0,199	1,139	
Kwartaal 2018-2	0,127	0,103	1	0,216	1,135	
Kwartaal 2018-3 en 2018-4	0b		0			
Softwarecluster=2,00	0,974	0,024	1	0,000	2,649	**
Softwarecluster=3,00	0,144	0,105	1	0,171	1,155	
Softwarecluster=4,00	0,480	0,052	1	0,000	1,616	*
Softwarecluster=overig	0b		0			

Bijlage 7. De MLR voor 55-jarigen en ouder

	Ebb vast Polis flex					
	B	SE	df	Sig.	Exp(B)	
Intercept	-2,508	0,106	1	0,000		
Leeftijdsklasse=55 tot 65 jaar	-0,475	0,060	1	0,000	0,622	*
Leeftijdsklasse=65 tot 75 jaar	0b		0			
SBI=Landbouw, bosbouw en visserij	0,290	0,128	1	0,023	1,337	
SBI=Detailhandel	0,595	0,051	1	0,000	1,813	*
SBI=Vervoer over water en luchtvaart	-0,683	0,251	1	0,006	0,505	*
SBI=Horeca	0,321	0,084	1	0,000	1,378	*
SBI=Financiële instellingen	-0,488	0,099	1	0,000	0,614	*
SBI=Arbeidsbemiddeling, uitzendbureaus en personeelsbeheer	-0,405	0,069	1	0,000	0,667	*
SBI=Overheids- en onderwijsinstellingen	-1,013	0,051	1	0,000	0,363	**
SBI=Zorginstellingen	-0,365	0,041	1	0,000	0,695	*
Overige SBI	0b		0			
Contracturen=<12 uur	0,666	0,048	1	0,000	1,947	*
Contracturen=12 tot 20 uur	0,523	0,044	1	0,000	1,687	*
Contracturen=20 tot 30 uur	0,256	0,040	1	0,000	1,292	*
Contracturen=30 uur en meer	0b		0			
Laatste respons=onbekend	-0,044	0,099	1	0,657	0,957	
Laatste respons=nee	-0,055	0,032	1	0,082	0,946	
Laatste respons=ja	0b		0			
migratieachtergrond=Marokko	0,181	0,278	1	0,514	1,199	
migratieachtergrond=Turkije	0,022	0,282	1	0,937	1,023	
migratieachtergrond=Suriname	0,029	0,115	1	0,803	1,029	
migratieachtergrond=Voorm. Nederlandse Antillen	-0,247	0,210	1	0,239	0,781	
migratieachtergrond=Overige niet-westerse landen	0,291	0,107	1	0,007	1,338	*
migratieachtergrond=Overige westerse landen	0,074	0,048	1	0,127	1,077	
migratieachtergrond=Nederland	0b		0			
opleidingsniveau=middel	-0,038	0,033	1	0,241	0,962	
opleidingsniveau=hoog	0,035	0,038	1	0,351	1,036	
opleidingsniveau=onbekend	-0,034	0,115	1	0,769	0,967	
opleidingsniveau=laag	0b		0			
Baanduur=<3 maanden	0,778	0,314	1	0,013	2,177	
Baanduur=3 tot 6 maanden	1,066	0,160	1	0,000	2,903	**
Baanduur=6 tot 12 maanden	1,232	0,092	1	0,000	3,429	**

	Ebb vast Polis flex (vervolg)					
	B	SE	df	Sig.	Exp(B)	
Baanduur=12 tot 24 maanden	1,309	0,059	1	0,000	3,702	**
Baanduur=24 tot 36 maanden	1,211	0,058	1	0,000	3,357	**
Baanduur=36 maanden en langer	0b		0			
Proxywaarneming=nee	-0,080	0,029	1	0,007	0,923	
Proxywaarneming=ja	0b		0			
Waarneemmodus= CAPI	0,005	0,043	1	0,903	1,005	
Waarneemmodus= CATI	0,014	0,038	1	0,719	1,014	
Waarneemmodus= CAWI	0b		0			
geslacht=man	0,016	0,035	1	0,647	1,016	
geslacht=vrouw	0b		0			
Kwartaal 2010-1	0,255	0,106	1	0,016	1,291	
Kwartaal 2010-2	0,134	0,111	1	0,228	1,143	
Kwartaal 2010-3	0,208	0,129	1	0,108	1,231	
Kwartaal 2010-4	0,291	0,120	1	0,015	1,338	
Kwartaal 2011-1	0,194	0,124	1	0,118	1,214	
Kwartaal 2011-2	0,205	0,124	1	0,100	1,227	
Kwartaal 2011-3	0,044	0,134	1	0,740	1,045	
Kwartaal 2011-4	0,260	0,122	1	0,034	1,297	
Kwartaal 2012-1	0,146	0,130	1	0,261	1,157	
Kwartaal 2012-2	0,185	0,095	1	0,052	1,204	
Kwartaal 2012-3	-0,038	0,099	1	0,703	0,963	
Kwartaal 2012-4	0,091	0,109	1	0,403	1,096	
Kwartaal 2013-1	-0,040	0,115	1	0,731	0,961	
Kwartaal 2013-2	0,060	0,112	1	0,592	1,062	
Kwartaal 2013-3	0,170	0,108	1	0,116	1,185	
Kwartaal 2013-4	0,222	0,109	1	0,041	1,249	
Kwartaal 2014-1	0,036	0,114	1	0,752	1,037	
Kwartaal 2014-2	0,282	0,107	1	0,009	1,326	*
Kwartaal 2014-3	0,210	0,108	1	0,053	1,234	
Kwartaal 2014-4	0,136	0,109	1	0,211	1,146	
Kwartaal 2015-1	0,271	0,105	1	0,010	1,312	*
Kwartaal 2015-2	0,067	0,111	1	0,548	1,069	
Kwartaal 2015-3	0,058	0,109	1	0,597	1,059	
Kwartaal 2015-4	0,164	0,104	1	0,113	1,179	
Kwartaal 2016-1	0,252	0,102	1	0,013	1,287	
Kwartaal 2016-2	0,267	0,101	1	0,008	1,306	*
Kwartaal 2016-3	0,177	0,104	1	0,088	1,193	
Kwartaal 2016-4	0,119	0,104	1	0,252	1,127	
Kwartaal 2017-1	-0,006	0,107	1	0,954	0,994	
Kwartaal 2017-2	0,007	0,105	1	0,946	1,007	
Kwartaal 2017-3	0,123	0,101	1	0,220	1,131	
Kwartaal 2017-4	0,073	0,101	1	0,466	1,076	
Kwartaal 2018-1	-0,099	0,100	1	0,323	0,906	
Kwartaal 2018-2	0,033	0,099	1	0,735	1,034	

	Ebb vast Polis flex (slot)					
	B	SE	df	Sig.	Exp(B)	
Kwartaal 2018-3 en 2018-4	0b		0			
Softwarecluster=2,00	0,145	0,034	1	0,000	1,156	
Softwarecluster=3,00	2,079	0,061	1	0,000	7,998	*
Softwarecluster=4,00	1,271	0,049	1	0,000	3,565	**
Softwarecluster=overig	0b		0			

	Ebb flex Polis vast (start)					
	B	SE	df	Sig.	Exp(B)	
Intercept	-3,791	0,169	1	0,000		
Leeftijdsklasse=55 tot 65 jaar	-1,133	0,072	1	0,000	0,322	**
Leeftijdsklasse=65 tot 75 jaar	0b		0			
SBI=Landbouw, bosbouw en visserij	0,317	0,189	1	0,094	1,372	
SBI=Detailhandel	-0,015	0,096	1	0,878	0,985	
SBI=Vervoer over water en luchtvaart	0,197	0,314	1	0,531	1,217	
SBI=Horeca	0,342	0,123	1	0,006	1,407	*
SBI=Financiële instellingen	-0,428	0,172	1	0,013	0,652	
SBI=Arbeidsbemiddeling, uitzendbureaus en personeelsbeheer	-0,590	0,110	1	0,000	0,555	*
SBI=Overheids- en onderwijsinstellingen	-0,600	0,082	1	0,000	0,549	*
SBI=Zorginstellingen	-0,392	0,066	1	0,000	0,676	*
Overige SBI	0b		0			
Contracturen=<12 uur	1,941	0,068	1	0,000	6,969	**
Contracturen=12 tot 20 uur	1,219	0,071	1	0,000	3,385	**
Contracturen=20 tot 30 uur	0,608	0,072	1	0,000	1,836	*
Contracturen=30 uur en meer	0b		0			
Laatste respons=onbekend	-0,240	0,184	1	0,194	0,787	
Laatste respons=nee	-0,042	0,052	1	0,416	0,958	
Laatste respons=ja	0b		0			
migratieachtergrond=Marokko	0,179	0,479	1	0,708	1,196	
migratieachtergrond=Turkije	0,000	0,518	1	0,999	1,000	
migratieachtergrond=Suriname	0,185	0,193	1	0,337	1,204	
migratieachtergrond=Voorm. Nederlandse Antillen	0,401	0,281	1	0,154	1,493	
migratieachtergrond=Overige niet-westerse landen	0,645	0,162	1	0,000	1,906	*
migratieachtergrond=Overige westerse landen	0,085	0,081	1	0,295	1,088	
migratieachtergrond=Nederland	0b		0			
opleidingsniveau=middel	0,025	0,053	1	0,636	1,026	
opleidingsniveau=hoog	-0,036	0,065	1	0,582	0,965	
opleidingsniveau=onbekend	0,455	0,152	1	0,003	1,577	*
opleidingsniveau=laag	0b		0			
Baanduur=<3 maanden	1,002	0,424	1	0,018	2,723	
Baanduur=3 tot 6 maanden	1,093	0,252	1	0,000	2,984	**

	Ebb flex Polis vast (vervolg)					
	B	SE	df	Sig.	Exp(B)	
Baanduur=6 tot 12 maanden	1,146	0,148	1	0,000	3,146	**
Baanduur=12 tot 24 maanden	1,214	0,092	1	0,000	3,366	**
Baanduur=24 tot 36 maanden	0,971	0,093	1	0,000	2,642	**
Baanduur=36 maanden en langer	0b		0			
Proxywaarneming=nee	0,012	0,050	1	0,815	1,012	
Proxywaarneming=ja	0b		0			
Waarneemmodus= CAPI	-0,122	0,072	1	0,090	0,885	
Waarneemmodus= CATI	-0,200	0,066	1	0,002	0,819	
Waarneemmodus= CAWI	0b		0			
geslacht=man	0,193	0,058	1	0,001	1,213	
geslacht=vrouw	0b		0			
Kwartaal 2010-1	0,479	0,179	1	0,007	1,615	*
Kwartaal 2010-2	0,224	0,192	1	0,245	1,251	
Kwartaal 2010-3	0,703	0,203	1	0,001	2,020	**
Kwartaal 2010-4	0,255	0,215	1	0,237	1,290	
Kwartaal 2011-1	0,520	0,206	1	0,012	1,682	
Kwartaal 2011-2	0,641	0,199	1	0,001	1,898	*
Kwartaal 2011-3	0,444	0,217	1	0,041	1,559	
Kwartaal 2011-4	0,355	0,213	1	0,096	1,426	
Kwartaal 2012-1	0,759	0,198	1	0,000	2,137	**
Kwartaal 2012-2	0,406	0,162	1	0,012	1,500	
Kwartaal 2012-3	0,299	0,163	1	0,067	1,348	
Kwartaal 2012-4	0,238	0,185	1	0,199	1,268	
Kwartaal 2013-1	0,480	0,174	1	0,006	1,616	*
Kwartaal 2013-2	0,437	0,178	1	0,014	1,548	
Kwartaal 2013-3	0,266	0,184	1	0,148	1,305	
Kwartaal 2013-4	0,469	0,178	1	0,008	1,598	*
Kwartaal 2014-1	0,287	0,186	1	0,123	1,332	
Kwartaal 2014-2	0,013	0,201	1	0,947	1,013	
Kwartaal 2014-3	0,390	0,184	1	0,033	1,477	
Kwartaal 2014-4	-0,004	0,201	1	0,983	0,996	
Kwartaal 2015-1	0,487	0,176	1	0,006	1,627	*
Kwartaal 2015-2	0,471	0,177	1	0,008	1,601	*
Kwartaal 2015-3	0,272	0,181	1	0,133	1,312	
Kwartaal 2015-4	0,320	0,177	1	0,071	1,377	
Kwartaal 2016-1	0,188	0,182	1	0,303	1,207	
Kwartaal 2016-2	0,433	0,171	1	0,011	1,542	
Kwartaal 2016-3	0,110	0,186	1	0,555	1,116	
Kwartaal 2016-4	0,167	0,185	1	0,366	1,182	
Kwartaal 2017-1	0,011	0,190	1	0,955	1,011	
Kwartaal 2017-2	0,191	0,178	1	0,284	1,211	
Kwartaal 2017-3	-0,084	0,188	1	0,655	0,919	

	Ebb flex Polis vast (slot)					
	B	SE	df	Sig.	Exp(B)	
Kwartaal 2017-4	0,022	0,183	1	0,904	1,022	
Kwartaal 2018-1	0,124	0,169	1	0,463	1,132	
Kwartaal 2018-2	-0,091	0,182	1	0,618	0,913	
Kwartaal 2018-3 en 2018-4	0b		0			
Softwarecluster=2,00	0,923	0,049	1	0,000	2,517	**
Softwarecluster=3,00	0,245	0,184	1	0,185	1,277	
Softwarecluster=4,00	0,204	0,115	1	0,078	1,226	
Softwarecluster=overig	0b		0			

Bijlage 8. De fit van de verschillende HMM-modellen

	2007*	2009*	2016	2017	2018	N par.
Model	BIC					
A': ICE Error EBB (1)	90031	66007	91460	85099	89873	44
A': ICE Error Polis (2)	147264	114052	146503	136811	142244	44
A: ICE Error EBB and Polis (3)	89363	65792	91221	84996	89737	50
B': non-ICE error EBB (4)	89113	65438	91118	84977	89683	54
B': non-ICE Error Polis (5)	79140	58408	85528	79849	85587	56
B: non-ICE Error EBB and Polis (6)	79157	58430	85516	79822	85543	60
C': non-ICE Error Polis with covariates (7)	79468	58836	85910	80171	85895	122
C': non-ICE Error Polis with covariates also initial state (8)	78666	58347	84804	79188	84777	144
C: non-ICE Error EBB and Polis with cov. also initial state (9)	78734	58367	84739	79162	84757	148
*: er gaat één vrijheidsgraad verloren door een actieve ongelijkheidsrestrictie op de foutkansen bij de EBB in de modellen A' en A)						

Bijlage 9. Bijdrage van de covariaten aan de onnauwkeurigheden in de meting in de Polis en in de EBB

De bijdrage van variabelen aan de verklaring van de onnauwkeurigheid EBB-meting is vast en modelschatting is flexibel voor 25- tot 55-jarigen in 2017 en 2018

Variabele	Chi-kwadraat waarde	Aantal vrijheidsgraden
Zelfde werkgever EBB	1552,5	2
SBI-groep	1336,6	6
Bedrijfsgrootte	505,5	1
Baanduur	316,1	3
Contracturen	219,8	2
Kwartaalnummer	132,0	7
Migratieachtergrond	121,6	8
Geslacht	43,9	1
Leeftijdsklasse	24,4	1
Opleidingsniveau	9,5	1

Daar hoort de volgende interpretatie bij:

Variabele	Kans op EBB vast, model flex
Zelfde werkgever EBB	Werknemers die dezelfde werkgever houden hebben een veel grotere kans op deze onnauwkeurigheid.
SBI-groep	Arbeidsbemiddeling, uitzendbureaus en personeelsbeheer alsmede Landbouw, bosbouw en visserij hebben een (veel) kleinere kans en Vervoer over water en luchtvaart een grotere kans hierop
Bedrijfsgrootte	Hoe groter het bedrijf hoe kleiner de kans op deze onnauwkeurigheid
Baanduur	Hoe langer de baanduur, hoe groter de kans op deze onnauwkeurigheid
Contractduur	Bij kleinere contracten is de kans op deze onnauwkeurigheid ook kleiner
Kwartaalnummer	Kans neemt monotoon toe tussen eerste kwartaal van 2017 en laatste kwartaal van 2018.
Migratieachtergrond	Bij alle niet-westerse migratieachtergronden is de kans op deze onnauwkeurigheid kleiner
Geslacht	niet belangrijk
Opleidingsniveau	niet belangrijk
Leeftijd	niet belangrijk

De bijdrage van variabelen aan de verklaring van de onnauwkeurigheid Polis-meting is vast en modelschatting is flexibel voor de 25- tot 55-jarigen in 2017 en 2018

Variabele	Chi-kwadraat waarde	Aantal vrijheids-graden
Softwarecluster	4020,4	4
Baanduur	1159,9	1
SBI-groep	938,5	8
Bedrijfs grootte	633,2	2
Zelfde baan in Polis	319,6	1
Migratieachtergrond	99,9	6
Leeftijdsklasse	67,5	1
Contracturen	47,3	1
Opleidingsniveau	43,8	3
Kwartaalnummer	28,1	7
Geslacht	24,2	1

Daar hoort de volgende interpretatie bij:

Variabele	Kans op Polis vast, model flex
Softwarecluster	In cluster 2 is de kans hoger en in de clusters 3,4 en 5 is de kans juist (veel) lager
Baanduur	De kans op deze onnauwkeurigheid neemt af bij banen langer dan 36 maanden
SBI-groep	Minder vaak bij Arbeidsbemiddeling, uitzendbureaus en personeelsbeheer, Zorginstellingen, Landbouw, bosbouw en visserij en Vervoer over water en luchtvaart hebben een kleinere kans hierop
Bedrijfs grootte	Vaker bij het midden- en kleinbedrijf dan bij de grote bedrijven
Zelfde baan in Polis	Minder vaak als men van baan verandert
Migratieachtergrond	Vaker bij een Antilliaanse en minder vaak met een Turkse migratieachtergrond
Leeftijdsklasse	niet belangrijk
Contracturen	niet belangrijk
Opleidingsniveau	niet belangrijk
Kwartaalnummer	niet belangrijk
Geslacht	niet belangrijk

De bijdrage van variabelen aan de verklaring van de onnauwkeurigheid EBB-meting is overig en modelschatting is flexibel voor 25- tot 55-jarigen in 2017 en 2018

Variabele	Chi-kwadraat waarde	Aantal vrijheidsgraden
Contracturen	1086,0	1
Zelfde werkgever EBB	472,9	2
Bedrijfsgrootte	214,5	2
Geslacht	150,0	1
SBI-groep	85,1	8
Migratieachtergrond	78,4	6
Baanduur	38,7	1
Leeftijdsklasse	13,7	1
Kwartaalnummer	11,9	7
Opleidingsniveau	11,4	3

Daar hoort de volgende interpretatie bij:

Variabele	Kans op EBB overig, model flex
Contracturen	Voor mensen met een contract van minder dan 30 uur een vier keer zo grote kans
Zelfde werkgever EBB	Een (veel) grotere kans voor werknemers die van werkgever zijn veranderd of van wie dat niet bekend is
Bedrijfsgrootte	Bij kleine bedrijven komt dit veel vaker voor
Geslacht	Een kleinere kans voor vrouwen
SBI-groep	Bij de detailhandel en zorginstellingen komt dit minder vaak voor
Migratieachtergrond	Komt vaker voor bij alle niet-Westerse migratieachtergronden
Baanduur	Komt vaker voor bij werknemers met een baanduur van meer dan 36 maanden
Leeftijdsklasse	niet belangrijk
Kwartaalnummer	niet belangrijk
Opleidingsniveau	niet belangrijk
Contracturen	niet belangrijk

De bijdrage van variabelen aan de verklaring van de onnauwkeurigheid Polis-meting is flexibel en modelschatting is vast voor 25- tot 55-jarigen in 2017 en 2018

Variabele	Chi-kwadraat waarde	Aantal vrijheidsgraden
Softwarecluster	10802,0	4
Bedrijfsgrootte	4388,2	2
Leeftijdsklasse	2156,2	1
SBI-groep	1644,7	8
Baanduur	1227,5	1
Zelfde baan in Polis	408,0	1
Migratieachtergrond	343,7	6
Kwartaalnummer	233,2	7
Opleidingsniveau	170,8	3
Geslacht	13,7	1
Contracturen	9,3	1

Daar hoort de volgende interpretatie bij:

Variabele	Kans op Polis flex, model vast
Softwarecluster	De clusters 3,4 en 5 sterk verhoogde kans in vergelijking met cluster 1
Bedrijfsgrootte	Komt vaker voor bij het midden- en kleinbedrijf dan bij grote bedrijven
Leeftijdsklasse	Vaker bij 25-39 jarigen dan bij 40-54 jarigen
SBI-groep	Vaker bij Detailhandel, Arbeidsbemiddeling, uitzendbureaus en personeelsbeheer, minder vaak bij Financiële instellingen
Baanduur	vaker bij baanduren <36 uur
Zelfde baan in Polis	vaker als men van baan wisselt
Migratieachtergrond	Bij alle niet-westerse migratieachtergronden is de kans op deze onnauwkeurigheid kleiner
Kwartaalnummer	niet belangrijk
Opleidingsniveau	niet belangrijk
Geslacht	niet belangrijk
Contracturen	niet belangrijk

De bijdrage van variabelen aan de verklaring van de onnauwkeurigheid EBB-meting is vast en modelschatting is flexibel voor de 55- tot 75-jarigen in 2017 en 2018

Variabele	Chi-kwadraat waarde	Aantal vrijheidsgraden
SBI-groep	604,5	8
Zelfde werkgever EBB	267,5	2
Bedrijfsgrootte	74,5	2
Contracturen	37,6	1
Kwartaalnummer	33,3	7
Baanduur	32,2	1
Leeftijdsklasse	16,5	1
Migratieachtergrond	10,6	6
Opleidingsniveau	9,5	3
Geslacht	6,1	1

Daar hoort de volgende interpretatie bij:

Variabele	Kans op EBB vast, model flex
SBI-groep	Arbeidsbemiddeling, uitzendbureaus en personeelsbeheer, Vervoer over water en luchtvaart en Onderwijs- en overheidsinstellingen hebben een kleinere kans
Zelfde werkgever EBB	Voor de werknemers die dezelfde werkgever houden hebben een veel grotere kans op deze onnauwkeurigheid.
Bedrijfsgrootte	vaker bij midden- en kleinbedrijf dan bij grootbedrijf
Contracturen	Bij kleinere contracten is de kans op deze onnauwkeurigheid ook kleiner
Kwartaalnummer	Kans neemt schoksgewijs toe tussen eerste kwartaal van 2017 en laatste kwartaal van 2018.
Baanduur	Bij kortere baanduren een kleinere kans op deze onnauwkeurigheid
Leeftijdsklasse	iets minder vaak bij 65 t/m 74 jaar dan bij 55 t/m 64 jaar
Migratieachtergrond	Kans is groter bij werknemers met een Turkse of Antilliaanse en kleiner bij werknemers met een Marokkaanse migratieachtergrond
Opleidingsniveau	niet belangrijk
Geslacht	niet belangrijk

De bijdrage van variabelen aan de verklaring van de onnauwkeurigheid Polis-meting is vast en modelschatting is flexibel voor de 55- tot 75-jarigen in 2017 en 2018.

Variabele	Chi-kwadraat waarde	Aantal vrijheidsgraden
Softwarecluster	2103,5	4
SBI-groep	1671,4	8
Baanduur	1045,2	1
Bedrijfsgrootte	199,3	2
Zelfde baan in Polis	99,5	1
Migratieachtergrond	49,5	6
Kwartaalnummer	28,9	7
Contracturen	13,0	1
Opleidingsniveau	11,7	3
Leeftijdsklasse	2,3	1
Geslacht	2,1	1

Daar hoort de volgende interpretatie bij:

Variabele	Kans op Polis vast, model flex
Softwarecluster	In cluster 2 is de kans groter en in de clusters 3,4 en 5 is de kans juist (veel) lager
SBI-groep	Veel minder vaak bij Arbeidsbemiddeling, uitzendbureaus en personeelsbeheer, en minder vaak bij Horeca en Vervoer over water en luchtvaart
Baanduur	De kans op deze onnauwkeurigheid is kleiner bij banen korter dan 36 maanden
Bedrijfsgrootte	Vaker bij kleine bedrijven dan bij de (middel)grote bedrijven
Zelfde baan in Polis	Minder vaak als men van baan verandert
Migratieachtergrond	Minder vaak bij een Antilliaanse en vaker met een Surinaamse migratieachtergrond
Kwartaalnummer	niet belangrijk
Contracturen	niet belangrijk
Opleidingsniveau	niet belangrijk
Leeftijdsklasse	niet belangrijk
Geslacht	niet belangrijk

De bijdrage van variabelen aan de verklaring van de onnauwkeurigheid EBB-meting is overig en modelschatting is flexibel voor de 55- tot 75-jarigen in 2017 en 2018:

Variabele	Chi-kwadraat waarde	Aantal vrijheidsgraden
SBI-groep	604,5	8
Zelfde werkgever EBB	267,5	2
Bedrijfsgrootte	74,5	2
Contracturen	37,6	1
Kwartaalnummer	33,3	7
Baanduur	32,2	1
Leeftijdsklasse	16,5	1
Migratieachtergrond	10,6	6
Opleidingsniveau	9,5	3
Geslacht	6,1	1

Daar hoort de volgende interpretatie bij:

Variabele	Kans op EBB overig, model flex
SBI-groep	Bij de Detailhandel, Horeca en Zorginstellingen komt dit minder vaak voor, bij Overheids- en onderwijsinstellingen komt dit vaker voor
Zelfde werkgever EBB	Een veel grotere kans voor werknemers die van werkgever zijn veranderd of van wie dat niet bekend is
Bedrijfsgrootte	Bij kleine bedrijven komt dit veel vaker voor
Contracturen	Voor mensen met een contract van minder dan 30 uur een veel grotere kans
Kwartaalnummer	niet belangrijk
Baanduur	Komt minder vaak voor bij werknemers met een baanduur van meer dan 36 maanden
Leeftijdsklasse	Grotere kans bij de 65-74 jarigen
Migratieachtergrond	Komt vaker voor bij werknemers met een Turkse, Surinaamse of ander niet-Westerse minder vaak bij Antilliaanse migratieachtergrond
Opleidingsniveau	Komt vaker voor bij hoger opgeleiden
Geslacht	niet belangrijk

De bijdrage van variabelen aan de verklaring van de onnauwkeurigheid Polis-meting is flexibel en modelschatting is vast voor de 15- tot 25-jarigen in 2017 en 2018:

Variabele	Chi-kwadraat waarde	Aantal vrijheidsgraden
Softwarecluster	2648,4	4
Baanduur	2226,2	1
SBI-groep	1344,8	8
Contracturen	94,7	1
Opleidingsniveau	93,7	3
Geslacht	81,0	1
Zelfde baan in Polis	68,6	1
Kwartaalnummer	55,9	7
Migratieachtergrond	46,4	6
Leeftijdsklasse	26,4	1
Bedrijfsgrootte	19,9	2

Daar hoort de volgende interpretatie bij:

Variabele	Kans op Polis vast, model flex
Softwarecluster	In cluster 2 is de kans iets kleiner en in de clusters 3,4 en 5 is de kans juist veel groter
Baanduur	De kans hierop is groter bij banen korter dan 36 maanden
SBI-groep	Groter bij Arbeidsbemiddeling, uitzendbureaus en personeelsbeheer, Detailhandel en Overheids- en onderwijsinstellingen, kleiner bij Landbouw, bosbouw en visserij
Zelfde baan in Polis	Groter als men van baan verandert
Contracturen	niet belangrijk
Opleidingsniveau	niet belangrijk
Geslacht	niet belangrijk
Kwartaalnummer	niet belangrijk
Migratieachtergrond	niet belangrijk
Leeftijdsklasse	niet belangrijk
Bedrijfsgrootte	niet belangrijk

De bijdrage van variabelen aan de verklaring van de onnauwkeurigheid EBB-meting is flex en modelschatting is vast voor de 15- tot 25-jarigen in 2017 en 2018 is:

Variabele	Chi-kwadraat waarde	Aantal vrijheids-graden
Contracturen	613,2	1
Zelfdewerkgever in EBB	410,2	2
SBI-groep	249,3	8
Kwartaalnummer	144,1	7
Bedrijfsgrootte	74,0	2
Baanduur	55,4	1
Leeftijdsklasse	47,7	1
Opleidingsniveau	7,7	3
Migratieachtergrond	5,6	6
Geslacht	1,4	1

Daar hoort de volgende interpretatie bij:

Variabele	Kans op EBB flex, model vast
Contracturen	Kleiner bij contracten van minder dan 30 uur per week
Zelfde werkgever in EBB	Veel grotere kans als men van werkgever verandert
SBI-groep	Groter bij Arbeidsbemiddeling, uitzendbureaus en personeelsbeheer, Landbouw, bosbouw en visserij en Horeca
Kwartaalnummer	Kans neemt langzamerhand af
Bedrijfsgrootte	Kans is groter bij kleine en middelgrote bedrijven
Baanduur	Kans is groter bij kortere baanduren
Leeftijdsklasse	Kans is groter bij de jongste leeftijdscategorie
Opleidingsniveau	Niet belangrijk
Migratieachtergrond	Niet belangrijk
Geslacht	Niet belangrijk

De bijdrage van variabelen aan de verklaring van de onnauwkeurigheid EBB-meting is vast en modelschatting is flex voor de 15- tot 25-jarigen in 2017 en 2018 is:

Variabele	Chi-kwadraat waarde	Aantal vrijheidsgraden
Zelfde werkgever in de EBB	1058,4	2
SBI-groep	270,5	8
Contracturen	257,5	1
Leeftijdsklasse	47,4	1
Bedrijfsgrootte	45,9	2
Migratieachtergrond	33,2	6
Kwartaalnummer	23,0	7
Opleidingsniveau	11,1	3
Geslacht	2,6	1
Baanduur	1,5	1

Daar hoort de volgende interpretatie bij:

Variabele	Kans op EBB vast, model flex
Zelfde werkgever in de EBB	Kleiner als men van baan wisselt
SBI-groep	Kleiner bij Arbeidsbemiddeling, uitzendbureaus en personeelsbeheer, Landbouw, bosbouw en visserij, groter bij Overheids- en onderwijsinstellingen
Contracturen	Kans is groter bij contracten van groter dan 30 uur
Leeftijdsklasse	Kans is groter bij de 19-24 jarigen
Bedrijfsgrootte	Vaker bij kleine dan bij grotere bedrijven
Migratieachtergrond	Vaker bij herkomst Nederland
Kwartaalnummer	Niet belangrijk
Opleidingsniveau	Niet belangrijk
Geslacht	Niet belangrijk
Baanduur	Niet belangrijk

Verklaring van tekens

Niets (blanco)	Een cijfer kan op logische gronden niet voorkomen
.	Het cijfer is onbekend, onvoldoende betrouwbaar of geheim
*	Voorlopig cijfer
**	Nader voorlopige cijfer
2017–2018	2017 tot en met 2018
2017/2018	Het gemiddelde over de jaren 2017 tot en met 2018
2017/'18	Oogstjaar, boekjaar, schooljaar enz., beginnend in 2017 en eindigend in 2018
2013/'14–2017/'18	Oogstjaar, boekjaar, enz., 2015/'16 tot en met 2017/'18

In geval van afronding kan het voorkomen dat het weergegeven totaal niet overeenstemt met de som van de getallen.

Colofon

Uitgever

Centraal Bureau voor de Statistiek
Henri Faasdreef 312, 2492 JP Den Haag
www.cbs.nl

Prepress

CCN Creatie en visualisatie

Ontwerp

Edenspiekermann

Inlichtingen

Tel. 088 570 70 70, fax 070 337 59 94
Via contactformulier: www.cbs.nl/infoservice

© Centraal Bureau voor de Statistiek, Den Haag/Heerlen/Bonaire, 2021.
Verveelvoudigen is toegestaan, mits het CBS als bron wordt vermeld.