



Haalbaarheidsonderzoek: regionale cijfers over specifieke groepen deeltijders en niet-werkenden

John Michiels

Juli, 2024

CBS Heerlen
CBS-weg 11
6412 EX Heerlen
Postbus 4481
6401 CZ Heerlen
+31 45 570 60 00
www.cbs.nl

projectnummer

PR002992
SAL
28 maart 2024

Inhoudsopgave

1.	Samenvatting en advies	4
2.	Inleiding	6
3.	Data	8
4.	Methoden	10
4.1	GREG-schatters	10
4.2	Kleinedomeinschatters	10
4.3	Methoden voor modelselectie	10
4.4	Methoden voor modevaluatie	11
5.	Resultaten	14
5.1	Resultaten modelselectie	14
5.2	Resultaten modevaluatie	14
5.3	Kwaliteitscriterium voor publicatie van regionale schattingen: de variatiecoëfficiënt	20
6.	Conclusies	22

1. Samenvatting en advies

Het CBS stelt cijfers samen over de binding van personen met de arbeidsmarkt, en meer in het bijzonder over het onbenut arbeidspotentieel onder zowel de werkzame beroepsbevolking als de niet-werkende bevolking. Daarbij wordt onder meer onderscheid gemaakt naar de opleidingsrichting van personen ([SL_bindingarbeidsmarktsonderwijs](#)). Vanuit het programma Arbeidsmarkt Zorg en Welzijn (AZW) is de wens geuit om een haalbaarheidsonderzoek te starten naar de regionale spreiding van een aantal groepen die voor het vervullen van de arbeidsvraag in de sector zorg en welzijn van betekenis zouden kunnen zijn. Het gaat om de volgende doelpopulaties:

1. Jongeren van 17 tot 25 jaar¹ die vanwege opleiding of studie niet kunnen of willen werken, niet direct beschikbaar zijn voor werk en ook niet recent naar werk hebben gezocht.
2. Onderbenutte deeltijdwerkers met opleidingsrichting gezondheidszorg en welzijn.
3. Niet-werkzame personen in de leeftijd van 65 tot 75 jaar die niet direct beschikbaar zijn voor werk en niet recent naar werk hebben gezocht vanwege vut/pensioen/hoge leeftijd met opleidingsrichting gezondheidszorg en welzijn.
4. Niet-werkzame personen in de leeftijd van 65 tot 75 jaar die niet direct beschikbaar zijn voor werk en niet recent naar werk hebben gezocht vanwege vut/pensioen/hoge leeftijd zonder opleidingsrichting gezondheidszorg en welzijn.

De gewenste mate van regionaal detail wordt verkregen door te publiceren naar RegioPlus-arbeidsmarktregio. Een belangrijke beperking hierbij is het aantal waarnemingen per regio in de Enquête beroepsbevolking (EBB). Met traditionele schattingsmethoden is het niet altijd mogelijk om voor de meeste RegioPlus-arbeidsmarktregio's voldoende betrouwbare cijfers op te leveren, maar de verwachting is dat dit met modelgebaseerde technieken (klein domein schatters) wel lukt.

In deze rapportage onderzoeken we de mogelijkheden om met modelgebaseerde technieken cijfers samen te stellen op regionaal niveau over de hiervoor genoemde groepen. De volgende onderdelen van deze methode zijn nader onderzocht:

1. De keuze van het beste model voor het schatten van de doelvariabelen (i.c. de omvang van de vier doelpopulaties), gegeven de beschikbare hulpinformatie.
2. De toetsing van de gevonden schatters.
3. Nagaan of deze schatters, na een positief toetsingsresultaat, ook voldoende nauwkeurige uitkomsten opleveren.

De schattingen, en de selectie en toetsing van de modellen, zijn gebaseerd op de Enquête beroepsbevolking (EBB) van de verslagjaren 2021-2023. Hierbij is ook gebruikgemaakt van registerbestanden. Deze registerbestanden bevatten demografische gegevens van personen, UWV-inschrijvingen, informatie over de sociaaleconomische categorie, het onderwijsniveau en de onderwijsrichting.

De conclusies zijn als volgt:

- Voor de vier doelpopulaties zijn over de verslagjaren 2022 en 2023 voor alle RegioPlus-arbeidsmarktregio's voldoende nauwkeurige jaarschattingen te realiseren.

¹ Aanduidingen zoals 17 tot 25 jaar (of 17 – 24 jaar) hebben betrekking op leeftijden vanaf 17 jaar (17 jaar is inbegrepen) tot en met 24 jaar (24 jaar is inbegrepen, maar 25 jaar is niet inbegrepen).

- Voor doelvariabele 'Onderbenutte deeltijdwerkers met opleidingsrichting gezondheidszorg en welzijn' zijn voor verslagjaar 2021 de kleinedomeinschattingen onvoldoende nauwkeurig. Dit hangt samen met de revisie van de EBB in dat verslagjaar².

² In 2021 was het steekproefpanel van de gereviseerde EBB nog in opbouw, waardoor de bruikbare respons (veel) lager uitviel.

2. Inleiding

Met de Enquête beroepsbevolking (EBB) worden cijfers samengesteld over de relatie tussen mens en arbeidsmarkt voor de groep personen woonachtig in particuliere huishoudens in Nederland. Het CBS presenteert deze cijfers doorgaans voor personen in de leeftijd van 15 tot 75 jaar. Deze cijfers bieden onder meer inzicht in de binding van personen met de arbeidsmarkt en de relatie met andere kenmerken: bijvoorbeeld geslacht, leeftijd en onderwijsniveau- en richting. Met het kenmerk binding met de arbeidsmarkt kunnen doelpopulaties worden onderscheiden die voor het *programma Arbeidsmarkt Zorg en Welzijn* gezien de arbeidsvraag in de sector zorg en welzijn van belang kunnen zijn. Niet alleen zijn landelijke gegevens gewenst, maar ook cijfers op het niveau van RegioPlus-arbeidsmarktregio. Deze regionale uitsplitsing maakt duidelijk in welke regio's nog potentieel arbeidsaanbod te vinden is.

Om cijfers op regionaal niveau te kunnen samenstellen zijn methoden nodig die deze schattingen voldoende nauwkeurig kunnen leveren. De traditionele methode, met het wegen van waarnemingen door de algemene regressieschatter³, schiet bij steekproefonderzoek veelal tekort, omdat deze methode bij het maken van schattingen per regio alleen de waarnemingen van het specifieke gebied betreft. Bij een gering aantal waarnemingen leidt dat tot grote steekproeffluctuaties in de uitkomsten. Een mogelijke oplossing bestaat erin door gebruik te maken van modelgebaseerde technieken: de kleinedomeinschatters. In geval van een groot aantal waarnemingen per domein lijken deze modelgebaseerde schatters op de hiervoor genoemde algemene regressieschatter. Bij een gering aantal waarnemingen krijgen de modelgebaseerde schatters een meer 'synthetisch karakter' en worden de uitkomsten bepaald door de algehele relatie tussen doelvariabelen en hulpvariabelen. Om in dit geval tot nauwkeurige schattingen te kunnen komen is het van belang dat de hulpinformatie sterk samenhangt met de doelvariabelen.

De vraag die in deze nota centraal staat is of, gegeven de beschikbare hulpinformatie en regio-indeling, de doelvariabelen met voldoende nauwkeurigheid door kleinedomeinschatters kunnen worden geschat op het niveau van RegioPlus-arbeidsmarktregio's. We nemen bij de berekening en evaluatie van de KDS-schattingen de volgende doelpopulaties van de EBB onder de loep:

1. Jongeren van 17 tot 25 jaar die vanwege opleiding of studie niet kunnen of willen werken, niet direct beschikbaar zijn voor werk en ook niet recent naar werk hebben gezocht.
2. Onderbenutte deeltijdwerkers met opleidingsrichting gezondheidszorg en welzijn.
3. Niet-werkzame personen in de leeftijd van 65 tot 75 jaar die niet direct beschikbaar zijn voor werk en niet recent naar werk hebben gezocht vanwege vut/pensioen/hoge leeftijd met opleidingsrichting gezondheidszorg en welzijn.
4. Niet-werkzame personen in de leeftijd van 65 tot 75 jaar die niet direct beschikbaar zijn voor werk en niet recent naar werk hebben gezocht vanwege vut/pensioen/hoge leeftijd zonder opleidingsrichting gezondheidszorg en welzijn.

Voor de hulpinformatie wordt onder meer gebruikgemaakt van de volgende gegevensbronnen:

- Demografische kenmerken van personen en huishoudens uit de bevolkingsregistraties;
- Inschrijvingskenmerken van personen die zijn ingeschreven bij het UWV (registerbestanden uit de EBB-productie);

³ Een synoniem voor algemene regressieschatter is GREG-schatter, wat staat voor **G**eneral **r**egression estimator.

- Loon- en baankenmerken van werknemers uit de Polis-administratie (SPOLISJJJJBUSVV en SPOLISHOOFDBAANJJJJBUSVV);
- Informatie over zelfstandigen, studerende, en ontvangers van een pensioenuitkering uit de registercomponent SECMBUSVV;
- Opleidingsniveau en –richting van personen uit registerbestand HOOGSTEOPLJJJJTABVV;

De inhoudelijke behandeling van het haalbaarheidsonderzoek begint in hoofdstuk 3 met een beschrijving van de gebruikte bronnen. In hoofdstuk 4 wordt aandacht besteed aan de methoden die in dit onderzoek zijn toegepast voor het schatten, de modelselectie en de modeltoetsing. Het zwaartepunt van de beschrijving ligt hierbij op de toetsing van KDS-schattingmethoden, omdat de schatters niet per se onvertkend zijn en de mate van vertekening in kaart moet worden gebracht. In hoofdstuk 5 komen de resultaten aan bod: er wordt een overzicht gegeven van de modelselectieresultaten voor ieder van de vier doelvariabelen en de uitkomsten van de modeltoetsingsmethoden worden beschreven. In het afsluitende hoofdstuk 6 leggen we de conclusies voor ten aanzien van de haalbaarheid van regionale publicaties over de gekozen doelvariabelen.⁴ Aanvullend zijn in dit document een literatuurlijst, een appendix, en een maatwerktablet opgenomen met regionale jaarcijfers (2023) voor de vier doelpopulaties.

⁴ Voor het woord ‘klein domeinschattingen’ wordt in deze nota ook vaak de afkorting KDS-schattingen gebruikt. Voor GREG-schattingen wordt in de tekst op diverse plaatsen ook de term directe schattingen gebruikt; GREG-schattingen zijn een type directe schattingen waarbij van hulpinformatie gebruik wordt gemaakt.

3. Data

Het haalbaarheidsonderzoek is uitgevoerd op de analysebestanden van de Enquête beroepsbevolking (EBB) voor de verslagjaren 2021, 2022 en 2023. Er is ook gebruikgemaakt van hulpinformatie uit registraties voor het berekenen van de diverse schattingen. De registerbestanden die voor dit doel zijn ingezet zijn bestanden die uiterlijk in het eerste kwartaal na het afsluiten van het verslagjaar beschikbaar komen. Dit betekent bijvoorbeeld dat voor de schattingen van verslagjaar 2023 gebruik is gemaakt van de demografische gegevens en UWV-inschrijfgegevens van dat jaar. Deze informatie is vervolgens gekoppeld met de registergegevens van de overige kenmerken ultimo het jaar voorafgaand aan het verslagjaar (31 december 2022). Hieronder is een beschrijving gegeven van de diverse bronnen in het haalbaarheidsonderzoek.

Enquête beroepsbevolking (EBB)

De Enquête beroepsbevolking (EBB) is een representatief onderzoek onder personen in Nederland. De EBB bevat onder andere informatie over de binding van werkzame en niet-werkzame personen met de arbeidsmarkt, het aantal gewerkte uren en of ze onderwijs volgen. De Enquête beroepsbevolking (EBB) is een zogenaamd roterend panelonderzoek met vijf peilingen. Sinds het vierde kwartaal van 2012 worden alle respondenten in eerste instantie gevraagd om via internet te responderen (CAWI). Een deel van de respondenten die niet responderen, wordt vervolgens face-to-face (CAPI) of telefonisch (CATI) benaderd. De vier vervolgbenederingen in het panel vinden via CAWI plaats, waarbij niet-responderende personen via CATI worden herbenaderd.

Demografische gegevens en inschrijvingen bij het UWV

De EBB-productie gebruikt een aantal registerbestanden die voor de weegprocedure van het onderzoek worden toegepast. Dit haalbaarheidsonderzoek neemt een drietal registerbestanden over: een bestand met demografische gegevens over personen (geslacht, leeftijd), een bestand met gegevens over het huishoudentype van personen die tot dat huishouden behoren, en een derde bestand met UWV-inschrijfgegevens van personen. Wat dat laatste bestand betreft: personen zijn wel of niet ingeschreven, en ook is het type inschrijving vastgelegd: of er bij inschrijving sprake is van een WW-uitkering of dat er een curriculum vitae in het dossier is opgenomen.

Andere bronnen voor hulpinformatie

De overige registerbestanden in dit haalbaarheidsonderzoek zijn afkomstig uit het Stelsel van Sociaal-statistische Bestanden (SSB). Het SSB is een uitgebreid stelsel van koppelbare registers en enquêtes, die in de vorm van zgn. SSB-componenten beschikbaar zijn. Met gegevens uit het SSB worden over uiteenlopende onderwerpen statistieken gemaakt en wordt sociaalwetenschappelijk onderzoek uitgevoerd.

Bestanden baankenmerken werknemers (SPOLISJJJBUSVV en SPOLISHOOFDBAANJJJBUSVV)

De SSB-component SPOLISJJJBUSVV bevat kwantitatieve en kwalitatieve gegevens over banen en lonen van werknemers bij Nederlandse bedrijven. De definitie van "baan" die ten grondslag ligt aan dit bestand, is de zogenaamde inkomstenverhouding (IKVID): de administratieve registratie van de arbeidsovereenkomst tussen werkgever en werknemer. Een werknemer kan per werkgever meerdere inkomstenverhoudingen tegelijkertijd hebben. Uit dit bestand zijn gegevens overgenomen over soort contract (bepaalde of onbepaalde tijd) en soort baan (uitzendkracht, oproepbaan, overig). Ook kan de hoofdbaan worden geïdentificeerd (via de component SPOLISHOOFDBAANJJJBUSVV).

Bestanden sociaaleconomische categorie (SECMBUSVV)

De SSB-component SECMBUSVV bevat onder meer gegevens over de sociaaleconomische categorie (SECM) van personen in een bepaalde maand. Van de afzonderlijke inkomensbronnen, die aan de afbakening van SECM ten grondslag liggen, wordt aangegeven of een persoon deze in de betreffende periode heeft. Ook wordt in de component aangegeven of een persoon als scholier of student in de betreffende maand stond ingeschreven. Dit laatste gegeven is gebaseerd op de onderwijsregistraties (SSB-component STUDERENDENJJJBUSVV).

Bestand hoogst behaald/gevolgd opleidingsniveau, en opleidingsrichting (HOOGSTEOPLJJJTABVV)

De SBB-component HOOGSTEOPLJJJTABVV stelt gegevens beschikbaar over het hoogst behaalde en hoogst gevolgde opleidingsniveau (en -richting) van de Nederlandse bevolking op 1 oktober van het verslagjaar. Het bestand is gebaseerd op gegevens uit diverse registers en de Enquête beroepsbevolking (EBB). Door het gebruik van meerdere bronnen heeft HOOGSTEOPLJJJTABVV een hoge dekkingsgraad (2018: ruim 11,5 miljoen personen). Desondanks vertegenwoordigt het bestand (nog) niet de gehele doelpopulatie. De informatie is voor een deel van de bevolking integraal en voor de overige personen op steekproefbasis (EBB). Het bestand bevat daarom een ophooggewicht (GEWICHTHOOGSTEOPL).

4. Methoden

4.1 GREG-schatters

De algemene regressieschatter (ook GREG-schatter genoemd) behoort tot het standaardarsenaal van schattingsmethodes voor officiële statistieken. De schatter maakt gebruik van modellen om de precisie van de schattingen te verbeteren en de effecten van selectiviteit door non-respons (het niet deelnemen van geselecteerde respondenten) te verminderen. In tegenstelling tot de modelgebaseerde schattingen wordt het kansmechanisme van deze schatters niet bepaald door het model (maar door het steekproefontwerp) en ze heten daarom ook wel model-assisted schatters (Särndal, Swensson en Wretman, 2003).

Een belangrijk kenmerk van GREG-schatters is dat voor het maken van schattingen in een klein regionaal gebied alleen gebruik wordt gemaakt van de waarnemingen in dat gebied. Vanwege steekproeffluctuaties kunnen de uitkomsten van zo'n schatter daarom onnauwkeurig zijn. GREG-schatters zijn vanwege die reden alleen goed bruikbaar voor regionale gebieden van enige omvang of als de enquête een zodanige omvang heeft dat er voldoende waarnemingen op regionaal niveau beschikbaar zijn.

In dit onderzoek zijn twee typen GREG-schatters toegepast:

- 1) De GREG-schatter zoals gebruikt in de EBB-weging, met een algemeen weegmodel.
- 2) De GREG-schatter op basis van het geselecteerde KDS-model. Dit model is voor iedere doelvariabele (over het algemeen) verschillend.

Voor meer informatie over het gebruik van GREG-schatters voor regionale toepassingen zie Särndal, Swensson en Wretman (2003) of Rao, J. (2003).

4.2 Kleinedomeinschatters

Kleinedomeinschatters compenseren voor een gering aantal waarnemingen in kleine regionale eenheden (ook wel domeinen genoemd) door gebruik te maken van de relatie tussen doelvariabelen en achtergrondkenmerken in de gehele steekproef. Op basis van deze relatie worden voor de niet-waargenomen eenheden in het kleine regionale gebied waarden geïmputeerd. De schatters zijn hierdoor minder gevoelig voor steekproeffluctuaties. De berekening van de KDS-schattingen vindt plaats met door het CBS ontwikkelde software in de statistische programmeertaal R (Boonstra, 2012). Voor meer informatie over KDS-schattingsmethoden zie Rao, J. (2003).

4.3 Methoden voor modelselectie

Kleinedomeinschattingen worden gebaseerd op een expliciet model. Voor een zo goed mogelijk model is het belangrijk dat er variabelen in het model worden opgenomen die deel uitmaken van het steekproefdesign, en variabelen die corrigeren voor selectieve respons en een sterke samenhang met de doelvariabelen hebben. Deze verzameling variabelen kan zodoende per doelvariabele verschillen.

Voor het selecteren van de modellen wordt gebruikgemaakt van een aantal evaluatiematen. Dit zijn grootheden die onder meer iets zeggen over hoe goed het model bij de data past. In dit onderzoek wordt van de volgende grootheden gebruikgemaakt:

- cAIC, meet modelfit en complexiteit van het model (mate van overfitting)
- standaardfouten van KDS-schattingen (indicatie voor precisie);

- tussendomeinvarianties KDS-schattingen (geeft informatie over het gewicht van de directe schatter binnen de KDS-schatter);

Voor meer informatie over de cAIC-fitmaat wordt verwezen naar Vaida en Blanchard (2005).

4.4 Methoden voor modevaluatie

Waarom modevaluatie?

Een belangrijke eigenschap van GREG-schatters is dat ze bij benadering design-zuiver zijn, mits de steekproefrespons niet sterk vertekend is. Dit betekent dat als de enquête vele malen herhaald zou worden, met iedere keer een nieuwe steekproef, het gemiddelde van de uitkomsten voor een bepaalde doelvariabele de overeenkomstige (onbekende) populatiewaarde zou moeten benaderen. Bij kleinedomeinschatters is deze eigenschap losgelaten. Daar staat tegenover dat KDS-schattingen over het algemeen wel veel preciezer (d.w.z. minder variabel) zijn dan GREG-schattingen. De vraag is of de hogere precisie van de KDS-schattingen opweegt tegen een mogelijke vertekening van deze schattingen. De nauwkeurigheid⁵ van KDS-schattingen moet om die reden worden getoetst. Hiervoor zijn diverse methoden beschikbaar.

Afwijkingen ten opzichte van directe schattingen⁶

Modelschattingen kunnen sterk vertekend zijn als het model niet goed gekozen is. Een vergelijking met de directe schattingen is een eerste stap om dit in beeld te krijgen. Visueel worden scatterplots van regionale puntschattingen van GREG- en KDS-schatters met elkaar vergeleken. Daarnaast wordt voor alle domeinen de afstand van de GREG-schatter tot de kleinedomeinschatter berekend, uitgedrukt in standaardfouten van de GREG-schatter (per domein j):

$$d_j = \left[\frac{y_j^{KDS} - y_j^{GREG}}{SE(y_j^{GREG})} \right],$$

De modelschattingen zijn plausibel als deze niet te sterk afwijken van de directe schattingen, gegeven de betrouwbaarheidsmarges van de directe schattingen. Het gaat hier niet om formele toetsen. Deze zouden ingewikkeld zijn om uit te voeren omdat de directe en modelschattingen gebaseerd zijn op dezelfde steekproef, en dus niet onafhankelijk zijn. In het geval van nauwkeurige modelschattingen die de populatieparameters dicht benaderen, zouden deze schattingen niet meer dan in ongeveer 5 procent van de gevallen meer dan twee standaardfouten van de directe schatting mogen afwijken.

Berekenen van geclusterde en landelijke totalen uit regionale schattingen

Als kleinedomeinschatters nauwkeurige schattingen opleveren dan zouden ook de geclusterde schattingen nauwkeurig moeten zijn. Deze geclusterde schattingen kunnen worden vergeleken met de resultaten van GREG-schattingen en zouden ermee overeen moeten stemmen. Een 'extreem' voorbeeld is het clusteren van regionale schattingen tot landelijke cijfers. Omdat RegioPlus-arbeidsmarktregio's al een redelijke omvang hebben, wordt er geclusterd tot op nationaal niveau.

Berekenen van standaardfouten

Omdat kleinedomeinschatters hun 'nauwkeurigheidswinst' vooral behalen met kleinere standaardfouten (= geringere variabiliteit) voor de schattingen, is het zaak na te gaan of de

⁵ Een nauwkeurige schatting is zowel precies (weinig variabel) als niet of slechts in geringe mate vertekend.

⁶ Bewerkte beschrijving van de tekst door Krieg en Burger (januari 2019)

standaardfouten van dergelijke schatters ook daadwerkelijk kleiner zijn dan die van hun GREG-tegenhangers.

Vergelijken van reeksen van KDS- en GREG schattingen

Een belangrijke stap in de toetsing van KDS-schattingen bestaat uit de vergelijking van dit type schattingen met een aantal benchmarks. Eén variant hiervan is in deze paragraaf al besproken, namelijk het vergelijken van geclusterde KDS-schattingen met overeenkomstige clusters van directe schattingen. Een andere benchmarkmethode bestaat uit het vergelijken van reeksen van KDS-schattingen over de tijd met overeenkomstige reeksen van directe schattingen (zelfde doelvariabele, zelfde domein). Deze methode is eerder toegepast voor het toetsen van regionale cijfers over arbeidspositie (Boonstra en Michiels, 2013). De reden dat voor deze benchmark wordt gekozen, is dat directe schatters, afgezien van non-respons, consistent en bij benadering design-zuiver zijn. Een KDS-reeks kan als goed worden beoordeeld als deze qua gemiddeld niveau de overeenkomstige reeks van directe schattingen volgt, ook al is de variatie in de directe schattingen bij kleinere domeinen doorgaans veel groter dan bij de KDS-schattingen.

Het aantal reeksen in de benchmark wordt bepaald door het aantal doelvariabelen (= aantal doelpopulaties) en het aantal domeinen (=aantal RegioPlus-arbeidsmarktregio's). Dit zijn 4 doelvariabelen⁷ x 28 RegioPlus-arbeidsmarktregio's = 112 reeksen. Gezien het grote aantal reeksen is ervoor gekozen om een samenvattende maat voor de verschillen tussen KDS-reeksen en GREG-reeksen samen te stellen: per doelvariabele en domein worden de reeksgemiddelden berekend van KDS-schattingen en directe schattingen. Het verschil van deze gemiddelden wordt geschaald met de standaardfout van het reeksgemiddelde van de directe schattingen. De op deze wijze berekende geschaalde verschillen zouden, net als eerder beschreven onder het kopje *Afwijkingen ten opzichte van directe schattingen*, merendeels niet meer dan 2 standaardfouten van 0 mogen afwijken. De figuren met alle reeksen zijn overigens wel in de appendix opgenomen.

Bij de berekening van de standaardfout van het reeksgemiddelde van de directe schattingen moet rekening worden gehouden met de correlatie tussen opeenvolgende schattingen in de tijd. In het voorliggende onderzoek is gebruikgemaakt van reeksen van 12 opeenvolgende kwartaalschattingen in de periode 2021-2023 ($m=12$).

Voor het reeksgemiddelde van de directe schattingen $\hat{t} = \frac{1}{m} \sum_{k=1}^m \hat{t}_k$ geldt de variantieformule:

$$(1) \quad V(\hat{t}) = \left(\frac{1}{m^2}\right) \sum_{k,l=1}^m Cov(\hat{t}_k, \hat{t}_l) = \left(\frac{1}{m^2}\right) \sum_{k=1}^m Var(\hat{t}_k) + \left(\frac{1}{m^2}\right) \sum_{k \neq l=1}^m Cov(\hat{t}_k, \hat{t}_l)$$

De covariantie van de kwartaalschattingen, $Cov(\hat{t}_k, \hat{t}_l)$, kan worden uitgedrukt in termen van de varianties van de afzonderlijke kwartaalschattingen en de omvang van de steekproeven in de betreffende kwartalen, n_k en n_l . Ook relevant in deze berekening zijn de mate van overlap van de steekproeven (n_{kl}) en de correlatie ($\rho_{k,l}$) tussen waarnemingen van de doelvariabele in kwartalen k en l (Meijers, Burger en Boonstra, 2011). De (gepoolde) correlatiecoëfficiënten worden hierbij berekend uit de overlappende steekproeven van kwartalen k en l .

$$(2) \quad Cov(\hat{t}_k, \hat{t}_l) = \frac{n_{kl}}{\sqrt{n_k n_l}} \cdot \rho_{k,l} \cdot \sqrt{Var(\hat{t}_k) Var(\hat{t}_l)}$$

⁷ Niet-werkzame personen in de leeftijd van 65 tot 75 jaar die niet willen of kunnen werken vanwege hoge leeftijd worden in twee groepen onderverdeeld: onderwijsrichting van de hoogst behaalde opleiding is in gezondheidszorg en welzijn, of daarbuiten.

Omdat de varianties en steekproefaantallen over de kwartalen niet veel wijzigen⁸, maken we de volgende benaderingen: $n_k \cong n_l$ en $Var(\hat{t}_k) \cong Var(\hat{t}_l) \cong \frac{1}{m} \sum_{k=1}^m Var(\hat{t}_l) \equiv Var(\hat{t}_l)$

Dan geldt in goede benadering: (3)

$$V(\hat{t}) = \frac{1}{m} \overline{Var(\hat{t}_l)} + \left(\frac{2}{m^2} \cdot \left[\sum_{l>k=1}^m f_{kl} \cdot \rho_{k,l} \right] \right) \cdot \overline{Var(\hat{t}_l)}$$

met $f_{kl} = \frac{n_{kl}}{n_k}$. De standaardfout van het reeksgemiddelde volgt dan uit $SE(\hat{t}) = \sqrt{V(\hat{t})}$. Voor de overlapfracties f_{kl} worden de volgende waarden gekozen (dit is eveneens een benadering, ingegeven door het roterend paneldesign van de Enquête beroepsbevolking).

f_{kl}	$l=1$	2	3	4	5	6	7	8	9	10	11	12
$k=1$	1	4/5	3/5	2/5	1/5	0	0	0	0	0	0	0
2		1	4/5	3/5	2/5	1/5	0	0	0	0	0	0
3			1	4/5	3/5	2/5	1/5	0	0	0	0	0
4				1	4/5	3/5	2/5	1/5	0	0	0	0
5					1	4/5	3/5	2/5	1/5	0	0	0
6						1	4/5	3/5	2/5	1/5	0	0
7							1	4/5	3/5	2/5	1/5	0
8								1	4/5	3/5	2/5	1/5
9									1	4/5	3/5	2/5
10										1	4/5	3/5
11											1	4/5
12												1

De correlatiecoëfficiënten zijn voor alle doelpopulaties berekend.

⁸ De kwartaalsteekproeven van verslagjaar 2021 zijn kleiner dan die van 2022 en 2023. De steekproefvarianties zijn voor deze kwartalen navenant groter. Er wordt hier aangenomen dat, gezien het beperkte aantal kleine kwartaalsteekproeven, de variantieformules nog steeds voldoen.

5. Resultaten

5.1 Resultaten modelselectie

In deze paragraaf vatten we de resultaten van de modelselectie samen. De eerste kolom van tabel 1 toont de namen van de doelvariabelen. De tweede kolom bevat de uitkomsten van de modelselectieprocedures (bij iedere variabele staat tussen haakjes het aantal categorieën). Voor meer informatie over de hulpvariabelen, zie [appendix A](#).

Tabel 1. Identificatie modellen en hulpvariabelen vaste effecten in het model (2023 Q1)

Doelvariabele	Hulpinformatie (vaste effecten, zie voor codering de technische toelichting)
Jongeren in de niet-beroepsbevolking, 17-24 jaar, die niet willen/kunnen werken vanwege opleiding	Gwu-inschrijving(7) + Onderwijsniveau(6) + Inkomen werknemer(6) + Indicatie zelfstandige(2) + Indicatie studerend(2)
Deeltijdwerkers met opleidingsrichting gezondheidszorg en welzijn	Geslacht(2) + Onderwijsrichting(13) + Contractsoort(4) + Arbeidsduur(6) + Inkomen werknemer(6) + Indicatie zelfstandige(2)
Niet-beroepsbevolking, 65-74 jr, die niet willen/kunnen werken vanwege hoge leeftijd, met opleidingsrichting gezondheidszorg en welzijn	Geslacht (2) + Opleidingsniveau(6) + Opleidingsrichting(13) + Contractsoort(4) + Indicatie zelfstandige(2) + Indicatie ziekte/arbeidsongeschiktheid(2) De doelpopulatie wordt gemeten ten aanzien van de populatie 65-74 jaar (het hier gepresenteerde model dient te worden uitgebreid met de interactie met Leeftijd(6) als de totale populatie van 15-74 jaar als uitgangspunt wordt gekozen)
Niet-beroepsbevolking, 65-74 jr, die niet willen/kunnen werken vanwege hoge leeftijd, andere of geen opleidingsrichting	Geslacht(2) + Leeftijd(24) + Type huishouden(7) + Gwu-inschrijving(7) + Opleidingsniveau(6) + Opleidingsrichting(13) + Contractsoort(4) + Indicatie zelfstandige(2) + Indicatie ziekte/arbeidsongeschiktheid(2) + Indicatie pensioen(2) De doelpopulatie wordt gemeten ten aanzien van de populatie 65-74 jaar (het hier gepresenteerde model dient te worden uitgebreid met de interactie met Leeftijd(6) als de totale populatie van 15-74 jaar als uitgangspunt wordt gekozen)

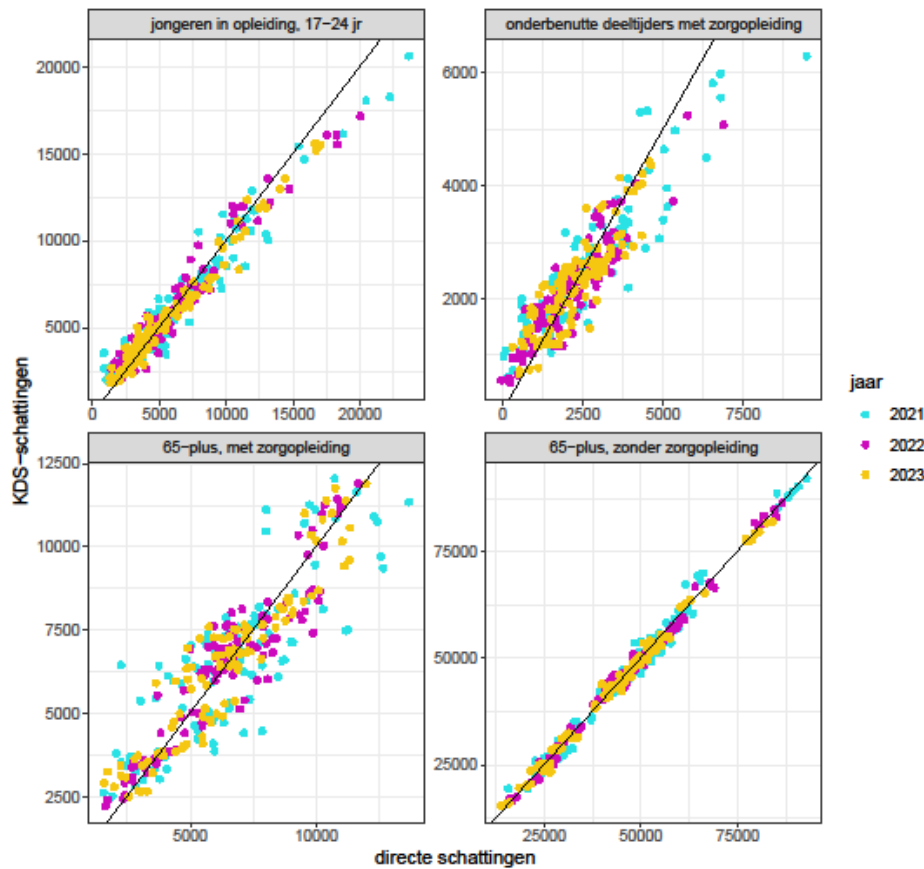
5.2 Resultaten modevaluatie

5.2.1 Afwijkingen kleinedomeinschattingen ten opzichte van directe schattingen

In figuur 1 is een illustratie opgenomen van KDS- en GREG-schattingen voor de vier doelvariabelen. Ieder punt in de figuur vertegenwoordigt een tweetal schattingen voor de doelvariabele in één bepaalde RegioPlus-arbeidsmarktregio (x-as: directe schatting, y-as: KDS-schatting). De diagonaal is

bedoeld als optische ondersteuning: punten die op de diagonaal liggen hebben identieke waarden voor de KDS- en directe schatting.

Figuur 1 Kleinedomein- vs. directe schattingen per RegioPlus-arbeidsmarktregio (2021-2023)



Uit de figuur valt op te maken dat de horizontale afwijkingen van de punten ten opzichte van de diagonaal over het algemeen groter zijn voor verslagjaar 2021. Dit duidt op een grotere variabiliteit van de (directe) schattingen in dat jaar: vanwege het feit dat direct na de start van de herontworpen EBB in 2021 het panel nog niet volledig is opgebouwd zijn de kwartaalsteekproeven kleiner (en de varianties navenant groter).

De eerste stap in de evaluatie van de KDS-schattingen bestaat erin na te gaan of de KDS-schattingen opvallend afwijken van de overeenkomstige GREG-schattingen. Een indicator hiervoor is het geschaalde verschil tussen KDS- en GREG-schattingen zoals beschreven in methodenparagraaf 4.4. Een populatieparameter valt doorgaans binnen 2 standaardfouten van de GREG-schatting (het betrouwbaarheidsinterval van de GREG-schatting). Ligt de preciezer geachte KDS-schatting buiten dat interval dan is dat een opvallende afwijking; de KDS-schatting ligt dan mogelijk ver verwijderd van de populatieparameter.

In tabel 2 is een overzicht gegeven van het percentage 'afwijkende' KDS-schattingen. Percentages van meer dan 4,6 procent (bij > 2 standaardfouten) zijn als verdacht te beschouwen. Bij geen van de doelvariabelen is hiervan sprake.

Tabel 2 Percentage afwijkende KDS-schattingen over het geheel van kleinedomeinschattingen met meer dan 1 resp. 2 standaardfouten (se) verschil t.o.v. overeenkomstige GREG-schattingen (2022-2023).

Doelvariabele	> 1 se	> 2 se
	%	%
Jongeren in de niet-beroepsbevolking, 17-24 jaar, die niet willen/kunnen werken vanwege opleiding	16,1	0,5
Onderbenutte deeltijdwerkers met opleidingsrichting gezondheidszorg en welzijn	18,3	1,3
Niet-beroepsbevolking, 65-74 jr, die niet willen/kunnen werken vanwege hoge leeftijd, met opleidingsrichting gezondheidszorg en welzijn	17,9	0,0
Niet-beroepsbevolking, 65-74 jr, die niet willen/kunnen werken vanwege hoge leeftijd, andere of geen opleidingsrichting	19,2	1,3

5.2.2 Clustering van schattingen

Na clustering van de KDS-schattingen zouden deze overeen moeten komen met GREG-schattingen op hetzelfde geaggregeerde niveau. Dat is in deze studie gerealiseerd door de schattingen te aggregeren naar het nationale niveau (tabel 3). De geaggregeerde GREG- en KDS-schattingen zijn over het algemeen goed vergelijkbaar. Bij de doelvariabele 'onderbenutte deeltijdwerkers met opleidingsrichting gezondheidszorg en welzijn' is sprake van een statistisch significant verschil (weliswaar op de grens van significantie). Dit verschil wordt mede veroorzaakt door het gebruik van verschillende modellen voor enerzijds de EBB-weging en anderzijds de KDS-schattingen.

Tabel 3 Gewogen schattingen en geaggregeerde KDS-schattingen op nationaal niveau, 2023 (Q1)

Doelvariabele	GREG-schatting	Geaggregeerde KDS-schattingen	Significant verschil ⁹ (Ja/Nee)
	%	%	
Jongeren in de niet-beroepsbevolking, 17-24 jaar, die niet willen/kunnen werken vanwege opleiding; Percenteerbasis: bevolking, 15-74 jaar	1,35	1,36	Nee
Onderbenutte deeltijdwerkers met opleidingsrichting gezondheidszorg en welzijn Percenteerbasis: bevolking, 15-74 jaar	0,59	0,52	Ja, net significant
Niet-beroepsbevolking, 65-74 jr, die niet willen/kunnen werken vanwege hoge leeftijd, met opleidingsrichting gezondheidszorg en welzijn Percenteerbasis: Bevolking, 65-74 jaar	8,56	9,16	Nee
Niet-beroepsbevolking, 65-74 jr, die niet willen/kunnen werken vanwege hoge leeftijd, andere opleidingsrichting Percenteerbasis: Bevolking, 65-74 jaar	63,44	64,32	Nee

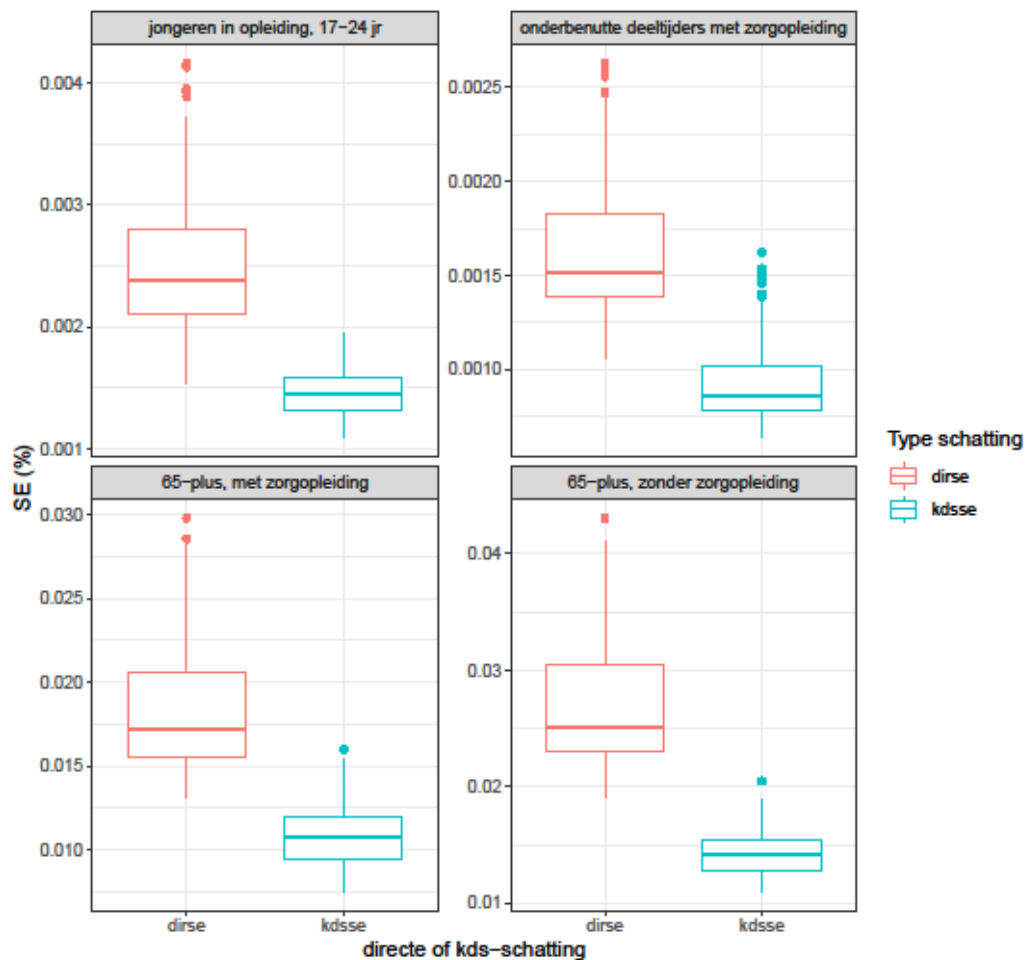
5.2.3 Standaardfouten kleindomeinschattingen

Wat de precisie van de schattingen betreft, presteren de KDS-schatters duidelijk beter dan de GREG-schatters voor overeenkomstige doelvariabelen (figuur 2). De mediane waarde van de geschatte

⁹ Betekent hier: KDS-schattingen vallen niet binnen het betrouwbaarheidsinterval van de GREG-schattingen, berekend met een variantieschatter voor een enkelvoudig aslect design zonder teruglegging.

standaardfouten (SE-schattingen) is bij KDS-schattingen (kdsse) ongeveer 2 maal kleiner dan bij de GREG-schattingen (dirse). Ook is de verdeling van de standaardfouten bij KDS-schattingen compacter dan bij de GREG-schattingen¹⁰. De ‘brede’ verdeling van de GREG SE-schattingen is voor een deel toe te schrijven aan de onnauwkeurigheid van de geschatte standaardfouten; deze schattingen zijn evenals de GREG-schattingen voor de doelvariabelen alleen gebaseerd op de waarnemingen binnen een bepaald domein.

Figuur 2 Vergelijking standaardfouten naar type schattingen, 2022-2023

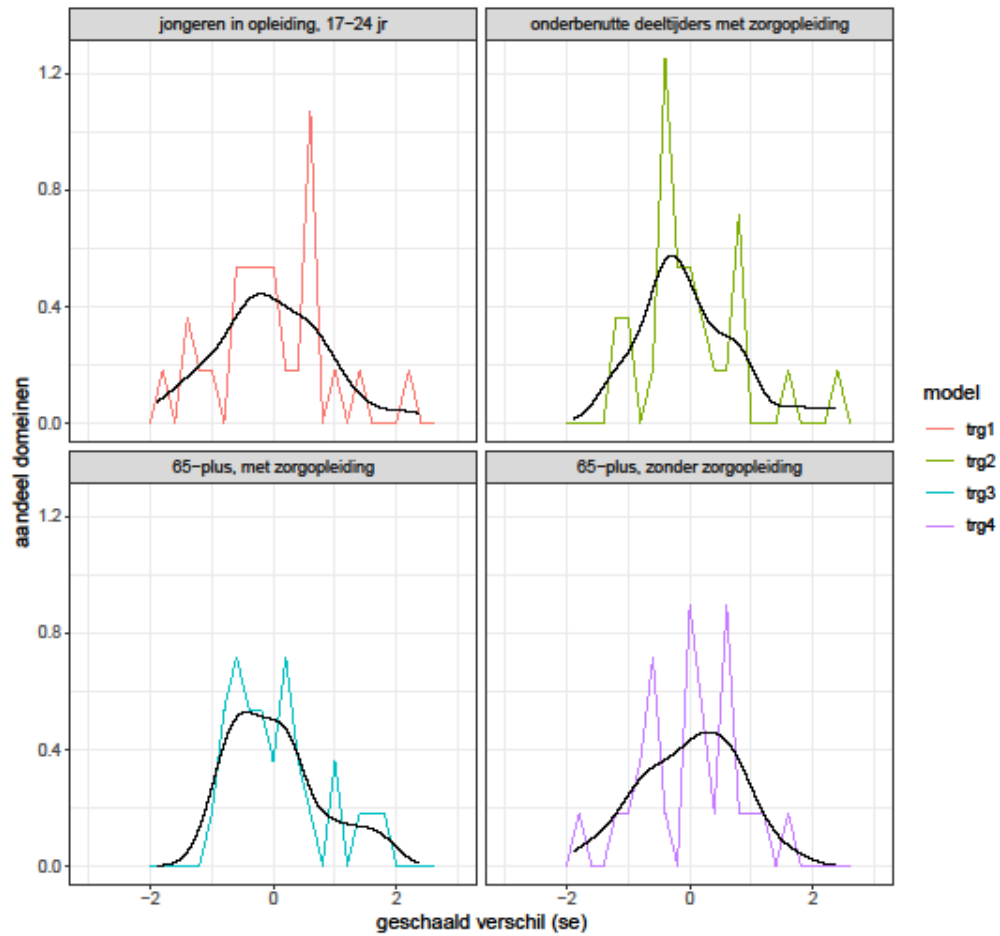


5.2.4 Vergelijking van reeksen KDS- en GREG-schattingen

De vergelijking van reeksgemiddelden van KDS- en GREG-schattingen in figuur 3 over de periode 2021-2023 laat geen opvallende afwijkingen zien: slechts een klein deel van de KDS-reeks-gemiddelden valt buiten het betrouwbaarheidsinterval van de reeksgemiddelden van de directe schattingen.

¹⁰ De verdeling van de standaardfouten over de domeinen (RegioPlus-arbeidsmarktregio's) wordt weergegeven met een zogenaamde box-plot. De 'doos' van de box-plot bevat alle cijfers die tussen het eerste en derde kwartiel liggen (de afstand tussen deze grenzen heet interkwartielafstand) en een aanduiding van de mediane waarde. De staarten van de box-plot geven de waarnemingen aan die anderhalf maal de interkwartielafstand buiten de doos van de box-plot liggen, waarbij de staarten niet verder reiken dan de uiterste waarden van de verdeling. Uitbijters zijn op te vatten als datapunten die buiten de doos en staarten liggen; deze punten worden afzonderlijk weergegeven.

Figuur 3 Verschillen van reeksgemiddelden 2021-2023 per doelvariabele (verdeling over domeinen).



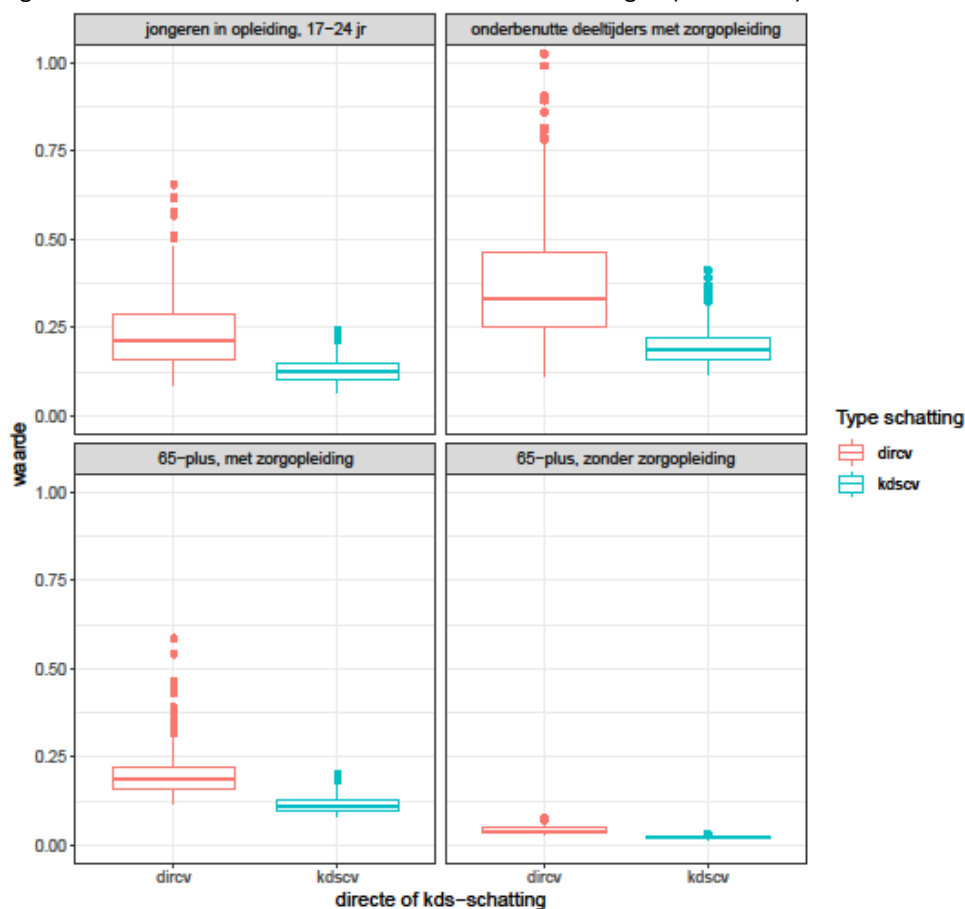
Voor een meer gedetailleerde vergelijking van GREG- en KDS-schattingen zie [appendix B](#). Het gaat in die vergelijking om kwartaalcijferreeksen van beide typen schatters per doelvariabele en RegioPlus-arbeidsmarktregio. De reeksen hebben betrekking op de periode 2021-2023. Over het algemeen wijken de reeksen van beide typen schatters niet veel van elkaar af.

5.3 Kwaliteitscriterium voor publicatie van regionale schattingen: de variatiecoëfficiënt

De kleinedomeinschattingen zijn voor alle vier de doelvariabelen preciezer dan voor vergelijkbare GREG-schattingen. Maar zijn ze ook precies genoeg voor publicatie? Om hierover uitsluitsel te geven wordt gebruikgemaakt van de zogenaamde variatiecoëfficiënt. De variatiecoëfficiënt (CV) van een schatting is eenvoudigweg de standaardfout van die schatting gedeeld door de schattingswaarde zelf. Bij een 'te hoge' waarde van de variatiecoëfficiënt wordt een schatting als onnauwkeurig bestempeld en wordt deze niet beschikbaar gesteld voor publicatie. Als grenswaarde wordt vaak 0,25 aangehouden: de standaardfout bedraagt dan een kwart van de geschatte waarde. Deze grenswaarde passen we in deze rapportage ook toe.

Figuur 4 bevat de variatiecoëfficiënten van de GREG- en KDS-schattingen (respectievelijk aangeduid met 'dircv' en 'kdscv'), waarbij per doelvariabele de verdeling van de variatiecoëfficiënten naar regio is weergegeven. Uit deze figuur blijkt dat voor alle vier de doelvariabelen de KDS-schattingen veelal precies genoeg zijn (ten minste 75 procent van de schattingen hebben een variatiecoëfficiënt kleiner dan 0.25) en bovendien veel preciezer zijn dan de overeenkomstige GREG-schattingen. De GREG-schatter levert bij twee van de vier doelvariabelen voldoende nauwkeurige schattingen.

Figuur 4 Variatiecoëfficiënten van GREG- en KDS-schattingen (2022-2023)



De cijfers in analyses van het haalbaarheidsonderzoek hebben betrekking op kwartaalcijfers, terwijl voor de publicatie van de regionale schattingen jaarcijfers worden beoogd. Deze jaarcijfers zijn uit de

kwartaalschattingen af te leiden, en zijn nauwkeuriger dan de kwartaalcijfers (de bijbehorende variatiecoëfficiënten van de jaarcijfers liggen voor alle regioplusgebieden onder de grenswaarde van 0,25). Zie ook de maatwerktabel met jaarcijfers voor ieder van de vier doelvariabelen aan het eind van deze rapportage.

6. Conclusies

Hoofdconclusies:

- Met behulp van kleinedomeinschatters kan de omvang van de vier doelpopulaties voor alle RegioPlus-arbeidsmarktregio's voldoende nauwkeurig worden geschat (verslagjaren 2022-2023):
 - o Jongeren van 17 tot 25 jaar die vanwege opleiding of studie niet kunnen of willen werken, niet direct beschikbaar zijn voor werk en ook niet recent naar werk hebben gezocht.
 - o Onderbenutte deeltijdwerkers met opleidingsrichting gezondheidszorg en welzijn.
 - o Niet-werkzame personen in de leeftijd van 65 tot 75 jaar die niet direct beschikbaar zijn voor werk en niet recent naar werk hebben gezocht vanwege vut/pensioen/hoge leeftijd met opleidingsrichting gezondheidszorg en welzijn.
 - o Niet-werkzame personen in de leeftijd van 65 tot 75 jaar die niet direct beschikbaar zijn voor werk en niet recent naar werk hebben gezocht vanwege vut/pensioen/hoge leeftijd zonder opleidingsrichting gezondheidszorg en welzijn.
- Een uitzondering hierop vormen de jaarcijfers over verslagjaar 2021 voor de doelpopulatie 'Onderbenutte deeltijdwerkers met opleidingsrichting gezondheidszorg en welzijn'. Voor dat verslagjaar zijn de jaarschattingen onvoldoende nauwkeurig. Dit hangt samen met de nog beperkte opbouw van het EBB-panel in dat jaar.

Andere conclusies:

Uit de modevaluaties blijkt dat de KDS-schattingen over het algemeen aan de toetscriteria voldoen:

- De KDS-schattingen wijken niet opvallend af van de directe schattingen, d.w.z. wijken grotendeels minder dan 1 standaardfout¹¹ af van de directe schattingen.
- De op landelijk niveau geaggregeerde KDS-schattingen komen goed overeen met de GREG-schattingen voor dezelfde doelvariabelen op basis van het EBB-weegmodel. De geaggregeerde KDS-schattingen van alle doelvariabelen, behalve de onderbenutte deeltijdwerkers, vallen binnen het betrouwbaarheidsinterval van de nationale GREG-schattingen (2023 Q1). De geaggregeerde KDS-schatting van de onderbenutte deeltijdwerkers wijkt significant maar gering af van het overeenkomstige GREG-cijfer. Dit verschil wordt mede veroorzaakt door het gebruik van verschillende modellen voor enerzijds de EBB-weging en anderzijds de KDS-schattingen.
- De standaardfouten van de KDS-schattingen zijn voor alle doelvariabelen beduidend kleiner dan voor de overeenkomstige directe schattingen (ongeveer twee maal kleiner).
- De KDS-reeksgegevens wijken voor het overgrote deel van de RegioPlus-arbeidsmarktregio's niet significant af van de overeenkomstige reeksgemiddelden van de directe schatter.
- De directe schatter is niet bruikbaar voor het maken van nauwkeurige regionale schattingen voor de twee doelpopulaties met 65-74 jarigen.

¹¹ Standaardfout behorend bij de directe schatting

Literatuur

Boonstra, H. J. (2012). hbsae: Hierarchical Bayesian Small Area Estimation. R package version 1.0.

Boonstra, H. J. en J.J.M. Michiels (2013). Evaluatie van kleine-domeinschatters (KDS) voor arbeidspositie op regionaal niveau. Interne CBS-nota.

Krieg, S. en J. Burger (2019). Regionale schattingen voor de statistiek investeringen. Interne CBS-nota.

Meijers, R, Burger, J. en H.J. Boonstra (2011). Nauwkeurigheid Enequête Beroepsbevolking. Interne CBS-nota.

Michiels, J.J.M. (2019). Haalbaarheidsonderzoek kleinedomeinschatters bij werknemerspanel AZW. Onderzoeksrapport CBS programma-AZW.

Rao, J. (2003). Small Area Estimation. Wiley-Interscience.

Särndal, C.-E., Swensson, B., en Wretman, J. (2003). Model assisted survey sampling. Springer Science & Business Media.

Vaida, F. and Blanchard, S. (2005). Conditional Akaike information for mixedeffects models. *Biometrika* 95 (2), 351–370.

Appendix

A categorieën van hulpvariabelen die als predictoren in de KDS-modellen zijn toegepast

Geslacht (2 categorieën)

- 1 Man,
- 2 Vrouw

Leeftijd (24 categorieën)

- 15jr, 16jr, 17jr, 18jr, 19jr, 20jr, 21jr, 22jr, 23jr, 24jr,
- 25-34jr, 35-44jr, 45-54jr, 55-64jr,
- 65jr, 66jr, 67jr, 68jr, 69jr, 70jr, 71jr, 72jr, 73jr, 74jr

Type huishouden (7 categorieën)

- 1 Eenpersoonshuishouden,
- 2 Niet gehuwd paar zonder kinderen,
- 3 Gehuwd paar zonder kinderen,
- 4 Niet gehuwd paar met kinderen,
- 5 Gehuwd paar met kinderen,
- 6 Eenouderhuishouden,
- 7 Overig huishouden

Onderwijsniveau (6 categorieën)

- 11 Basisonderwijs
- 12 Vmbo, avo onderbouw, mbo 1
- 21 Havo, vwo, mbo
- 31 Hbo-, wo-bachelor
- 32 Hbo-, wo-master, doctor
- 97 Onbekend

Onderwijsrichting (13 categorieën)

- 0000 Algemeen
- 0100 Onderwijs
- 0200 Vormgeving, kunst, talen en geschiedenis
- 0300 Journalistiek, gedrag en maatschappij
- 0400 Recht, administratie, handel en zakelijke dienstverlening
- 0500 Wiskunde, natuurwetenschappen
- 0600 Informatica
- 0700 techniek, industrie en bouwkunde
- 0800 Landbouw, diergeneeskunde en -verzorging
- 0900 Gezondheidszorg en welzijn
- 1000 Dienstverlening
- 9997 Missende waarde
- 9999 Weet niet/weigert/onbekend

GWU-inschrijving (7 categorieën)

- 1 Niet ingeschreven,
- 2 Ingeschreven met ww of cv, 15-44jr,
- 3 Ingeschreven met ww of cv, 55-54jr,
- 4 Ingeschreven met ww of cv, 55plus,
- 5 Ingeschreven overig, 15-44jr,
- 6 Ingeschreven overig, 45-54jr,
- 7 Ingeschreven overig, 55plus

Contractsoort (4 categorieën)

- 1 Contract voor bepaalde tijd
- 2 Contract voor onbepaalde tijd
- 3 Directeur-Groootaandeelhouder
- 4 Geen werknemer

Arbeidsduur (6 categorieën, uren per week)

- 1 0-12u,
- 2 12-19u,
- 3 20-24u,
- 4 25-29u,
- 5 30-34u,
- 6 35u of meer

Maandloon werknemer (6 categorieën, maandloon middelste maand kwartaal)

- 1 0 euro
- 2 1 – 1100 euro
- 3 1101 – 1200 euro
- 4 1201 – 3100 euro
- 5 3101 – 4400 euro
- 6 meer dan 4400 euro

Indicatie zelfstandige (2 categorieën)

- 0 Geen zelfstandige
- 1 Is zelfstandige

Indicatie studerende (2 categorieën)

- 0 Geen studerende
- 1 Is studerende

Indicatie ziekte/arbeidsongeschiktheid (2 categorieën)

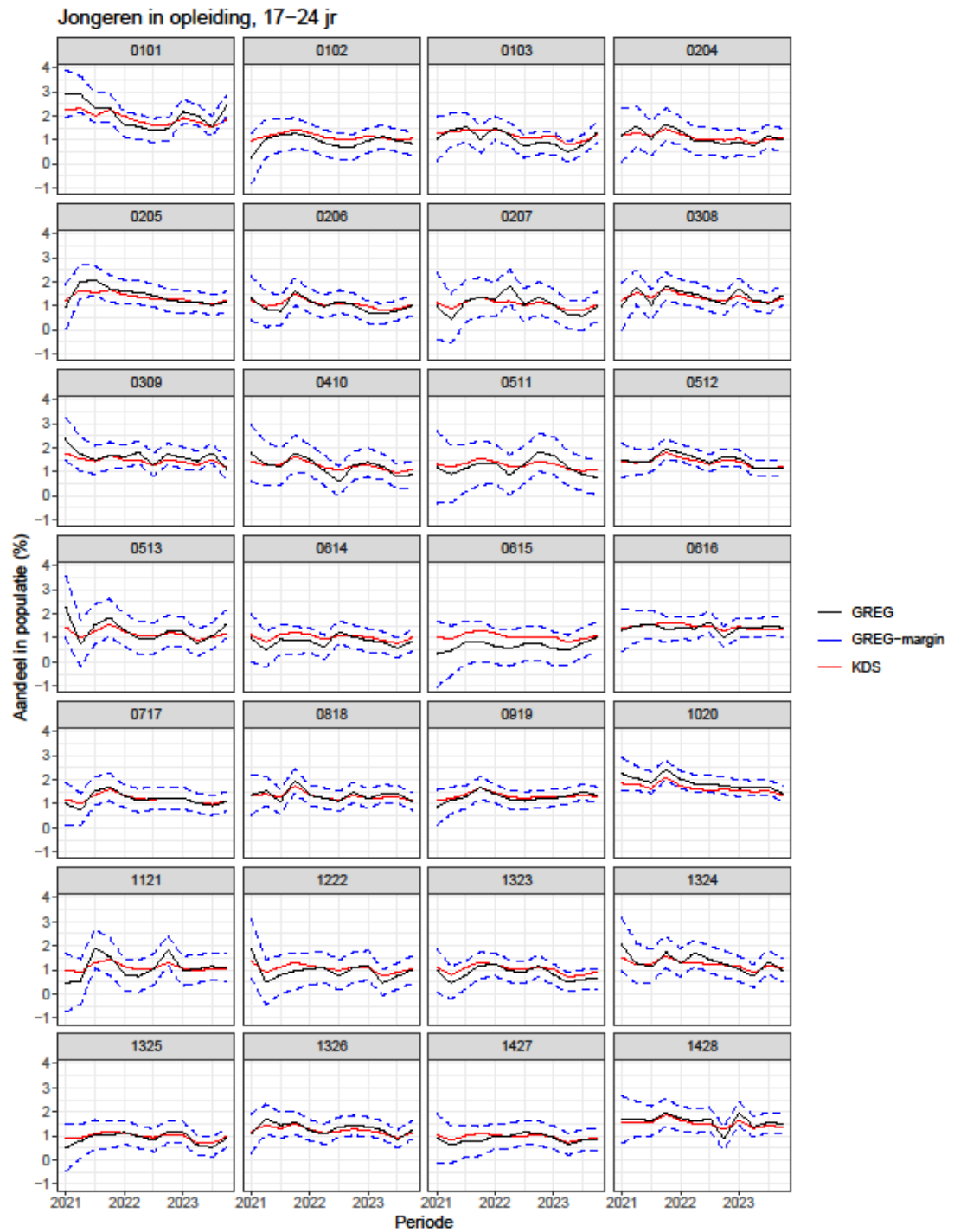
- 0 Geen ontvanger van een ziektewet- of arbeidsongeschiktheidsuitkering
- 1 Is uitkeringsontvanger (ziekte/arbeidsongeschiktheid)

Indicatie pensioen (2 categorieën)

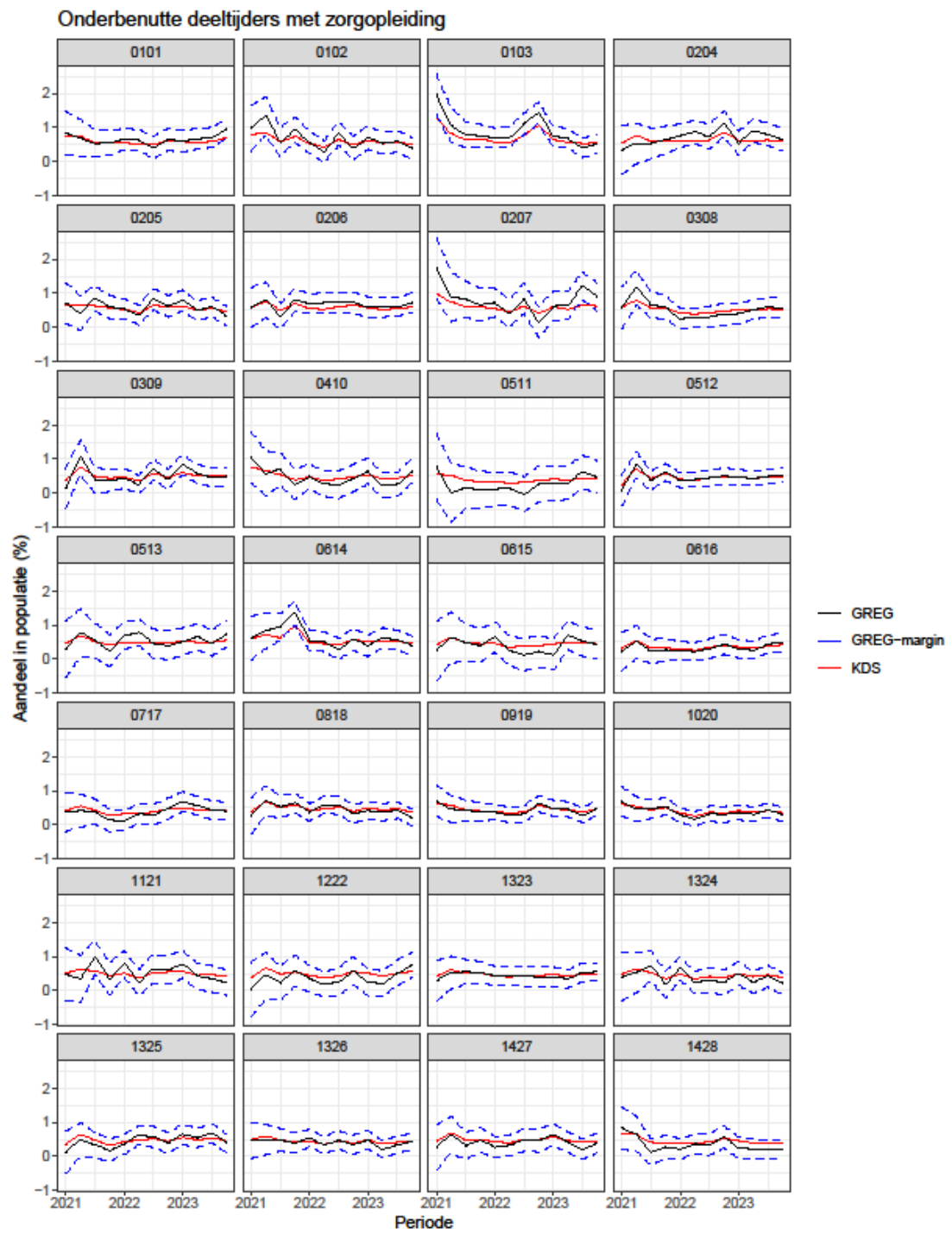
- 0 Geen ontvanger van een pensioenuitkering
- 1 Is uitkeringsontvanger (pensioen)

B reekscijfers doelvariabelen voor ieder van de RegioPlus-arbeidsmarktregio's

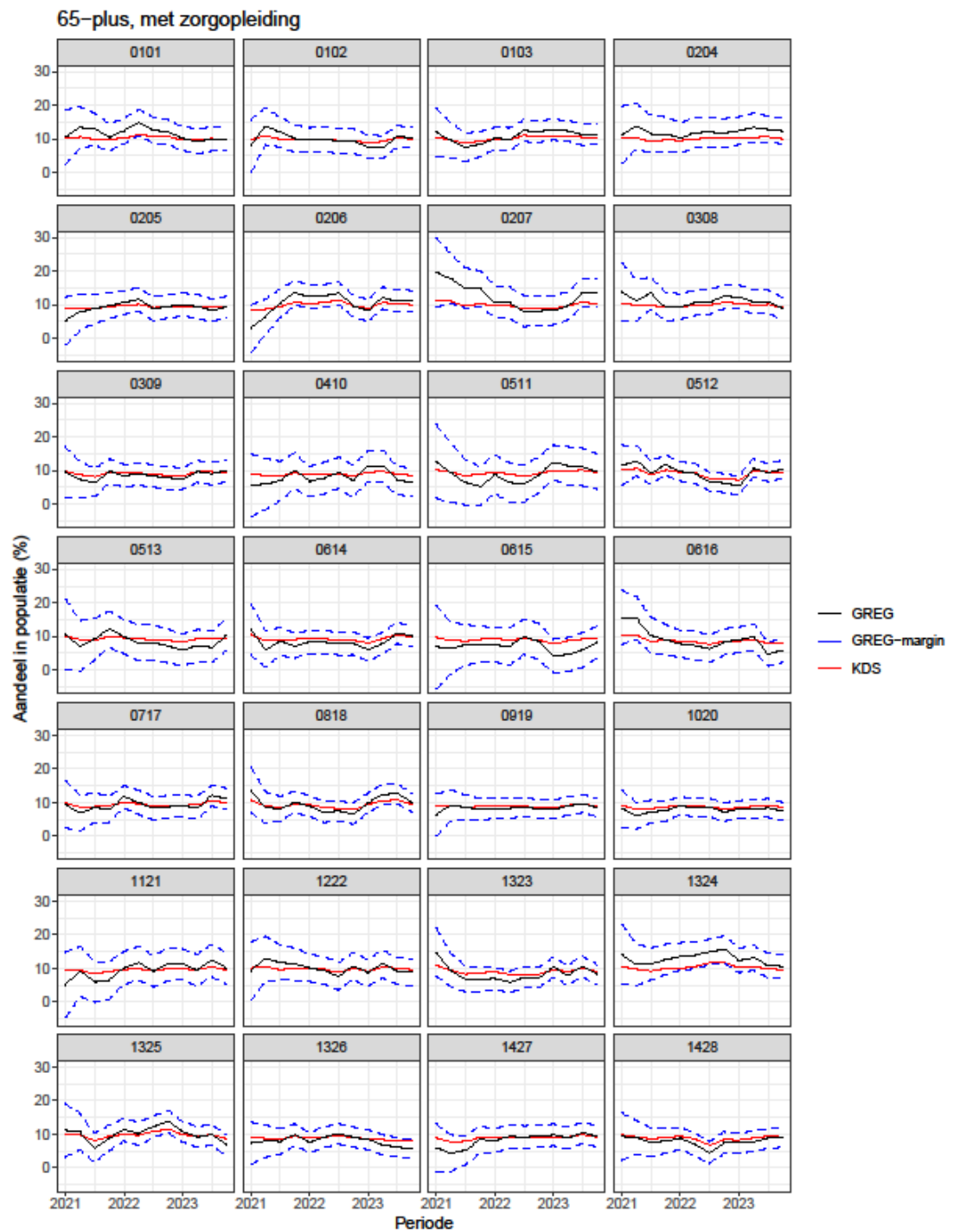
Figuur 5a



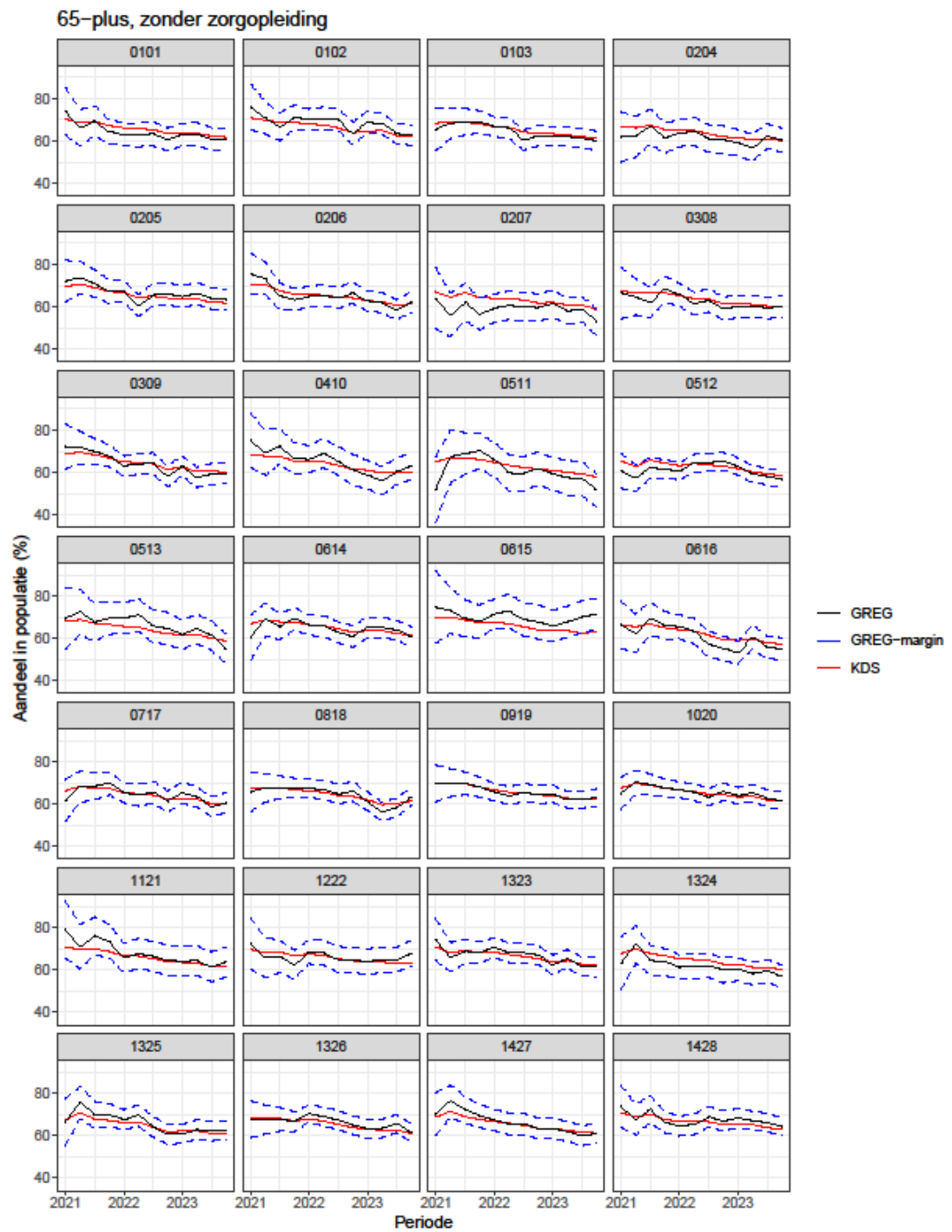
Figuur 5b.



Figuur 5c.



Figuur 5d.



Maatwerktabel regionale jaarcijfers, 2023

RegioPlus- arbeidsmarktregio	Jongeren 17-24 jr, niet werkend vanwege opleiding	Onderbenutte deeltijdwerkers met zorgopleiding	65-plus, werkt niet vanwege hoge leeftijd, met zorgopleiding	65-plus, werkt niet vanwege hoge leeftijd, geen zorgopleiding
	x1000			
Groningen	8	3	6	41
Drenthe	4	2	6	41
Friesland	5	3	8	49
Regio Zwolle	3	2	5	28
Twente	6	3	6	44
Stedendriehoek & Noord-Veluwe	5	3	7	46
Achterhoek	2	1	4	24
Zuid-West Gelderland	6	3	6	38
Midden-Gelderland	7	3	6	43
Flevoland	4	2	4	25
Gooi- en Vechtstreek	2	1	3	16
Utrecht en omgeving	11	4	9	63
Amersfoort en omgeving	3	1	3	20
Noord-Holland Noord	5	3	7	48
Zaanstreek en Waterland	3	1	3	23
Amsterdam	11	3	6	43
Amstelland, Kennemerland en Meerlanden	6	3	7	48
Rijn Gouwe	8	3	9	54
Rijnmond	13	5	11	77
Haaglanden en Nieuwe Waterweg Noord	15	5	11	79
Zuid-Holland Zuid	3	1	4	24
Zeeland	3	2	5	31
West-Brabant	5	3	8	52
Midden-Brabant	4	2	5	31
Noordoost-Brabant	4	3	7	45
Zuidoost-Brabant	7	3	7	54
Noord- en Midden- Limburg	4	2	6	43
Zuid-Limburg	7	2	7	51
Nederland¹²	161	72	174	1181

¹² De regio-schattingen zijn zodanig geschaald dat ze optellen tot de nationale cijfers. Deze afgeronde cijfers zijn in de maatwerktabel opgenomen. Door afrondingsfouten kunnen alsnog geringe verschillen optreden als de regionale cijfers in de tabel worden gesommeerd en vergeleken met de gepresenteerde nationale cijfers.