



rijksuniversiteit
 groningen

Wortels of vleugels?

*De rol van sociaal kapitaal bij het vertrek van jonge mensen uit
dorpen in niet-stedelijk gebied*

Masterscriptie van Kylian Halmingh

Studentnummer: S3807142

Begeleider: Dr. Jacob Dijkstra

Referent: Dr. Jaap Nieuwenhuis

Inleverdatum: 25-07-2025

Abstract

Jongere mensen vertrekken relatief vaker uit dorpen in niet-stedelijke gebieden. Dit draagt bij aan veroudering en bedreigt de leefbaarheid en voorzieningen. Deze scriptie onderzoekt kennishiaten over de rol van sociaal kapitaal bij dit vertrek. De onderzoeksvraag luidt: Welke rol speelt de locatie van sociaal kapitaal bij de hogere vertrekkans van jongere mensen uit dorpen in niet-stedelijke gebieden?

Op basis van bestaande inzichten wordt verwacht dat een hogere leeftijd samenhangt met een lagere intentie om uit het dorp te vertrekken en dat een hogere leeftijd gepaard gaat met relatief meer sociaal kapitaal binnen het eigen dorp. Op basis van het aspirations-capability framework wordt verwacht dat meer sociaal kapitaal binnen het eigen dorp samenhangt met een lagere vertrekintentie en dat sociaal kapitaal een mediërende rol speelt bij de hogere vertrekintentie van jongere mensen. Als alternatief op deze mediatiehypothese wordt de sociaalemotionele selectiviteitstheorie toegepast en wordt op basis hiervan verwacht dat de invloed van sociaal kapitaal op vertrekintentie sterker is bij oudere mensen dan bij jongere mensen.

Regressieanalyses zijn uitgevoerd op vragenlijstdata van een steekproef (16 jaar en ouder) uit drie dorpen in de Nederlandse gemeente Dantumadiel. Ontbrekende waarden zijn geïmputeerd via multiële imputatie (*MICE*). Vertrekintentie is gebruikt als voorspeller van vertrek. Relatief cognitief en structureel sociaal kapitaal binnen het eigen dorp zijn gemeten als percentage van iemands totale sociaal kapitaal dat zich binnen het eigen dorp bevindt.

Analyses bevestigen dat jongere mensen een hogere vertrekintentie hebben dan oudere mensen. De locatie van sociaal kapitaal verschilt echter niet significant tussen leeftijdsgroepen. De verwachte mediërende rol van sociaal kapitaal wordt daarmee niet bevestigd: het toevoegen van sociaal kapitaal aan het model verandert het verband tussen leeftijd en vertrekintentie nauwelijks. Het verband tussen cognitief sociaal kapitaal binnen dorp en vertrekintentie wordt, tegen de verwachting in, zwakker naarmate de leeftijd toeneemt en is alleen significant bij jongere respondenten. Vanaf 61 jaar wordt het verband zelfs positief. Ook relatief meer structureel sociaal kapitaal binnen het eigen dorp hangt zoals verwacht samen met lagere vertrekintentie. De verwachte modererende invloed van leeftijd op dit verband werd niet bewezen.

Terugkomend op de onderzoeksvraag blijkt dat de locatie van sociaal kapitaal invloed heeft op de vertrekkans van jongere mensen: bij relatief meer kapitaal binnen het eigen dorp is deze kans lager. De locatie van sociaal kapitaal verklaart echter niet waarom zij relatief vaker willen vertrekken omdat er geen leeftijdsverschillen zijn gevonden in de locatie van sociaal kapitaal.

Gemeenten met dorpen in niet-stedelijke gebieden wordt aangeraden beleid te richten op het versterken van sociaal kapitaal van jongere mensen binnen het eigen dorp en dit beleid te monitoren.

Inhoudsopgave

Abstract	2
Inhoudsopgave	3
1. Inleiding	5
1.1. Aanleiding.....	5
1.2. Centrale theoretische concepten.....	6
1.3. Bestaande inzichten	8
1.4. Hiaten in de bestaande inzichten	9
1.5. Onderzoeksvraag en deelvragen	9
1.6. Relevantie van het onderzoek	10
1.7. Hypothesen op basis van bestaande inzichten	10
2. Theoretisch kader	12
2.1. Aspirations-capabilities framework.....	12
2.1.1. Het aspirations-capabilities framework	12
2.1.2. Toepassing van het aspirations-capabilities framework.....	13
2.2. Sociaalemotionele selectiviteitstheorie	17
2.2.1. De sociaalemotionele selectiviteitstheorie	17
2.2.2. Toepassing van de sociaalemotionele selectiviteitstheorie.....	18
3. Methodologie	20
3.1. Dataverzameling	20
3.2. Operationalisatie van theoretische concepten	21
3.3. Ontbrekende waarden	29
3.4. Analyseplan	31
3.4.1. Imputeren van ontbrekende waarden	31
3.4.2. Kwaliteit van imputaties	33
3.4.3. Univariate statistieken voor en na imputatie.....	33
3.4.4. Bivariate statistieken voor en na imputatie	34
3.4.5. Modellen en plan van aanpak in MICE.....	34
3.4.6. Evaluatie invloed van geïmputeerde data	36
3.4.7. Modevaluatie	36
4. Voorbereidende analyses.....	37
4.1. Analyse kwaliteit van imputaties	37
4.2. Univariate statistieken voor en na imputatie	38
4.3. Bivariate statistieken voor en na imputatie	40
4.4. Analyse invloed van geïmputeerde data	43

4.5. Analyse modevaluatie	43
5. Resultaten.....	45
5.1. Hypothesetoetsing deelvraag 1: leeftijd en vertrek uit dorpen in niet-stedelijke gebieden	45
5.2. Hypothesetoetsing deelvraag 2: locatie sociaal kapitaal en leeftijd	46
5.3. Hypothesetoetsing deelvraag 3: sociaal kapitaal en vertrek uit dorpen in niet-stedelijke gebieden	46
5.4. Hypothesetoetsing deelvraag 4: Mediatie van leeftijdsverschillen in vertrek uit dorpen in niet-stedelijke gebieden door locatie sociaal kapitaal	47
5.5. Hypothesetoetsing deelvraag 5: Moderatie van verschillen in invloed sociaal kapitaal op vertrek uit dorpen in niet-stedelijke gebieden door leeftijd	48
6. Conclusie en discussie	50
6.1. Conclusie	50
6.2. Discussie	50
6.2.1. Interpretatie van resultaten	51
6.2.2. Validiteit en betrouwbaarheid.....	51
6.2.3. Beperkingen	52
6.2.4. Implicaties voor beleid	52
6.2.5. Suggesties voor wetenschappelijk vervolgonderzoek	53
7. Literatuur	54
Bijlage A: grafieken kwaliteit van imputaties	57
Bijlage B: invloed van ontbrekende waarden.....	77
Bijlage C: R-script.....	79
Bijlage D: R-namen bij vragenlijstitems en variabelen	93

1. Inleiding

De Nederlandse bevolking veroudert. De gemiddelde leeftijd van de Nederlander is van 30,8 jaar in 1950 gestegen naar 42,2 jaar in 2022 (Centraal Bureau voor de Statistiek, z.d.-d). Dit is een stijging van 37,0%. Ter vergelijking: over dezelfde periode is de levensverwachting in Nederland slechts met 14,3% gestegen van 71,4 naar 81,6 jaar (Centraal Bureau voor de Statistiek, z.d.-e).

Niet-stedelijke gebieden vergrijzen sneller dan stedelijke gebieden (Verwey-Jonker Instituut, 2022). Dit heeft volgens dit rapport te maken met de trek van jongeren uit het platteland. Data van het Centraal Bureau voor de Statistiek (hierna CBS) geven hier concreet inzicht in. Door de leeftijdsverdeling per gemeente in 2024 (via Centraal Bureau voor de Statistiek, z.d.-c) te koppelen aan de indeling op verstedelijkingsniveau van het CBS (via Centraal Bureau voor de Statistiek, z.d.-b), wordt duidelijk dat niet-stedelijke gebieden sterker verouderen dan stedelijke gebieden.

In de gemiddelde niet-stedelijke (minder dan 500 adressen per km²) gemeente is 24,9% van de inwoners 65 jaar of ouder. Dit percentage daalt naar 23,9% bij weinig stedelijk (500 tot 1000 per km²), 22,7% bij matig stedelijk (1000 tot 1500 per km²), 21,4% bij sterk stedelijk (1500 tot 2500 per km²) en 18,1% bij zeer sterk stedelijk (meer dan 2500 per km²).

Tegenover het grotere aandeel ouderen in niet-stedelijke gemeenten, staat juist een kleiner aandeel jonge mensen. In de gemiddelde niet-stedelijke gemeente is 25,4% van de inwoners jonger dan 25 jaar. Dit percentage stijgt naar 26,2% bij weinig stedelijk, 27,0% bij matig stedelijk, 27,2% bij sterk stedelijk en 28,5% bij zeer sterk stedelijk.

1.1. Aanleiding

Het rapport Staatscommissie Demografische Ontwikkelingen 2050 (Ministerie van Sociale Zaken en Werkgelegenheid, 2024) slaat alarm voor de negatieve gevolgen voor gebieden met een ouder wordende populatie. Een ouder wordende populatie gaat volgens dit rapport gepaard met een grotere vraag naar zorg terwijl er tegelijkertijd juist minder personeel beschikbaar is voor zorgaanbieders zoals huisartsen, thuiszorg en zorginstellingen. Ook kan het onderwijs onder druk komen te staan doordat leraren moeilijker te vervangen zijn en doordat het niet meer haalbaar is om onderwijs in elke regio voor alle doelgroepen overeind te houden. Verder blijkt uit Amerikaans onderzoek dat het voorzieningenniveau in verouderende dorpen afneemt (Thiede et al., 2016). Dit effect werkt vaak dubbelop in niet-stedelijke gebieden omdat daar vaak ook een laag voorzieningenniveau is in naburige dorpen. Daarnaast biedt veroudering van dorpen uitdagingen voor de beschikbaarheid van mantelzorg (Cohen & Greaney, 2023). Als de gemiddelde leeftijd hoger wordt, zijn er immers meer mensen die mantelzorg nodig hebben en minder die dit kunnen geven. Doordat de veroudering in niet-stedelijke gebieden sterker is dan in stedelijke gebieden zijn deze gebieden extra kwetsbaar voor deze negatieve gevolgen van veroudering. Bovendien blijkt uit onderzoek dat jonge mensen geneigd zijn te vertrekken uit regio's waar de voorzieningen en economische kansen afnemen (Giannone et al., 2024). Dit kan leiden tot een zelfversterkend effect waarbij jonge mensen wegtrekken, de voorzieningen afnemen, en daardoor weer meer jonge mensen wegtrekken.

Dit is aanleiding om te onderzoeken welke oorzaken er zijn voor het fenomeen dat jongere mensen eerder vertrekken uit deze niet-stedelijke gebieden. De focus ligt in dit onderzoek op de rol die sociaal kapitaal hierbij speelt vanwege aanwijzingen in de bestaande literatuur dat

sociaal kapitaal afzonderlijk zowel samenhangt met leeftijd als met vertrek. In dit onderzoek worden de onderlinge relaties tussen deze drie theoretische concepten gezamenlijk onderzocht om bij te dragen aan kennis over de redenen waarom jongere mensen eerder vertrekken uit niet-stedelijke gebieden.

Allereerst worden de drie centrale theoretische concepten besproken. Vervolgens worden bestaande inzichten in kaart gebracht, waarna de probleemstelling en deelvragen worden geformuleerd. Dit hoofdstuk wordt afgesloten met een bespreking van de relevantie van dit onderzoek.

1.2. Centrale theoretische concepten

In deze paragraaf worden de drie theoretische concepten die centraal staan in dit onderzoek toegelicht. Daarmee wordt duidelijk wat er met deze concepten bedoeld wordt in de verdere uitwerking van het onderzoek en bij het beantwoorden van de onderzoeksvraag.

Vertrekintentie

Dit onderzoek richt zich op het verklaren van de mate waarin mensen willen vertrekken uit het dorp waarin zij wonen. Idealiter zou je hiervoor een beeld krijgen van de mate waarin mensen vertrekken en van de factoren die daar een rol in spelen. De periode waarin de stageopdracht gecombineerd met de scriptieopdracht is uitgevoerd, beslaat echter een te korte periode om een dubbele meting te doen waarin getoetst kan worden of de mensen ook daadwerkelijk vertrekken. Daarom wordt aan respondenten gevraagd in hoeverre zij de intentie hebben om te vertrekken.

Volgens de Groot, Manting en Mulder (2011) is de kans beduidend groter dat mensen die de intentie hebben om binnen twee jaar te vertrekken, ook daadwerkelijk vertrekken. Zij toetsten namelijk met het vragenlijstitem ‘Wilt u binnen twee jaar verhuizen?’ de algemene intentie van Nederlanders om te verhuizen. Van degenen zonder vertrekintentie (antwoordopties ‘Zeer zeker niet’ en ‘Weet ik niet’) vertrok binnen twee jaar slechts 6 procent. Van mensen met een minder sterke vertrekintentie (antwoordopties ‘Mogelijk wel, misschien’ of ‘Ik zou wel willen, maar ik kan niets vinden’) vertrok 21 procent en van degenen met een sterke vertrekintentie (antwoordoptie ‘Zeer zeker wel’) vertrok liefst 50 procent. In deze scriptie is er daarom voor gekozen om de vertrekintentie te gebruiken als voorspeller van een daadwerkelijk vertrek.

Leeftijd

Met de leeftijd wordt in dit onderzoek bedoeld op de leeftijd in jaren van personen.

Sociaal kapitaal

In dit onderzoek wordt de veelgebruikte definitie van Putnam (1995) voor sociaal kapitaal gehanteerd. Hij definieert sociaal kapitaal als “Kenmerken van sociale organisatie, zoals netwerken, normen en sociaal vertrouwen, die coördinatie en samenwerking voor wederzijds voordeel bevorderen.” Omdat sociaal kapitaal voordelen oplevert, is het een vorm van kapitaal.

Sociaal kapitaal is een breed concept dat op verschillende manieren kan worden geïnterpreteerd en toegepast. Naast deze algemene definitie van sociaal kapitaal wordt daarom nu toegelicht hoe sociaal kapitaal specifiek wordt toegepast in dit onderzoek. Dit gebeurt via de vragen: ‘Aan wie behoort sociaal kapitaal toe?’, ‘Op welk niveau wordt sociaal kapitaal in dit

onderzoek gemeten?', 'Waar bevindt het sociaal kapitaal zich?' en 'Welke vormen van sociaal kapitaal worden onderscheiden?'.

Aan wie behoort sociaal kapitaal toe?

Binnen de wetenschap bestaat discussie over de vraag of sociaal kapitaal behoort aan het individu (volgens bijvoorbeeld Fukuyama, 2001), aan het collectief (volgens bijvoorbeeld Bourdieu, 1986) of allebei (volgens bijvoorbeeld Putnam, 2000). Onderzoek van Alguezaui & Filieri (2010) ondersteunt dat sociaal kapitaal aan het individu behoort aangezien individuen kunnen investeren in hun sociaal kapitaal, hier een bepaalde mate van eigenaarschap en controle over hebben, en exclusieve voordelen uit hun sociale kapitaal kunnen halen.

Toegepast op de definitie van Putnam (1995) kan een persoon dus investeren in bijvoorbeeld de eigen netwerken, normen en sociaal vertrouwen, en daar zelf ook voordelen uit behalen. In dit onderzoek wordt de lijn van sociaal kapitaal als behorend aan het individu aangehouden. Deze keuze is passend omdat dit onderzoek zich specifiek richt op de invloed van het sociale kapitaal van het individu zelf op diens eigen vertrekvens.

Op welk niveau wordt sociaal kapitaal in dit onderzoek gemeten?

Sociaal kapitaal wordt gemeten op verschillende niveaus, namelijk micro (niveau van het individu), meso (niveau van groepen) en macro (niveau van de samenleving) (Halpern, 2005). In dit onderzoek ligt de focus op het sociale kapitaal dat het individu zelf heeft. Dit is dus het microniveau. Concreet wordt dus in lijn met de definitie van Putnam (1995) gemeten in hoeverre individuen zelf kenmerken van sociaal kapitaal zoals netwerken, normen en sociaal vertrouwen bezitten.

Waar bevindt het sociaal kapitaal zich?

In dit onderzoek ligt de nadruk op de rol van sociaal kapitaal bij het vertrek uit dorpen. Hiervoor is de locatie waar een individu het sociaal kapitaal bezit relevant. Er worden twee locaties onderscheiden: sociaal kapitaal dat men binnen het eigen dorp bezit en sociaal kapitaal dat men buiten het eigen dorp bezit. In lijn met de definitie van Putnam (1995) betekent dit dat bijvoorbeeld netwerken, normen en sociaal vertrouwen zich zowel binnen als buiten het eigen dorp kunnen bevinden.

Een concreet voorbeeld is deelname aan een maandelijkse bingoavond bij een vereniging. Als iemand deelneemt aan zo'n activiteit in het eigen dorp bevindt dit deel van het sociaal kapitaal zich binnen het dorp. Vindt diezelfde activiteit plaats in een nabijgelegen stad dan behoort dit deel van het sociaal kapitaal tot buiten het dorp.

Dit onderscheid zorgt ervoor dat berekend kan worden hoeveel sociaal kapitaal iemand binnen het eigen dorp bezit en hoeveel daarbuiten. Hiermee kan weer worden berekend in welke mate het sociale kapitaal relatief binnen of buiten het eigen dorp gericht is.

Welke vormen van sociaal kapitaal worden onderscheiden?

Sociaal kapitaal zoals gedefinieerd door Putnam (1995) is een breed en overkoepelend begrip. Putnam beschrijft sociaal kapitaal als "Kenmerken van sociale organisatie, zoals netwerken, normen en sociaal vertrouwen, die coördinatie en samenwerking voor wederzijds voordeel bevorderen." Deze omschrijving benoemt de centrale onderdelen van sociaal kapitaal, maar maakt nog geen onderscheid tussen verschillende vormen sociaal kapitaal.

Om verschillende vormen van sociaal kapitaal concreet in dit onderzoek te kunnen onderscheiden wordt in dit onderzoek het model van Bain en Hicks (1998) gehanteerd. Sociaal kapitaal wordt hierin uitgesplitst in cognitief en structureel sociaal kapitaal (Eriksson, 2011). Cognitief sociaal kapitaal wordt omschreven als de minder tastbare kant van sociaal kapitaal (Krishna & Shrader, 2000). Hieronder vallen bijvoorbeeld normen van vertrouwen, solidariteit en wederkerigheid. Structureel sociaal kapitaal refereert daarentegen aan de opbouw, omvang en activiteit van bijvoorbeeld instituties of netwerken. Kort gezegd gaat cognitief sociaal kapitaal dus over wat mensen voelen met betrekking tot hun sociale relaties terwijl structureel sociaal kapitaal gaat over wat mensen doen.

In eerder onderzoek van Ejlskov et al. (2014) is deze tweedeling toegepast. Hierbij hanteren zij de volgende concrete definities: Cognitief sociaal kapitaal geeft de mate van vertrouwen in anderen en de perceptie van delen en wederkerigheid weer terwijl structureel sociaal kapitaal betrekking heeft op de dichtheid van het sociale netwerk van een persoon en diens betrokkenheid bij maatschappelijke activiteiten. Deze definities worden in dit onderzoek aangehouden voor de operationalisatie van cognitief en structureel sociaal kapitaal.

Deze tweedeling sluit aan bij de definitie van sociaal kapitaal zoals geformuleerd door Putnam (1995) die sociaal kapitaal beschrijft als kenmerken van sociale organisatie zoals netwerken, gedeelde normen en sociaal vertrouwen. Binnen deze benadering vallen normen en sociaal vertrouwen onder cognitief sociaal kapitaal, terwijl netwerken behoren tot het structurele sociaal kapitaal.

In dit onderzoek wordt dus uitgegaan van sociaal kapitaal behorend aan het individu en wordt het op microniveau gemeten. Hierbij zal zowel het kapitaal dat een persoon binnen het eigen dorp bezit als het kapitaal dat een persoon buiten het eigen dorp bezit worden gemeten. Ook wordt een onderscheid gemaakt tussen cognitief en structureel sociaal kapitaal.

1.3. Bestaande inzichten

Verhuizen en leeftijd

Uit onderzoek van Long (1992) in verschillende westerse landen komt consistent het patroon naar voren dat het percentage individuen dat in een jaar verhuist het hoogst is bij de leeftijdsgroep 20-24 jaar. Het percentage van elke leeftijdsgroep dat verhuist neemt daarna met de leeftijd af. Ook bij jonge kinderen ligt het percentage relatief hoog doordat zij met hun tevens relatief jonge ouders meegaan. Hieruit blijkt dat jongere mensen gemiddeld vaker verhuizen dan oudere leeftijdsgroepen.

Dit vormt een uitdaging voor dorpen als deze jonge individuen verhuizen naar woonplaatsen buiten het dorp. In de Nederlandse context beweren Haartsen & Venhorst (2010) dat jongere mensen eerder wegtrekken uit niet-stedelijke gemeenten terwijl deze gemeenten juist aantrekkelijk zijn voor oudere leeftijdsgroepen. De gemiddelde leeftijd van vertrekkers is in deze gebieden dus lager dan die van nieuwkomers. Dit fenomeen draagt bij aan de veroudering van de bevolking in deze niet-stedelijke gebieden.

Leeftijdsverschillen in locatie van sociaal kapitaal

De locatie van sociaal kapitaal blijkt te variëren tussen leeftijdsgroepen. Deens onderzoek toont aan dat bij oudere leeftijdsgroepen het vertrouwen in leden van de lokale gemeenschap ten

opzichte van vertrouwen in mensen in het algemeen hoger is dan bij jongere leeftijdsgroepen (Sørensen, 2014).

Uit Zwitsers onderzoek blijken jongere individuen daarnaast de hoogste kans te hebben om sterke banden te hebben met mensen buiten hun eigen gemeenschap en blijkt deze kans af te nemen naarmate de leeftijd vordert (Ohnmacht, 2009). De oudste leeftijdsgroep heeft de laagste kans op het onderhouden van sterke banden met mensen buiten hun eigen gemeenschap.

Locatie van sociaal kapitaal en vertrek uit woonplaats

Als zowel de kans op vertrek uit het dorp als de locatie van het sociaal kapitaal verschillen per leeftijdsgroep, is het interessant om te onderzoeken of een verschillende locatie waar het sociaal kapitaal zich bevindt invloed heeft op de vertrekans van mensen uit hun woonplaats. David, Janiak & Wasmer (2010) tonen in een onderzoek in verschillende Europese landen, waaronder Nederland, namelijk aan dat het hebben van meer lokaal sociaal kapitaal de kans verkleint dat het individu die regio verlaat. Aan de andere kant blijkt dat het hebben van sociaal kapitaal buiten de woonplek de kans vergroot dat een individu wil verhuizen naar de plek waar deze al sociaal kapitaal heeft opgebouwd (Büchel et al., 2020).

1.4. Hiaten in de bestaande inzichten

Dit onderzoek richt zich op twee hiaten in de bestaande inzichten over de relatief hoge vertrekcijfers van jongere mensen uit niet-stedelijke gebieden. Volgens het aspirations-capabilities framework van de Haas (2021) vertrekken mensen eerder als ze de wens hebben en in staat zijn om die stap te zetten. Het is aannemelijk dat sociaal kapitaal hierin een rol speelt: wie meer sociaal kapitaal buiten het dorp heeft zal eerder de wens en mogelijkheid ervaren om te vertrekken. Omdat jongere mensen relatief meer sociaal kapitaal buiten het dorp hebben, is het aannemelijk dat dit hun hogere vertrekans verklaart. Deze mogelijke mediërende rol van sociaal kapitaal bij het hogere vertrek van jongere mensen is echter nog niet onderzocht.

Een alternatieve verklaring komt voort uit de sociaalemotionele selectiviteitstheorie (Carstensen, 1995). Deze stelt dat sociaal kapitaal belangrijker wordt naarmate mensen ouder worden. Het is aannemelijk dat sociaal kapitaal dan ook zwaarder meeweegt in vertrekbeslissingen van ouderen dan van jongere mensen. Omdat oudere mensen relatief meer sociaal kapitaal binnen het dorp hebben en dit kapitaal bij hen sterker doorwerkt, zouden zij minder snel geneigd zijn te vertrekken dan jongere mensen. Deze mogelijke modererende rol van leeftijd in de relatie tussen sociaal kapitaal en vertrekintentie is tot nu toe echter ook nog niet onderzocht.

1.5. Onderzoeksvraag en deelvragen

Op basis van de bestaande inzichten en de besproken hiaten in de bestaande inzichten is de onderzoeksvraag als volgt: Welke rol speelt de locatie van sociaal kapitaal bij de hogere vertrekans van jongere mensen uit dorpen in niet-stedelijke gebieden? Om tot een antwoord op deze onderzoeksvraag te komen, worden de volgende deelvragen onderzocht:

Deelvraag 1: Hoe beïnvloedt leeftijd de kans op vertrek uit dorpen in niet-stedelijke gebieden?

Deelvraag 2: Hoe verschilt de locatie van sociaal kapitaal tussen leeftijdsgroepen?

Deelvraag 3: Welke rol speelt sociaal kapitaal bij de vertrekkans uit het eigen dorp?

Deelvraag 4: In hoeverre medieert de locatie van sociaal kapitaal de relatie tussen leeftijd en vertrekintentie uit dorpen in niet-stedelijke gebieden?

Deelvraag 5: In hoeverre modereert leeftijd de relatie tussen de locatie van sociaal kapitaal en vertrekintentie uit dorpen in niet-stedelijke gebieden?

1.6. Relevantie van het onderzoek

De hoofdreden waarom dit onderzoek relevant is voor de sociologische wetenschap is dat het kan helpen om de besproken kennishiaten in de wetenschap over het relatief hoge vertrekcijfer van jongere mensen op te vullen.

Aangezien een relatief hoog vertrekpercentage van jongere mensen de gemiddelde leeftijd in de woonplaats verhoogt, geeft dit onderzoek ook meer kennis over oorzaken van veroudering van dorpen in niet-stedelijke gebieden. Deze veroudering levert zoals in de aanleiding besproken problemen op voor de betreffende gemeenschappen: zorg, onderwijs en voorzieningen kunnen hierdoor namelijk onder druk komen te staan.

Het opvullen van de kennishiaten kan beleidsmakers helpen om beleidsmaatregelen betreffende sociaal kapitaal effectiever in te zetten doordat meer bekend wordt over de rol van sociaal kapitaal bij de intentie om de eigen woonplaats te verlaten. Zo kunnen beleidsinstrumenten specifiekere worden ingericht om inwoners in deze leeftijdsgroepen te stimuleren om het dorp niet te verlaten.

1.7. Hypothesen op basis van bestaande inzichten

Voordat in hoofdstuk 2 het theoretisch kader wordt uitgewerkt, worden eerst op basis van bestaande inzichten hypothesen opgesteld die horen bij de eerste twee deelvragen. Deze deelvragen richten zich op mogelijke leeftijdsverschillen in vertrekintentie en in de locatie van sociaal kapitaal. Inzicht in deze verschillen is belangrijk om later te kunnen onderzoeken welke rol sociaal kapitaal speelt bij de relatief hogere vertrekintentie van jongere mensen uit dorpen in niet-stedelijke gebieden.

De eerste deelvraag is ‘Hoe beïnvloedt leeftijd de kans op vertrek uit dorpen in niet-stedelijke gebieden?’. Uit de bestaande inzichten komt naar voren dat jonge mensen gemiddeld vaker verhuizen dan oudere mensen (Long, 1992). In de Nederlandse context trekken jongere mensen eerder weg uit niet-stedelijke gemeenten (Haartsen & Venhorst, 2010).

Dit ondersteunt dat er een negatief verband is tussen leeftijd en vertrekkans. In dit onderzoek wordt vertrekintentie gebruikt als voorspeller van de daadwerkelijke vertrekkans. Hieruit volgt de volgende hypothese:

H1: Hogere leeftijd hangt samen met een lagere vertrekintentie uit dorpen in niet-stedelijke gebieden.

De tweede deelvraag is: ‘Hoe verschilt de locatie van sociaal kapitaal tussen leeftijdsgroepen?’. Uit de bestaande inzichten komt naar voren dat bij oudere leeftijdsgroepen het vertrouwen in leden van de lokale gemeenschap ten opzichte van mensen in het algemeen hoger is dan bij jongere leeftijdsgroepen (Sørensen, 2014). Ook is besproken dat de kans op sterke sociale

banden met mensen buiten de eigen gemeenschap hoger is bij jongere leeftijdsgroepen dan bij oudere leeftijdsgroepen (Ohnmacht, 2009).

Dit ondersteunt dat er een positief verband is tussen leeftijd en sociaal kapitaal binnen het eigen dorp en een negatief verband tussen leeftijd en sociaal kapitaal buiten het eigen dorp. Hieruit volgt de volgende hypothese:

H2: Hogere leeftijd hangt samen met relatief meer sociaal kapitaal binnen het eigen dorp en relatief minder sociaal kapitaal buiten het eigen dorp.

2. Theoretisch kader

In dit hoofdstuk worden twee theorieën toegelicht en toegepast en worden op basis hiervan hypothesen opgesteld voor deelvragen 3 tot en met 5. In paragraaf 2.1. wordt het aspirations-capabilities framework van de Haas (2021) toegelicht en toegepast en in paragraaf 2.2 de sociaalemotionele selectiviteitstheorie van Carstensen (1995).

2.1. Aspirations-capabilities framework

Om theoretisch onderbouwde hypothesen voor deelvraag 3 en 4 te formuleren, wordt het aspirations-capabilities framework (vanaf nu ACF) toegelicht en toegepast op de relatie tussen sociaal kapitaal en vertrekintentie.

2.1.1. Het aspirations-capabilities framework

Het ACF is een migratietheorie ontwikkeld door Hein de Haas (2021) waarin de aspiratie en capaciteit om te migreren centraal staan. Volgens deze theorie is de beslissing om wel of niet te migreren een functie van deze twee elementen.

Met aspiratie wordt bedoeld op de mate waarin iemand wenst om te migreren. De aspiratie om al dan niet te migreren is volgens het ACF een functie van de algemene wensen die een persoon in het leven heeft en de veronderstelde geografische kansenstructuren (de Haas, 2021). Concreet betekent dit dat de aspiratie afhankelijk is van wat de persoon graag in het leven wil en van de plek waar de persoon dit doel kan bereiken. Bijvoorbeeld: als een belangrijke wens in het leven is om naast een natuurgebied te wonen en je woont nu naast een natuurgebied, dan verlaagt dit de aspiratie om te vertrekken. Het ACF onderscheidt instrumentele en intrinsieke aspiraties. Instrumentele aspiraties verwijzen naar migratie als een middel om een specifiek doel te bereiken, zoals een hoger inkomen, betere sociale status, gezondheidszorg of onderwijs. Intrinsieke aspiraties daarentegen zien migratie als een doel op zich, waarbij de ervaring, ontdekking en het sociale prestige van migratie zelf als waardevol worden beschouwd. In de praktijk kunnen beide soorten aspiraties gelijktijdig optreden en elkaar versterken.

Met capaciteit wordt bedoeld op de mate waarin iemand zelf de keuze kan maken om wel of juist niet te migreren. Deze capaciteit is afhankelijk van de positieve en negatieve vrijheid van de persoon (de Haas, 2021). Positieve vrijheid is hier de mate waarin iemand de vrijheid heeft om controle over het eigen leven te nemen en zijn fundamentele doelen te realiseren. Denk bijvoorbeeld aan het bezitten van voldoende geld om een huis te kopen. Negatieve vrijheid is de mate waarin de vrijheid niet ingeperkt wordt door obstakels, barrières of beperkingen. De negatieve vrijheid om te vertrekken wordt bijvoorbeeld ingeperkt wanneer teamgenoten uit het dorp iemand proberen te overtuigen om te blijven door te benadrukken dat het lokale voetbalteam zonder hem niet kan doorgaan. Volgens het ACF hangt meer capaciteit om zelf te beslissen over wel of niet migreren daarnaast samen met een hogere aspiratie.

2.1.2. Toepassing van het aspirations-capabilities framework

Eerst wordt uitgelegd waarom sociaal kapitaal in het licht van het ACF van de Haas (2021) met een relatieve maat worden gemeten. Om verwachtingen voor de derde en vierde deelvraag over de rol van sociaal kapitaal bij vertrekintentie te kunnen opstellen wordt dit ACF daarna toegepast. Hiervoor is relevant dat sociaal kapitaal volgens Putnam (1995) een middel is om voordelen te verkrijgen die mensen zonder sociaal kapitaal niet zouden verkrijgen. Als men sociaal kapitaal in het eigen dorp heeft dan levert dit voordelen binnen het dorp op. Als men juist sociaal kapitaal buiten het eigen dorp heeft dan zijn hiermee voordelen buiten het dorp te behalen.

In dit onderzoek wordt sociaal kapitaal geoperationaliseerd als het relatieve aandeel van iemands totale cognitieve en het relatieve aandeel van iemands totale structurele sociaal kapitaal dat zich binnen het eigen dorp bevindt. De concrete operationalisatie hiervan wordt toegelicht in paragraaf 3.2.

Deze keuze voor een relatieve maat sluit allereerst goed aan bij het aspiratie-element van het ACF waarin wordt gesteld dat migratie-aspiraties voortkomen uit algemene wensen die een persoon in het leven heeft en veronderstelde geografische kansenstructuren (de Haas, 2021). Deze aspiraties worden dus mede gevormd doordat mensen hun kansen binnen het dorp vergelijken met hun kansen daarbuiten. Dit zijn dus relatieve kansen. In de eerste alinea van deze paragraaf is besproken dat sociaal kapitaal een middel is om voordelen te verkrijgen. Sociaal kapitaal binnen het dorp levert voordelen binnen het dorp op en sociaal kapitaal buiten het dorp levert voordelen buiten het dorp op. In de context van het ACF kunnen deze voordelen van sociaal kapitaal worden gezien als kansen binnen of buiten het dorp. Deze kansen hebben ook invloed op de aspiratie om te vertrekken doordat de relatieve kansen binnen en buiten het dorp zorgen voor een bepaalde mate van migratie-aspiratie. Een relatieve maat voor sociaal kapitaal binnen het eigen dorp sluit hier theoretisch beter bij aan dan een absolute maat omdat hiermee duidelijker naar voren komt in hoeverre iemand relatief meer of minder sociaal kapitaal binnen het eigen dorp heeft en daarmee relatief meer of minder kansen binnen het dorp ten opzichte van daarbuiten. Dit maakt vervolgens inzichtelijker hoe dit sociaal kapitaal via die kansen invloed kan hebben op migratie-aspiraties en daardoor op vertrekintentie.

Daarnaast sluit de keuze voor een relatieve maat ook aan bij het capaciteitselement van het ACF. Binnen dit framework wordt capaciteit gezien als de mogelijkheid om migratie daadwerkelijk te realiseren en wordt deze bepaald door zowel positieve als negatieve vrijheid (de Haas, 2021). Deze vrijheden zijn plaatsgebonden. Meer sociaal kapitaal binnen het eigen dorp vergroot bijvoorbeeld de positieve vrijheid om binnen het dorp doelen te realiseren maar beperkt tegelijkertijd de positieve vrijheid om deze doelen buiten het dorp te realiseren. Ook negatieve vrijheid is afhankelijk van de locatie van sociaal kapitaal: sociaal kapitaal binnen het eigen dorp gaat gepaard met minder negatieve vrijheid om te vertrekken terwijl sociaal kapitaal buiten het dorp juist zorgt voor minder negatieve vrijheid om te blijven. Belangrijk is hierbij dat capaciteit binnen het ACF, net als aspiratie, relatief is. De positieve en negatieve vrijheid die mensen in het huidige dorp hebben, staan namelijk altijd in verhouding tot de vrijheden die zij elders hebben. Dit betekent dat de bijdrage van sociaal kapitaal aan deze capaciteiten beter inzichtelijk wordt wanneer ook sociaal kapitaal relatief wordt gemeten. Door sociaal kapitaal uit te drukken als het aandeel sociaal kapitaal dat iemand binnen het eigen dorp heeft, wordt duidelijker hoe dit kapitaal bijdraagt aan de mate van positieve en negatieve vrijheid binnen en

buiten het eigen dorp. Daarmee wordt inzichtelijker in hoeverre iemand daadwerkelijk de capaciteit heeft om te vertrekken of juist te blijven.

Na deze theoretische onderbouwing voor het gebruik van relatieve maten voor sociaal kapitaal wordt het ACF nu concreet toegepast om te onderbouwen wat de verwachte rol van sociaal kapitaal bij vertrekintentie is.

Toegepast op aspiraties (aspirations) is het aannemelijk dat mensen van wie het sociaal kapitaal sterker binnen het eigen dorp gericht is meer kansen zien om hun wensen binnen het dorp te vervullen en minder kansen om die daarbuiten te vervullen. De verwachte functie van algemene wensen in het leven en veronderstelde kansen om deze buiten het eigen dorp te vervullen is dus lager als het sociaal kapitaal relatief meer binnen het eigen dorp gericht is. Hierdoor zal de verwachte vertrekintentie lager zijn als sociaal kapitaal sterker binnen het eigen dorp gericht is.

Andersom geldt dat mensen van wie het sociaal kapitaal meer buiten het eigen dorp gericht is meer kansen zien om hun wensen buiten het dorp te vervullen en minder kans om deze binnen het dorp te vervullen. De verwachte functie van algemene wensen in het leven en veronderstelde kansen om deze buiten het eigen dorp te vervullen is dus hoger als het sociaal kapitaal relatief meer buiten het eigen dorp gericht is. Hierdoor zal de verwachte vertrekintentie hoger zijn als het sociaal kapitaal relatief meer buiten het eigen dorp gericht is.

Naast aspiraties spelen ook capaciteiten (capabilities) een rol. Deze capaciteiten worden, zoals in 2.1.1 besproken, beïnvloed door twee factoren: positieve en negatieve vrijheid.

Voor positieve vrijheid is relevant dat relatief meer sociaal kapitaal binnen het eigen dorp in lijn met Putnam (1995) relatief meer voordelen oplevert binnen het dorp. Denk bijvoorbeeld aan hulp bij huisvesting, sociale steun of de juiste contacten voor werkgelegenheid. Deze voordelen zorgen ervoor dat de persoon meer positieve vrijheid heeft om de controle over het eigen leven te nemen en zijn fundamentele doelen te behalen als hij in het huidige dorp blijft wonen. Aan de andere kant zorgt relatief meer sociaal kapitaal binnen het dorp juist voor minder voordelen en dus positieve vrijheid buiten het dorp. Hoe meer sociaal kapitaal binnen het eigen dorp gericht is, hoe hoger dus de positieve vrijheid binnen het dorp en hoe lager de positieve vrijheid buiten het dorp, en daarmee hoe lager de verwachte vertrekintentie.

Andersom geldt dat relatief meer sociaal kapitaal buiten het dorp ervoor zorgt dat mensen minder van deze voordelen en daarmee positieve vrijheid kunnen verkrijgen in hun dorp. Deze groep kan juist meer voordelen en daarmee positieve vrijheid verkrijgen buiten hun dorp. Hoe meer sociaal kapitaal buiten het eigen dorp gericht is, hoe lager dus de positieve vrijheid binnen het dorp en hoe hoger de positieve vrijheid buiten het dorp, en daarmee hoe hoger de verwachte vertrekintentie.

Voor de toepassing van sociaal kapitaal op negatieve vrijheid is het relevant dat sociaal kapitaal sociale druk kan uitoefenen op individuele keuzes. Volgens Coleman (1988) beïnvloeden de normen binnen sociaal kapitaal de keuzes van mensen: sommige keuzes worden gestimuleerd en andere keuzes juist beperkt. Omdat sociaal kapitaal voordelen oplevert, is het aannemelijk dat mensen binnen het dorp sociale druk uitoefenen op dorpsgenoten die deel uitmaken van

hun sociaal kapitaal om hen te overtuigen om te blijven. Zij behalen immers voordelen uit het sociaal kapitaal en dit kapitaal is makkelijker te benutten als de persoon in het dorp blijft wonen. Anderzijds is het aannemelijk dat sociaal kapitaal buiten het dorp juist sociale druk oplevert om het dorp te verlaten omdat deze mensen juist meer kunnen profiteren als de persoon dichterbij hun nabijheid komt wonen.

Voor negatieve vrijheid geldt dus dat relatief meer sociaal kapitaal binnen het eigen dorp zorgt voor meer sociale druk om te blijven. Het sociaal kapitaal vormt dan meer een obstakel om te vertrekken. Anderzijds ontvangt deze groep relatief minder sociale druk die hen stimuleert om te vertrekken uit het dorp. Het sociaal kapitaal vormt dan minder een obstakel om te blijven. De negatieve vrijheid om te vertrekken is dus lager en de negatieve vrijheid om te blijven is hoger. Daarmee is de verwachte vertrekintentie dus lager.

Relatief meer sociaal kapitaal buiten het dorp zorgt juist voor relatief minder sociale druk om te blijven. Het sociaal kapitaal vormt dan minder een obstakel om te vertrekken. Aan de andere kant zorgt relatief meer sociaal kapitaal buiten het eigen dorp juist voor meer sociale druk die stimuleert om te vertrekken. Het sociaal kapitaal vormt dan meer een obstakel om in het dorp te blijven. De negatieve vrijheid om te vertrekken is dus hoger, de negatieve vrijheid om te blijven is lager en de verwachte vertrekintentie is daardoor ook hoger.

De capaciteiten tot vertrek zijn daarmee lager voor mensen met relatief meer sociaal kapitaal binnen het eigen dorp doordat zij meer positieve vrijheid hebben binnen en minder positieve vrijheid hebben buiten het dorp, doordat zij minder negatieve vrijheid hebben om te vertrekken en doordat zij meer negatieve vrijheid hebben om te blijven. De capaciteiten tot vertrek zijn juist hoger bij mensen met relatief meer sociaal kapitaal buiten het eigen dorp doordat zij minder positieve vrijheid hebben binnen en meer positieve vrijheid hebben buiten het dorp, doordat zij meer negatieve vrijheid hebben om te vertrekken en doordat zij minder negatieve vrijheid hebben om te blijven. Zoals besproken versterkt de capaciteit volgens het ACF ook de aspiratie: als men minder de capaciteit heeft om te vertrekken, heeft men ook minder die aspiratie (de Haas, 2021).

Op basis van deze toepassing van het ACF kunnen hypothesen worden opgesteld voor deelvraag 3. De derde deelvraag luidt: ‘Welke rol speelt sociaal kapitaal bij de vertrekken uit het eigen dorp?’. Het toepassen van het ACF op de rol van sociaal kapitaal ondersteunt dat mensen met relatief meer sociaal kapitaal binnen het dorp eerder de aspiratie en capaciteit hebben om in het dorp te blijven. Mensen met relatief meer sociaal kapitaal buiten het dorp zullen juist eerder de aspiratie en capaciteit hebben om het dorp te verlaten.

Dit is in lijn met de bestaande inzichten. Uit de bestaande inzichten kwam naar voren dat het hebben van meer lokaal sociaal kapitaal de kans verkleint dat een individu die regio verlaat (David, Janiak & Wasmer, 2010). Het hebben van sociaal kapitaal buiten de woonplek vergroot de kans dat de persoon wil verhuizen naar de plek waar deze sociaal kapitaal heeft opgebouwd (Büchel et al., 2020).

Dit ondersteunt dat er een negatief verband is tussen relatief meer sociaal kapitaal binnen het eigen dorp en vertrekken, en een positief verband tussen relatief meer sociaal kapitaal buiten het eigen dorp en vertrekken. In dit onderzoek wordt vertrekintentie gebruikt als voorspeller van de daadwerkelijke vertrekken. Hieruit volgt de volgende hypothese:

H3: Relatief meer sociaal kapitaal binnen het eigen dorp hangt samen met een lagere vertrekintentie, terwijl relatief meer sociaal kapitaal buiten het eigen dorp samenhangt met een hogere vertrekintentie.

De vierde deelvraag, die bij het eerste hiaat hoort, is: 'In hoeverre medieert de locatie van sociaal kapitaal de relatie tussen leeftijd en vertrekintentie uit dorpen in niet-stedelijke gebieden?'. Door voort te bouwen op de inzichten en theorie die de basis vormen voor hypothesen 1, 2 en 3 gezamenlijk kan een onderbouwde verwachting voor deze deelvraag worden ontleend. In hypothese 3 is uitgebreid besproken hoe de relatieve locatie van sociaal kapitaal een rol speelt bij vertrekintentie. Deze redenering vormt het uitgangspunt om voor deelvraag 4 te verkennen in hoeverre leeftijdsverschillen in de relatieve locatie van sociaal kapitaal een verklaring bieden voor de hogere vertrekintentie onder jongere mensen.

In hypothese 1 wordt verwacht dat jongere mensen een hogere vertrekintentie hebben dan oudere mensen. Hypothese 2 stelt dat jongere mensen relatief minder sociaal kapitaal binnen het eigen dorp bezitten en juist relatief meer sociaal kapitaal buiten het dorp dan oudere mensen. Voor hypothese 3 is onderbouwd dat relatief meer sociaal kapitaal binnen het dorp samenhangt met een lagere vertrekintentie terwijl relatief meer sociaal kapitaal buiten het dorp samenhangt met een hogere vertrekintentie. Deze werking van het ACF wordt nu toegepast op de hogere vertrekintentie van jongere mensen om een verwachting voor deelvraag 4 af te leiden.

Binnen het ACF wordt de beslissing om wel of niet te migreren beschouwd als een functie van de aspiratie en capaciteit om wel of niet te migreren (de Haas, 2021). Deze aspiratie en capaciteit wordt volgens de onderbouwing van hypothese 3 beïnvloed door de locatie waar het sociaal kapitaal van een persoon zich bevindt. Om het ACF op deelvraag 4 toe te passen, is het relevant dat jongere mensen op basis van hypothese 2 naar verwachting relatief minder sociaal kapitaal binnen het eigen dorp bezitten en relatief meer buiten het eigen dorp.

Toegepast op de aspiratie om te vertrekken geldt dat jongere mensen hierdoor minder kansen zien om hun wensen in hun eigen dorp te vervullen. Doordat zij juist relatief meer sociaal kapitaal buiten het dorp bezitten zien zij meer kansen om die wensen elders te vervullen. De aspiratie van jongere mensen om uit hun dorp te vertrekken is daarom naar verwachting hoger.

Ook hun capaciteiten om te vertrekken worden beïnvloed doordat zij naar verwachting relatief minder sociaal kapitaal binnen hun eigen dorp hebben. De capaciteit is, zoals in paragraaf 2.1.1 besproken, afhankelijk van positieve en negatieve vrijheid (de Haas, 2021).

Voor positieve vrijheid is relevant dat sociaal kapitaal voordelen oplevert (Putnam, 1995). Jongere mensen met minder sociaal kapitaal binnen het dorp verkrijgen in hun dorp minder van deze voordelen en hebben daardoor minder positieve vrijheid om in hun eigen dorp de controle over hun leven te nemen en hun fundamentele doelen te realiseren. Tegelijkertijd hebben zij relatief meer sociaal kapitaal buiten het dorp wat hun positieve vrijheid vergroot om deze doelen juist buiten het dorp te realiseren.

Daarnaast speelt negatieve vrijheid een rol bij de capaciteit om te blijven of vertrekken. Hiervoor is relevant dat sociale druk invloed heeft op individuele keuzes (Coleman, 1988). Dit kan dus een obstakel vormen om te blijven of te vertrekken en beïnvloedt dus de negatieve vrijheid. Als jongere mensen weinig sociale druk ontvangen om te blijven doordat zij relatief minder sociaal kapitaal bezitten binnen hun dorp vormt dit minder een obstakel om uit het dorp te vertrekken. Anderzijds ontvangen zij juist meer sociale druk om te vertrekken doordat zij

relatief meer sociaal kapitaal buiten het dorp hebben. Dit vormt juist een obstakel om te blijven. De negatieve vrijheid om te blijven is voor jongere mensen dus lager en de negatieve vrijheid om te vertrekken hoger. De capaciteiten om te blijven zijn in lijn met voorgaande bij jongere mensen naar verwachting dus lager en de capaciteiten om te vertrekken hoger dan bij oudere mensen.

De beslissing om wel of niet te migreren is volgens het ACF een functie van zowel de aspiratie als de capaciteit om te vertrekken (de Haas, 2021). Jongere mensen hebben dus naar verwachting een lagere aspiratie en capaciteit om te blijven en een hogere aspiratie en capaciteit om te vertrekken dan oudere mensen. Daarmee is hun verwachte vertrekkans hoger. Deze hogere vertrekkans wordt, zoals hierboven onderbouwd, naar verwachting (deels) verklaard doordat jongere mensen relatief minder sociaal kapitaal binnen en relatief meer sociaal kapitaal buiten het dorp hebben. Er is dan sprake van mediatie van het verband tussen leeftijd en vertrekkans door de relatieve locatie van sociaal kapitaal. In dit onderzoek wordt vertrekintentie gebruikt als voorspeller van de daadwerkelijke vertrekkans. Hieruit volgt de volgende hypothese:

H4: De hogere vertrekintentie van jongere mensen wordt (deels) verklaard doordat zij relatief minder sociaal kapitaal binnen en relatief meer sociaal kapitaal buiten het eigen dorp bezitten.

2.2. Sociaalemotionele selectiviteitstheorie

Als alternatief voor de besproken mediatieverwachting in hypothese 4 wordt in deze paragraaf een ander theoretisch perspectief gepresenteerd. Hierbij wordt niet uitgegaan van een verklarende rol van sociaal kapitaal maar van een modererende werking van leeftijd op het verband tussen sociaal kapitaal en vertrekintentie. Voor deze benadering wordt de sociaalemotionele selectiviteitstheorie van Carstensen (1995) toegelicht. Deze theorie wordt vervolgens toegepast op het ACF van de Haas (2021) om te onderbouwen hoe leeftijd de werking van sociaal kapitaal binnen dit framework beïnvloedt. Op basis daarvan wordt een hypothese opgesteld voor deelvraag 5.

2.2.1. De sociaalemotionele selectiviteitstheorie

De sociaalemotionele selectiviteitstheorie van Carstensen (1995) is een levenslooptheorie die beschrijft hoe sociale doelen van mensen veranderen naarmate zij ouder worden. Deze veranderingen worden verklaard doordat oudere mensen hun tijdshorizon doorgaans als beperkter ervaren dan jongere mensen. De emotionele behoeften van mensen worden groter naarmate ze ouder worden. Het opgebouwde sociaal kapitaal speelt voor oudere mensen dan ook een erg belangrijke rol. Volgens de theorie gaat ouder worden daarnaast gepaard met een afnemende behoefte aan informatiezoekend gedrag waarbij men de wereld ontdekt, onafhankelijker wordt van het bestaande sociale kapitaal van familie en jeugdvriendschappen en juist nieuwe sociale contacten zoekt.

Jongere mensen zijn dus minder afhankelijk van bestaand sociaal kapitaal en bouwen eerder nieuw sociaal kapitaal op, terwijl oudere mensen afhankelijker zijn van hun bestaande kapitaal en minder snel nieuw sociaal kapitaal opbouwen.

2.2.2. Toepassing van de sociaalemotionele selectiviteitstheorie

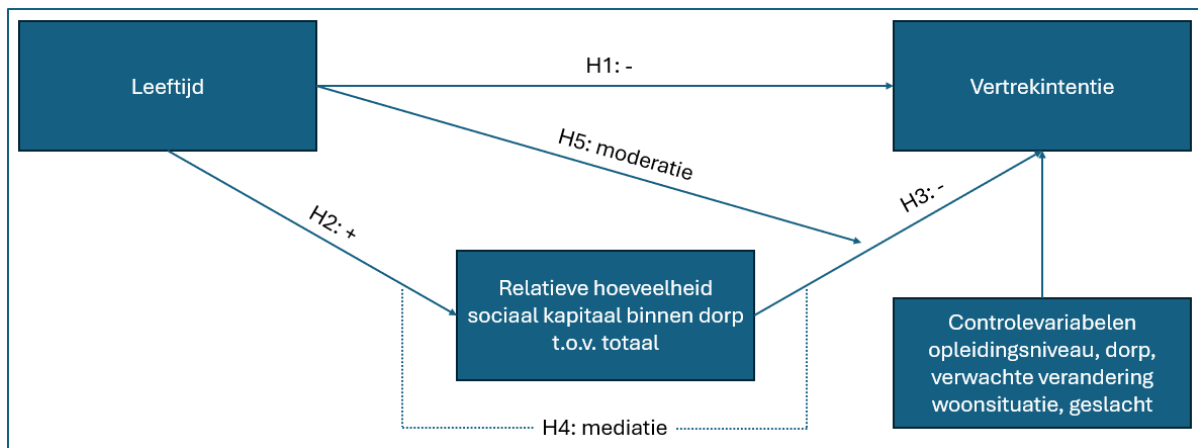
Als het bestaande sociaal kapitaal een belangrijkere rol gaat spelen bij oudere mensen, dan is het in het licht van het ACF van de Haas (2021) ook aannemelijk dat het bestaande sociaal kapitaal voor oudere mensen een sterkere invloed heeft op de vertrekintentie dan voor jongere mensen. Het bestaande sociaal kapitaal en de locatie daarvan speelt ten eerste een sterkere rol bij hun mogelijke aspiratie om te vertrekken doordat het benutten van dit sociaal kapitaal voor hen een sterkere wens in het leven is. Volgens het ACF is de aspiratie namelijk een functie van de algemene wensen die een persoon in het leven heeft en de veronderstelde geografische kansenstructuren. Doordat het benutten van het bestaande sociaal kapitaal een belangrijkere wens is het leven wordt, is de invloed van sociaal kapitaal op de aspiratie om wel of niet te willen vertrekken sterker bij oudere mensen.

Ook speelt de leeftijd een rol bij de capaciteiten die men heeft om te vertrekken. Doordat het bestaande sociaal kapitaal belangrijker is voor oudere mensen zal de locatie van het sociaal kapitaal voor hen meer invloed hebben op de vraag of ze op de huidige plek controle over hun leven kunnen nemen en hun fundamentele doelen kunnen verwezenlijken. Als mensen relatief meer sociaal kapitaal binnen hun dorp bezitten, verkrijgen ze hieruit meer van deze positieve vrijheid in het dorp en minder positieve vrijheid buiten het dorp. Als ze juist meer sociaal kapitaal buiten hun dorp bezitten, verkrijgen ze juist minder positieve vrijheid in het dorp en meer positieve vrijheid buiten het dorp om daar de doelen te verwezenlijken. Qua negatieve vrijheid zorgt relatief meer sociaal kapitaal binnen het dorp voor minder negatieve vrijheid om te vertrekken en meer om te blijven. Relatief meer sociaal kapitaal buiten het dorp zorgt juist voor relatief meer negatieve vrijheid om te vertrekken en minder om te blijven. Doordat het bestaande sociaal kapitaal voor oudere mensen een belangrijkere rol speelt dan voor jongere mensen, is het aannemelijk dat deze gevolgen voor de positieve en negatieve vrijheid door sociaal kapitaal sterker doorwerken in de capaciteit van oudere mensen om te vertrekken.

De sociaalemotionele selectiviteitstheorie in combinatie met de toepassing van het ACF ondersteunt dat de leeftijd het negatieve effect van de relatieve hoeveelheid sociaal kapitaal binnen het eigen dorp op vertrekkans modereert. Verwacht wordt dat het effect sterker wordt bij een hogere leeftijd doordat oudere mensen meer belang hechten aan hun bestaande sociaal kapitaal, en dat kapitaal bevindt zich dicht bij hun huidige woonplek. Anderzijds ondersteunt het ook dat leeftijd het positieve effect van sociaal kapitaal buiten het eigen dorp op de vertrekintentie modereert: verwacht wordt dat het effect sterker wordt bij een hogere leeftijd doordat oudere mensen meer belang hechten aan hun bestaande sociaal kapitaal, en dat kapitaal bevindt zich verder weg van hun huidige woonplek. In dit onderzoek wordt vertrekintentie gebruikt als voorspeller van de daadwerkelijke vertrekkans. Hieruit volgt de volgende hypothese:

H5: De relatie tussen sociaal kapitaal en vertrekintentie wordt gemodereerd door leeftijd, waarbij de effecten van sociaal kapitaal binnen en buiten het eigen dorp op vertrekintentie sterker worden naarmate de leeftijd hoger is.

In figuur 1 zijn de opgestelde hypothesen gevisualiseerd in een conceptueel model.



Figuur 1
Conceptueel model

3. Methodologie

In dit hoofdstuk wordt toegelicht hoe het onderzoek is uitgevoerd. In paragraaf 3.1 wordt de dataverzameling besproken. In paragraaf 3.2 wordt uitgelegd hoe de theoretische concepten zijn geoperationaliseerd. Paragraaf 3.3 bespreekt hoe wordt omgegaan met ontbrekende waarden. In paragraaf 3.4 wordt tot slot het analyseplan besproken.

3.1. Dataverzameling

De data zijn door mij verzameld in de periode van 6 juni 2024 tot en met 4 juli 2024. Er is hiervoor een willekeurige steekproef van in totaal 3750 inwoners van Damwâld, De Westereen en Feanwâlden van 16 jaar of ouder getrokken. Op de peildatum 1 juni 2024 woonden er in totaal 11976 inwoners van 16 jaar of ouder in een van deze dorpen.

De vragenlijst voor dit onderzoek heb ik onafhankelijk samengesteld. De gemeente Dantumadiel heeft toestemming gegeven voor de dataverzameling en heeft het verzenden van uitnodigingen gefaciliteerd. De inwoners in de steekproef zijn per brief uitgenodigd om deel te nemen aan het onderzoek. De envelop bevatte de uitnodigingsbrief in het Nederlands en Fries met een deelnamecode om online deel te nemen in de Nederlandse of de Friese taal. Verder zat in de envelop de papieren vragenlijst in het Nederlands en een informatieformulier over het onderzoek in het Nederlands. Op het informatieformulier is verwezen naar een Friese vertaling van het document.

Uitgenodigden konden dus online in het Nederlands en het Fries deelnemen via een link naar de enquêtetool van Onderzoekdoen.nl. Ook konden zij deelnemen door de bijgevoegde Nederlandstalige papieren vragenlijst in te vullen en in te leveren bij de brievenbus van het gemeentehuis in Damwâld of bij een van de Buurtkamers in Damwâld, De Westereen en Feanwâlden. Deze papieren vragenlijsten zijn vervolgens ingevoerd in de enquêtetool.

De vragenlijst bestond uit meerkeuzevragenlijstitems en vragenlijstitems waarbij respondenten een getal moesten invullen. In totaal hebben 361 mensen deelgenomen aan het onderzoek. Het responspercentage is dus 9,63%

Eén persoon heeft echter slechts vragenlijstitems over twee van de negen variabelen beantwoord, namelijk die over controlevariabele woonplaats en het vragenlijstitem over leeftijd. Imputatie van de ontbrekende waarden zou dan plaatsvinden op basis van zeer beperkte gegevens. Dit zou de kwaliteit van de imputatie en daarmee de betrouwbaarheid van de resultaten verzwakken. Daarom wordt deze persoon uitgesloten van het onderzoek.

Daarnaast zijn er drie respondenten die op alle acht vragenlijstitems over structureel sociaal kapitaal 0 hebben geantwoord. Bij hen is het niet mogelijk een relatieve maat voor structureel sociaal kapitaal binnen het eigen dorp te berekenen. Omdat deze variabele een cruciale rol speelt in de analyses worden ook deze drie respondenten uitgesloten.

Tot slot is er één persoon die op zeven van de acht vragenlijstitems over structureel sociaal kapitaal heeft geantwoord met een 0 en op het achtste vragenlijstitem niets heeft ingevuld. In de geïmputeerde datasets kwam deze persoon vaak alsnog op alle acht vragenlijstitems uit op een score van 0. Omdat hierdoor geen zinvolle maat voor relatief structureel sociaal kapitaal binnen het eigen dorp kan worden berekend, wordt ook deze persoon uitgesloten.

Er blijven daarmee 356 respondenten over die worden meegenomen in de analyses.

3.2. Operationalisatie van theoretische concepten

In deze paragraaf wordt besproken welke theoretische concepten centraal staan in dit onderzoek en hoe deze concepten meetbaar zijn gemaakt en geoperationaliseerd tot variabelen op basis van vragenlijstitems. Eerst vind je in tabel 1 een opsomming van alle vragenlijstitems die zijn gebruikt om de theoretische concepten te meten en zo tot variabelen te komen. Per vragenlijstitem is aangegeven of het een vraag of een stelling betreft. Daarna wordt de operationalisatie tot variabele besproken per theoretisch concept. Bij de variabelen ‘relatief cognitief sociaal kapitaal binnen dorp’, ‘relatief structureel sociaal kapitaal binnen dorp’, ‘(verwacht) opleidingsniveau’ en ‘verwachte verandering woonsituatie’ wordt imputatie toegepast (zie paragraaf 3.4.1). Voor deze variabelen wordt ook besproken of deze imputatie plaatsvindt vóór of na het hercoderen van de vragenlijstitems naar de variabelen.

Tabel 1

Vragenlijstitems en antwoordopties per theoretisch concept

Vertrekintentie	
Inleidende tekst: Geef een score tussen 0 en 10. Een score van 0 betekent dat u het zeker niet wilt, 10 betekent dat u het zeker wel wilt.	
Vraag: Wilt u binnen twee jaar verhuizen uit het dorp waar u nu woont?	Schaal met 0= zeker niet t/m 10= zeker wel
Leeftijd	
Vraag: Wat is uw leeftijd?	_____ jaar
Cognitief sociaal kapitaal binnen dorp	
Inleidende tekst: Geef aan in hoeverre u het eens bent met deze stellingen over mensen die wel in uw dorp wonen:	
Stelling: De meeste van mijn dorpsgenoten zijn te vertrouwen	Likert 5 opties (helemaal niet mee eens t/m helemaal mee eens)
Stelling: De meeste van mijn dorpsgenoten proberen meestal eerlijk te zijn	Likert 5 opties (helemaal niet mee eens t/m helemaal mee eens)
Stelling: De meeste van mijn dorpsgenoten zouden van je profiteren als ze de kans kregen	Likert 5 opties (helemaal niet mee eens t/m helemaal mee eens)
Stelling: Je kunt niet te voorzichtig zijn bij het omgaan met mijn dorpsgenoten	Likert 5 opties (helemaal niet mee eens t/m helemaal mee eens)
Cognitief sociaal kapitaal buiten dorp	
Inleidende tekst: Geef aan in hoeverre u het eens bent met deze stellingen over mensen die niet in uw dorp wonen:	
Stelling: De meeste mensen die geen dorpsgenoten van mij zijn, zijn te vertrouwen	Likert 5 opties (helemaal niet mee eens t/m helemaal mee eens)
Stelling: De meeste mensen die geen dorpsgenoten van mij zijn, proberen meestal eerlijk te zijn	Likert 5 opties (helemaal niet mee eens t/m helemaal mee eens)
Stelling: De meeste mensen die geen dorpsgenoten van mij zijn, zouden van je profiteren als ze de kans kregen	Likert 5 opties (helemaal niet mee eens t/m helemaal mee eens)
Stelling: Je kunt niet te voorzichtig zijn bij het omgaan met mensen die geen dorpsgenoten van mij zijn	Likert 5 opties (helemaal niet mee eens t/m helemaal mee eens)

Structureel sociaal kapitaal (zowel binnen als buiten dorp)	
Inleidende tekst: Geef aan hoe vaak dit in een gemiddelde maand van 30 dagen voorkomt. Het gaat niet om het aantal dagen waarin het voorkomt, maar om het aantal keren. Het kan dus voorkomen dat u iets vaker dan 30 keer per maand doet.	
Vraag: Hoe vaak spreekt u in een gemiddelde maand van 30 dagen af met vrienden of kennissen die in hetzelfde dorp als uzelf wonen? <i>(geen huisgenoten)</i>	_____ keer per maand
Vraag: Hoe vaak spreekt u in een gemiddelde maand van 30 dagen af met familieleden die in hetzelfde dorp als uzelf wonen? <i>(geen huisgenoten)</i>	_____ keer per maand
Vraag: Hoe vaak bent u in een gemiddelde maand van 30 dagen actief bij verenigingen in het dorp waar u woont (bijvoorbeeld sportverenigingen, hobbyclubs etc.)?	_____ keer per maand
Vraag: Hoe vaak bezoekt u in een gemiddelde maand van 30 dagen een kerk en/of andere religieuze activiteiten in het dorp waar u woont?	_____ keer per maand
Vraag: Hoe vaak spreekt u in een gemiddelde maand van 30 dagen af met vrienden of kennissen die niet in hetzelfde dorp als uzelf wonen? <i>(geen huisgenoten)</i>	_____ keer per maand
Vraag: Hoe vaak spreekt u in een gemiddelde maand van 30 dagen af met familieleden die niet in hetzelfde dorp als uzelf wonen? <i>(geen huisgenoten)</i>	_____ keer per maand
Vraag: Hoe vaak bent u in een gemiddelde maand van 30 dagen actief bij verenigingen buiten het dorp waar u woont (bijvoorbeeld sportverenigingen, hobbyclubs etc.)?	_____ keer per maand
Vraag: Hoe vaak bezoekt u in een gemiddelde maand van 30 dagen een kerk en/of andere religieuze activiteiten buiten het dorp waar u woont?	_____ keer per maand
(Verwacht) opleidingsniveau	
Vraag: Welk van deze opties is op u van toepassing? <i>(bijscholing of cursussen vanuit werk worden niet gezien als opleiding of school)</i>	Ik studeer niet meer; Ik studeer of ga naar school; Ik studeer nu niet, maar begin in de komende 12 maanden aan een opleiding of school
Alleen mensen die nog (gaan) studeren Vraag: Op welk niveau volgt u een opleiding of gaat u een opleiding volgen?	Praktijkonderwijs; Vmbo; Havo of vwo; Mbo-1; Mbo-2, mbo-3, mbo-4; Hbo-opleiding; Universitaire opleiding
Vraag: Wat is uw hoogst behaalde opleidingsniveau?	Basisonderwijs; Praktijkonderwijs; Vmbo, havo-onderbouw, vwo-onderbouw, mbo-1 (inclusief mavo, ulo, mulo, lts, lhno, Vbo); Havo-bovenbouw, vwo-bovenbouw (inclusief hbs, mms); Mbo-2, mbo-3, mbo-4; Hbo-opleiding; Universitaire opleiding
Woonplaats	
Vraag: Wat is uw woonplaats?	Damwâld; Feanwâlden; De Westereen
Verwachte verandering woonsituatie	
Vraag: Wat is uw woonsituatie?	Ik woon alleen; Ik woon met mijn partner en kind(eren); Ik woon met kind(eren), zonder partner; Ik woon met mijn partner, zonder kinderen; Ik woon met mijn ouder(s)/verzorger(s); Ik heb een andere woonsituatie

Vraag: Welke woonsituatie verwacht u dat over twee jaar voor u geldt?	Ik woon alleen; Ik woon met mijn partner en kind(eren); Ik woon met kind(eren), zonder partner; Ik woon met mijn partner, zonder kinderen; Ik woon met mijn ouder(s)/verzorger(s); Ik heb een andere woonsituatie
Geslacht	
Vraag: Wat is uw geslacht?	Man; Vrouw; Anders; Zeg ik liever niet

Vertrekintentie

De interesse in dit onderzoek gaat specifiek uit naar het vertrek uit het eigen dorp. Zoals besproken wordt vertrekintentie binnen twee jaar gebruikt als voorspeller voor vertrek uit het dorp. Aan respondenten is dus gevraagd om aan te geven in hoeverre zij de intentie hebben om binnen 2 jaar te verhuizen uit het dorp waar zij nu wonen. Het antwoord op dit vragenlijstitem op een schaal van 0 (zeker niet) tot en met 10 (zeker wel) is de waarde voor de variabele 'vertrekintentie'.

Dit vragenlijstitem is grotendeels hetzelfde opgesteld als door de Groot, Manting & Mulder (2011). Omdat in dit onderzoek specifiek wordt onderzocht of mensen uit hun dorp willen verhuizen en niet of ze willen verhuizen in het algemeen is het vragenlijstitem licht aangepast op deze context. Aangezien de sterkte van de vertrekintentie invloed blijkt te hebben op de mate waarin mensen deze intentie in actie omzetten, zijn de antwoordopties aangepast ten opzichte van de antwoordopties die de Groot, Manting en Mulder (2011) gebruikten. Om de respondenten de mogelijkheid te geven duidelijk onderscheid te maken in de sterkte van hun vertrekintentie is gebruik gemaakt van een 11-punt likertschaal. Volgens Alwin (1997) zorgt een groter aantal schaalpunten ervoor dat een variabele preciezer gemeten wordt. Ook kan de variabele hiermee als continu behandeld worden (Hodge & Gillespie, 2007). Het gebruik van een groter aantal antwoordopties om preciezer te meten en de variabele als continu te kunnen behandelen, is zowel empirisch als met simulaties onderbouwd (Wu & Leung, 2017).

De validiteit van deze operationalisatie is over het algemeen hoog doordat het vragenlijstitem expliciet betrekking heeft op vertrekintentie binnen een duidelijke tijdshorizon van twee jaar en doordat er specifiek wordt benoemd dat het om vertrek uit het eigen dorp gaat. Hierdoor dekt deze operationalisatie het theoretische concept nauwkeurig. Doordat er gebruik is gemaakt van een 11-punt likertschaal wordt de sterkte van deze vertrekintentie nog beter in kaart gebracht dan door de Groot, Manting en Mulder (2011) die minder antwoordopties hanteerden. De validiteit wordt wel beperkt doordat er slechts één algemene vraag wordt gesteld terwijl de daadwerkelijke kans op vertrek nog concreter kan worden blootgelegd door gedragscomponenten zoals actief zoeken naar een woning mee te nemen. Een andere beperkende factor voor de validiteit is dat vertrekintentie is gebruikt als voorspeller van vertrek omdat het niet mogelijk was om tweemaal te toetsen en dus daadwerkelijk vertrek te meten.

De betrouwbaarheid van deze operationalisatie is redelijk. De formulering van het vragenlijstitem is gebaseerd op eerder onderzoek van De Groot, Manting en Mulder (2011) waarin het effectief is toegepast in de Nederlandse context. Dit vergroot de kans dat respondenten de vraag op consistente wijze interpreteren. Dit versterkt de betrouwbaarheid. Wel kan het antwoord beïnvloed worden door persoonlijke omstandigheden of stemmingen op het moment van invullen, zoals een recente negatieve of juist positieve ervaring in het dorp. Dit maakt de meting gevoelig voor contextuele invloeden. Dit is een beperkende factor voor de betrouwbaarheid.

Leeftijd

Om de leeftijd te meten is respondenten gevraagd om hun leeftijd in jaren aan te geven. Het antwoord op dit vragenlijstitem is de waarde voor de variabele 'leeftijd'. Leeftijd is in dit onderzoek direct en eenduidig gemeten met een open invulveld in jaren en meet dus precies én uitsluitend het onderliggende theoretische concept leeftijd. De validiteit is hoog omdat de vraag exact aansluit bij het bedoelde concept zonder interpretatieruimte. Ook de betrouwbaarheid is hoog aangezien leeftijd een objectief kenmerk waarbij het ingevulde antwoord bij herhaalde meting niet verandert.

Relatief cognitief sociaal kapitaal binnen dorp en relatief structureel sociaal kapitaal binnen dorp

Om sociaal kapitaal te meten wordt de vragenlijst van Ejlskov et al. (2014) gebruikt. In deze vragenlijst wordt het theoretische concept sociaal kapitaal opgedeeld in cognitief sociaal kapitaal en structureel sociaal kapitaal. In de vragenlijst meten de onderzoekers het sociale kapitaal ongeacht of dat sociaal kapitaal zich binnen of buiten het eigen dorp bevindt. Het verschil tussen die benadering en de benadering in dit onderzoek is het gegeven dat in dit onderzoek sociaal kapitaal binnen en buiten het eigen dorp los van elkaar wordt gemeten. Daarom wordt elk vragenlijstitem zowel met betrekking tot sociaal kapitaal binnen het eigen dorp als met betrekking tot sociaal kapitaal buiten het eigen dorp gesteld. Verder zijn een aantal kleine aanpassingen doorgevoerd die de vraagmodule passender maken voor dit onderzoek zonder aan te vergaren informatie in te boeten. In totaal zijn er 16 vragenlijstitems opgenomen over sociaal kapitaal: vier over cognitief sociaal kapitaal binnen het eigen dorp, vier over cognitief sociaal kapitaal buiten het eigen dorp, vier over structureel sociaal kapitaal binnen het eigen dorp en vier over structureel sociaal kapitaal buiten het eigen dorp.

In dit onderzoek worden de 16 vragenlijstitems over sociaal kapitaal samengevoegd tot twee schalen om de theoretische concepten 'relatief cognitief sociaal kapitaal binnen dorp' en 'relatief structureel sociaal kapitaal binnen dorp' te meten. Deze schalen worden in de analyse van alle modellen allebei opgenomen waardoor er in de resultaten onderscheid te maken is tussen de rol van cognitief en de rol van structureel sociaal kapitaal.

Cognitief sociaal kapitaal wordt gemeten aan de hand van stellingen over de mate waarin men dorpsgenoten of niet-dorpsgenoten vertrouwt, hen eerlijk vindt, denkt dat zij van je zouden profiteren als ze de kans kregen en of men voorzichtig denkt te moeten zijn in contact met hen. Deze items brengen in kaart hoe individuen hun sociale omgeving beoordelen in termen van betrouwbaarheid en intenties van anderen. Daarmee sluiten ze aan bij de definitie van cognitief sociaal kapitaal door Ejlskov et al. (2014) als vertrouwen in anderen en percepties van delen en wederkerigheid.

Voor cognitief sociaal kapitaal worden de antwoordopties van de vragenlijstitems over profiteren van anderen en voorzichtig zijn met anderen vóór de imputatie omgezet zodat de score 0 betekent dat men het helemaal eens is met de stelling en 4 betekent dat men het er helemaal niet mee eens is. De vragenlijstitems over of anderen te vertrouwen zijn en of anderen eerlijk zijn, zijn gecodeerd zodat 0 staat voor helemaal niet mee eens en 4 staat voor helemaal mee eens. Hierdoor betekent een hogere score voor deze vragenlijstitems dat men positiever is over anderen.

Omdat de vier vragenlijstitems geacht worden één onderliggend construct te meten wordt Cronbach's alpha berekend vóór imputatie om de interne consistentie te toetsen. Dit gebeurt apart voor de vragenlijstitems over cognitief sociaal kapitaal binnen het eigen dorp en die over

cognitief sociaal kapitaal buiten het eigen dorp. Een score van 0,7 of hoger wordt volgens Civelek (2018) gezien als een indicatie van goede interne consistentie.

Voor cognitief sociaal kapitaal binnen het eigen dorp bedraagt Cronbach's alpha 0,80, wat wijst op een sterke samenhang tussen de vragenlijstitems. De vragenlijstitems over cognitief sociaal kapitaal buiten het eigen dorp laten een vergelijkbare samenhang zien met een Cronbach's alpha van 0,82. In beide gevallen verhoogt het verwijderen van een vragenlijstitem de waarde van Cronbach's alpha niet. Dit duidt erop dat alle vier vragenlijstitems een waardevolle bijdrage leveren aan de meting van het onderliggende construct en dat er in beide gevallen dus sprake is van interne consistentie.

Hoewel de oorspronkelijke vragenlijstitems zijn gemeten op ordinaal niveau via Likert-schalen onderbouwen Carifio en Perla (2008) dat samengestelde schalen bestaande uit meerdere samenhangende vragenlijstitems met Likert-schalen als intervalniveau mogen worden behandeld. Om de vragenlijstitems te kunnen samenvoegen tot een schaal is het belangrijk dat ze vergelijkbare antwoordopties en richting hebben. Zo voorkom je dat sommige vragenlijstitems zwaarder meewegen dan anderen of dat hoge scores iets negatiefs betekenen bij het ene vragenlijstitem en iets positiefs bij een andere. Alle vragenlijstitems waren al op een schaal van 0 tot en met 4 gemeten, maar zijn qua richting zoals besproken geharmoniseerd naar een schaal van 0 (negatief) tot en met 4 (positief). De hoge Cronbach's alpha scores van 0,80 (binnen dorp) en 0,82 (buiten dorp) ondersteunen verder het toepassen van deze benadering aangezien deze scores wijzen op een sterke interne consistentie tussen de vragenlijstitems. Omdat beide subschalen dus als intervalniveau mogen worden toegepast is het verantwoord om een relatieve schaal te berekenen op basis van hun onderlinge verhouding.

Om deze relatieve schaal te berekenen worden na imputatie op itemniveau (zie paragraaf 3.4.1) de waarden van de vragenlijstitems over cognitief sociaal kapitaal binnen het eigen dorp opgeteld en gedeeld door de som van de waarden van de vragenlijstitems over cognitief sociaal kapitaal binnen en buiten het dorp. De schaalscore kan dus variëren van 0% (alleen cognitief sociaal kapitaal buiten eigen dorp) tot 100% (alleen cognitief sociaal kapitaal binnen eigen dorp). Op deze manier wordt de relatieve individuele oriëntatie binnen of buiten het dorp blootgelegd.

Op basis van het bovenstaande kan worden geconcludeerd dat de operationalisatie van cognitief sociaal kapitaal goed aansluit bij het theoretisch concept zoals afgebakend in paragraaf 1.2 en dat er dus sprake is van hoge validiteit. De vragenlijstitems sluiten inhoudelijk namelijk goed aan op de definitie van Ejlskov et al. (2014). Het concept wordt gemeten op individueel niveau via percepties die het cognitieve sociaal kapitaal van het individu weerspiegelen. De vragenlijstitems richten zich op percepties ten aanzien van mensen binnen en buiten het dorp, waarmee de locatiecomponent expliciet tot uiting komt. Wel gaat het om algemene inschattingen van bijvoorbeeld vertrouwen in dorpsgenoten, los van specifieke contexten of situaties. Daarmee wordt slechts één dimensie van het theoretische concept gemeten. Dit beperkt de validiteit enigszins.

De hoge interne consistentie ondersteunt dat de vragenlijstitems gezamenlijk één theoretisch construct meten en dat de schaal dus over het algemeen betrouwbaar is. Wel kan het antwoord op de vragen beïnvloed worden door tijdelijke contextuele invloeden. Denk hierbij bijvoorbeeld aan een recente diefstal door een dorpsgenoot waardoor de persoon negatiever over dorpsgenoten is dan op andere momenten. Dit verlaagt de betrouwbaarheid van de metingen.

Structureel sociaal kapitaal wordt gemeten met vragenlijstitems over hoe vaak men contact heeft met vrienden en kennissen, hoe vaak men contact heeft met familie, hoe vaak men actief

is bij verenigingen en hoe vaak men religieuze activiteiten bezoekt, zowel binnen als buiten het dorp. Alle vragenlijstitems hebben betrekking op de frequentie per gemiddelde maand van 30 dagen. Deze vragenlijstitems geven inzicht in de dichtheid van het sociale netwerk en hoe actief de persoon is in of juist buiten het dorp. Daarmee sluiten ze aan bij de definitie van structureel sociaal kapitaal door Ejlskov et al. (2014) die dit omschrijven als de dichtheid van iemands sociale netwerk en zijn maatschappelijke betrokkenheid.

De vragenlijstitems over structureel sociaal kapitaal gaan dus over verschillende soorten activiteiten zoals contact met vrienden en activiteit bij verenigingen. Deze activiteiten verschillen inhoudelijk en hoeven niet sterk samen te hangen maar vormen samen wel het totale structureel sociaal kapitaal. Omdat de vragenlijstitems dus niet allemaal hetzelfde meten, wordt hier geen betrouwbaarheidscoëfficiënt uitgerekend.

In enkele gevallen zijn extreem hoge aantallen van 90 keer of meer per maand op een vragenlijstitem ingevuld. Dit komt neer op gemiddeld drie keer per dag. Dergelijke waarden worden als onrealistisch beschouwd en daarom vóór de analyse als missing genoteerd. Ze worden later geïmputeerd (zie paragraaf 3.4.1) om de kwaliteit van de data te waarborgen zonder onnodige uitsluiting van respondenten.

De vragenlijstitems betreffen frequenties en zijn dus gemeten op intervalniveau. Het optellen van de waarden van de vragenlijstitems is verantwoord doordat ze allemaal zijn gemeten als frequenties per gemiddelde maand van 30 dagen. Hierdoor zijn de waarden onderling vergelijkbaar en kan een relatieve schaal worden berekend.

Na imputatie op itemniveau (zie paragraaf 3.4.1) worden de antwoorden op de afzonderlijke vragenlijstitems over structureel sociaal kapitaal hiervoor samengevoegd tot één schaal. Dit gebeurt door de waarden voor de vragenlijstitems over structureel sociaal kapitaal binnen het eigen dorp op te tellen en te delen door de totale som van de scores op structureel sociaal kapitaal binnen en buiten het eigen dorp. Deze schaalscore kan variëren van 0% (alleen structureel sociaal kapitaal buiten eigen dorp) tot 100% (alleen structureel sociaal kapitaal binnen eigen dorp) en geeft daarmee inzicht in de relatieve oriëntatie van het structureel sociaal kapitaal binnen of buiten het dorp.

Ook voor structureel sociaal kapitaal kan op basis van bovenstaande worden geconcludeerd dat de operationalisatie goed aansluit bij het theoretische concept zoals deze is afgebakend in paragraaf 1.2 en dat er dus over het algemeen sprake is van hoge validiteit. De vragenlijstitems sluiten inhoudelijk namelijk goed aan op de definitie van Ejlskov et al. (2014). De vragenlijstitems zijn daarnaast op individueel niveau gemeten via de aantallen keren dat het individu per gemiddelde maand contact heeft of deelneemt aan activiteiten. Ook is er sprake van onderscheid tussen structureel sociaal kapitaal binnen en buiten het dorp waardoor de locatiecomponent tot uiting komt. Wel richt de operationalisatie zich uitsluitend op de frequentie en niet op de kwaliteit of betekenis van deze sociale interacties. Hierdoor wordt slechts één dimensie van het theoretisch concept gemeten. Dit beperkt validiteit enigszins.

De betrouwbaarheid van de meting is over het algemeen hoog. Deze wordt versterkt doordat de antwoorden niet afhankelijk zijn van contextuele invloeden. Er wordt namelijk gevraagd om het aantal keren dat iemand iets doet en dit aantal is niet afhankelijk van bijvoorbeeld de stemming van de respondent. Wel kan het zijn dat mensen een verkeerde inschatting maken van de frequentie per gemiddelde maand. Doordat de relatieve maat van structureel sociaal kapitaal binnen het eigen dorp wordt berekend, levert dit geen problemen voor de betrouwbaarheid op als mensen zowel de frequenties binnen als buiten het eigen dorp in dezelfde mate onderschatten of overschatten. Als respondenten bijvoorbeeld al hun

frequenties structureel iets te laag inschatten, blijft de verhouding correct. Als zij bijvoorbeeld alleen de frequenties buiten het eigen dorp overschatten zal dit de relatieve maat vertekenen. In zulke gevallen wordt de betrouwbaarheid wel aangetast.

Om de interne validiteit van de uitkomsten te vergroten en daardoor met meer zekerheid causaliteit vast te kunnen stellen, wordt in dit onderzoek gecontroleerd voor de volgende variabelen: (verwacht) opleidingsniveau, woonplaats, verwachte verandering woonsituatie en geslacht.

(Verwacht) opleidingsniveau

De eerste controlevariabele is het (verwachte) opleidingsniveau van de respondent. Aan respondenten is gevraagd wat hun hoogst behaalde opleidingsniveau is. Ook is gevraagd of zij nog studeren of dit binnen een jaar verwachten (weer) te gaan doen. Aan de mensen die nog (gaan) studeren is gevraagd op welk opleidingsniveau deze opleiding wordt gevolgd.

De antwoorden worden gecategoriseerd in laagopgeleid, middelbaar opgeleid en hoogopgeleid aan de hand van de Standaard Onderwijsindeling 2021 van het Centraal Bureau voor de Statistiek (z.d.-g).

Bij het vragenlijstitem wat het hoogste afgeronde opleidingsniveau is vallen de antwoordopties 'Basisonderwijs', 'Praktijkonderwijs' en 'Vmbo, havo-onderbouw, vwo-onderbouw, mbo-1 (inclusief mavo, ulo, mulo, lts, lhno, Vbo)' onder laagopgeleid. 'Havo-bovenbouw, vwo-bovenbouw (inclusief hbs, mms)' en 'Mbo 2, mbo 3, mbo 4' vallen onder middelbaar opgeleid. 'Hbo-opleiding' en 'Universitaire opleiding' vallen onder hoogopgeleid.

Bij het vragenlijstitem voor respondenten die nog (gaan) studeren vallen de antwoordopties 'Praktijkonderwijs' en 'Mbo-1' onder laagopgeleid. 'Vmbo' en 'Mbo 2, mbo 3, mbo 4' vallen onder middelbaar opgeleid. 'Havo of vwo', 'Hbo-opleiding' en 'Universitaire opleiding' vallen onder hoogopgeleid.

Als respondenten nog (gaan) studeren telt het hoogste opleidingsniveau van deze twee vragenlijstitems. Elke respondent heeft dus één waarde voor de variabele '(verwacht) opleidingsniveau': laagopgeleid (gecodeerd als 0), middelbaar opgeleid (gecodeerd als 1) of hoogopgeleid (gecodeerd als 2). Deze hercodering vindt plaats vóór de imputatie die op de variabele (Verwacht) opleidingsniveau wordt toegepast. In paragraaf 3.4.1 vind je hier een uitleg over.

Met deze vragenlijstitems wordt het onderliggende theoretische concept (verwacht) opleidingsniveau zo nauwkeurig mogelijk benaderd. De validiteit van de operationalisatie is daardoor over het algemeen hoog. Het opleidingsniveau van mensen die niet meer studeren en ook niet van plan zijn om te gaan studeren wordt ten eerste nauwkeurig gemeten. Bij respondenten die momenteel studeren, wordt het verwachte opleidingsniveau zo nauwkeurig mogelijk ingeschat op basis van het niveau van de opleiding die zij daadwerkelijk volgen. Hun uiteindelijke opleidingsniveau wordt daarmee niet bepaald op basis van hun eigen inschatting maar op basis van een concreet gegeven. Bij respondenten die verwachten te gaan studeren, wordt het verwachte opleidingsniveau afgeleid uit het niveau waarop zij deze opleiding gaan volgen. Deze inschatting is minder valide omdat deze gebaseerd is op een verwachting en een verwachting is gevoeliger voor subjectieve of tijdelijke invloeden. Hierdoor zal in sommige gevallen een verschil met de daadwerkelijke uitkomst ontstaan. Dit beperkt de validiteit enigszins.

De betrouwbaarheid van de operationalisatie is over het algemeen hoog. Bij respondenten met een afgeronde opleiding is dit antwoord een vaststaand feit dat niet verandert bij een herhaalde meting. Ook bij respondenten die nog studeren is de betrouwbaarheid hoog omdat

het om concrete opleidingsdeelname gaat. Het antwoord is dus niet afhankelijk van bijvoorbeeld de stemming van de respondent. Bij respondenten die binnen een jaar verwachten weer te gaan studeren is de betrouwbaarheid iets lager doordat het opleidingsniveau bij hen is gebaseerd op het niveau waarop zij een opleiding verwachten te gaan volgen. Deze verwachting is gevoeliger voor contextuele invloeden op het moment van deelname aan het onderzoek. Denk bijvoorbeeld aan een recent gesprek met iemand die de opleiding volgt en daar negatief over is. Hierdoor kan de respondent twijfelen of hij deze opleiding wil volgen en bij herhaalde metingen verschillende antwoorden geven.

Woonplaats

De tweede controlevariabele is de woonplaats waar de respondent woont. Respondenten konden namelijk woonachtig zijn in de drie dorpen Damwâld (gecodeerd als 1), De Westereen (gecodeerd als 2) en Feanwâlden (gecodeerd als 3). Aan de respondenten is gevraagd in welk dorp zij wonen. Respondenten die niet in één van deze drie dorpen woonden zijn uit de vragenlijst geleid. Het antwoord op dit vragenlijstitem is de waarde van variabele 'woonplaats'.

De woonplaats is gemeten via een vragenlijstitem met de drie woonplaatsen waaruit de steekproef is getrokken als antwoordopties. Hierdoor sluit de meting goed aan bij het theoretische concept. De validiteit is daardoor over het algemeen hoog. Bij respondenten met meerdere woonplekken kan twijfel ontstaan over welke woonplaats bedoeld wordt. De uitnodiging is verstuurd naar het adres waar men staat ingeschreven en het is waarschijnlijk dat respondenten de vragenlijst daar invullen. Het is daardoor waarschijnlijk dat deze officiële woonplaats ook als antwoord is opgegeven. Tegelijkertijd is het mogelijk dat dit adres op het moment van invullen niet de feitelijke hoofdverblijfplaats is. Denk hierbij bijvoorbeeld aan personen met meerdere verblijfplaatsen. Dit beperkt de validiteit in geringe mate.

Ook de betrouwbaarheid is over het algemeen hoog omdat de woonplaats in veruit de meeste gevallen een vaststaand gegeven is dat bij herhaalde metingen hetzelfde wordt ingevuld. Bij respondenten met meerdere verblijfplaatsen kan het antwoord in theorie per moment verschillen. Doordat de brief wordt gestuurd naar het adres waar iemand is ingeschreven en de vragenlijst daar waarschijnlijk wordt ingevuld, is de kans echter groot dat bij herhaalde meting steeds die woonplaats als antwoord zou worden ingevuld. De betrouwbaarheid wordt daarom slechts in geringe mate beperkt.

Verwachte verandering woonsituatie

De derde controlevariabele is het al dan niet verwachten van een verandering in de woonsituatie in de komende twee jaren. Hiervoor worden respondenten gecategoriseerd op basis van de vragenlijstitems 'Wat is uw woonsituatie?' en 'Welke woonsituatie verwacht u dat over twee jaar voor u geldt?'. De antwoordopties bij beide vragenlijstitems waren 'Ik woon alleen', 'Ik woon met mijn partner en kind(eren)', 'Ik woon met kind(eren), zonder partner', 'Ik woon met mijn partner, zonder kinderen', 'Ik woon met mijn ouder(s)/verzorger(s)' en 'Ik heb een andere woonsituatie'. Als het antwoord over de huidige woonsituatie hetzelfde is als het antwoord op het vragenlijstitem over de verwachte woonsituatie wordt er geen verandering verwacht. Als er verschillende antwoorden zijn gegeven wordt er wel een verandering verwacht. Hieruit volgt de variabele 'verwachte verandering woonsituatie', een categorische variabele met geen verwachte verandering (gecodeerd als 1) en wel een verwachte verandering (gecodeerd als 2) als mogelijke waarden. Deze variabele wordt gecodeerd na de imputaties van ontbrekende waarden van de twee vragenlijstitems die hieraan ten grondslag liggen. In paragraaf 3.4.1 vind je hier een uitleg over.

In de vragenlijstitems wordt gevraagd naar de huidige woonsituatie en de verwachte woonsituatie over twee jaar. Deze vragenlijstitems hebben dezelfde antwoordopties. Hieruit wordt beoordeeld of deze waarden verschillen. De operationalisatie richt zich dus niet op het soort verandering maar puur op het wel of niet verwachten ervan.

De validiteit van deze variabele is redelijk. De meting sluit goed aan bij het theoretische concept omdat wordt vastgesteld of iemand wel of geen verandering in woonsituatie verwacht. De huidige woonsituatie is daarbij in veruit de meeste gevallen een eenduidig gegeven. De verwachte woonsituatie is echter gebaseerd op een persoonlijke inschatting van de situatie over twee jaar. Niet alle respondenten kunnen die toekomst nauwkeurig inschatten. Hierdoor wordt het concept niet altijd exact gemeten. Dit beperkt de validiteit.

Ook de betrouwbaarheid van de meting is redelijk. De huidige woonsituatie is in veruit de meeste gevallen een eenduidig gegeven dat niet afhankelijk is van contextuele invloeden en daardoor bij herhaalde meting hetzelfde antwoord oplevert. Tegelijkertijd is de variabele mede gebaseerd op een toekomstverwachting die bij herhaalde meting kan veranderen door contextuele invloed. Het antwoord van iemand die met een kind woont kan bijvoorbeeld per moment verschillen als dat kind veel twijfelt om wel of niet uit huis te gaan. Daardoor bestaat de kans dat dezelfde persoon bij herhaalde meting een andere inschatting maakt. Dit beperkt de betrouwbaarheid van de meting.

Geslacht

De vierde controlevariabele is het geslacht van de respondent. Bij de dataverzameling is respondenten gevraagd om aan te geven wat hun geslacht is. Hun antwoord is de waarde voor variabele 'geslacht'. De antwoordopties zijn man (gecodeerd als 1), vrouw (gecodeerd als 2), anders (gecodeerd als 3) en zeg ik liever niet (gecodeerd als 4).

Geslacht is in dit onderzoek direct en eenduidig gemeten met verschillende opties voor het geslacht als antwoordopties en meet dus goed het onderliggende theoretische concept. De validiteit van deze meting is daardoor hoog. De optie 'zeg ik liever niet' vormt daarop een uitzondering: in die gevallen wordt geen invulling aan het concept gegeven en is geslacht dus niet meetbaar. Deze antwoordoptie is echter slechts door twee respondenten gekozen en als aparte categorie meegenomen in de analyses. De invloed op de validiteit van de variabele als geheel is daardoor verwaarloosbaar.

Ook de betrouwbaarheid van deze meting is hoog. Geslacht is voor veruit de meeste mensen een stabiel kenmerk dat niet afhankelijk is van contextuele invloeden en bij herhaalde meting hetzelfde antwoord oplevert.

3.3. Ontbrekende waarden

In de dataset komt het voor dat respondenten niet alle vragenlijstitems hebben beantwoord. Ook zijn enkele zeer extreme waarden bij structureel sociaal kapitaal als ontbrekend gecodeerd (zie paragraaf 3.2). Van de 356 respondenten zijn bij 305 alle vragenlijstitems ingevuld die in dit onderzoek worden gebruikt. Vanwege de ontbrekende waarden wordt *MICE* (Multivariate Imputation by Chained Equations) gebruikt. *MICE* is een pakket in programmeertaal R waarin men multiële imputatie kan toepassen (van Buuren & Groothuis-Oudshoorn, 2011). Multiële imputatie wordt toegepast aan de hand van drie stappen.

Allereerst worden meerdere datasets gemaakt. Hierin worden de ontbrekende waarden vervangen door plausibele schattingen op basis van een door de gebruiker gespecificeerd model waarmee waarden voor andere variabelen worden gebruikt om ontbrekende waarden te schatten (van Buuren & Groothuis-Oudshoorn, 2011). Dit resulteert in meerdere datasets waarin

de geobserveerde waarden gelijk blijven, maar waarbij er verschillen zijn in de invulling van de ontbrekende waarden. Op basis van deze datasets worden in stap 2 de analyses uitgevoerd en ontstaan dus meerdere resultaten. Na het analyseren poolt men deze resultaten in stap 3 tot één uitkomst. Het gebruik van imputatie zorgt er hiermee voor dat de analyses kunnen worden gebaseerd op zoveel mogelijk respondenten en dat dit zo min mogelijk ten koste gaat van de betrouwbaarheid.

Imputeren gebeurt met een vastgesteld aantal iteraties. Een iteratie houdt in dat voor elke variabele met ontbrekende waarden een regressiemodel wordt geschat op basis van de overige variabelen (Azur et al., 2011). Hierna worden deze modellen gebruikt om de ontbrekende waarden in te vullen. Eén cyclus door alle variabelen vormt één iteratie. Dit proces wordt *chained equations* genoemd omdat elke variabele met ontbrekende waarden iteratief wordt geschat op basis van de overige variabelen. Bij elke iteratie worden de imputaties geüpdatet: de ontbrekende waarden worden telkens opnieuw geschat op basis van de geobserveerde én de meest recente geïmputeerde waarden.

Bij het toepassen van multiële imputatie wordt uitgegaan van de assumptie ‘Missing at random’. Dit betekent dat het ontbreken van gegevens verklaard kan worden aan de hand van andere geobserveerde variabelen en niet afhankelijk is van de ontbrekende waarde zelf (van Buuren, 2018). Er zal in de plausibiliteitstoetsing (zie paragraaf 3.4.2), bij de univariate statistieken (zie paragraaf 3.4.3) en bij de bivariate statistieken (zie paragraaf 3.4.4) worden uitgelegd hoe zal worden onderzocht of deze assumptie niet wordt geschonden.

Het imputeren van de ontbrekende waarden wordt verschillend uitgevoerd afhankelijk van het type vragenlijstitem waarvan er waarden ontbreken. De methode wordt gekozen op basis van het meetniveau van dat vragenlijstitem. Voor dit onderzoek zijn de methoden voor continue, voor ordinale en voor categorische vragenlijstitems met meer dan 2 antwoordopties nodig. Deze methoden worden nu besproken.

Imputatiemethode continue vragenlijstitems

Het R-pakket *MICE* implementeert predictive mean matching via de functie *mice.impute.pmm()* als de standaard imputatiemethode voor ontbrekende waarden bij vragenlijstitems met continue antwoordopties (van Buuren & Groothuis-Oudshoorn, 2011). Predictive mean matching is een semi-parametrische imputatiemethode. Bij predictive mean matching wordt een groep van vergelijkbare, complete gevallen (kandidaatdonoren) geselecteerd met voorspelde waarden die het dichtst liggen bij de voorspelde waarde van de incomplete casus (van Buuren, 2018). Eén donor wordt willekeurig gekozen en deze geobserveerde waarde wordt gebruikt om de ontbrekende waarde te vervangen. Hierbij wordt verondersteld dat de verdeling van de ontbrekende waarde overeenkomt met die van de geselecteerde kandidaatdonoren. Een belangrijk voordeel van deze methode is dat de imputaties alleen plaatsvinden op basis van geobserveerde waarden (van Buuren & Groothuis-Oudshoorn, 2011). Dit geldt alleen voor de daadwerkelijke imputaties. Het onderliggende model waarmee de kandidaatdonoren gekozen worden maakt zoals gebruikelijk in het chained equation proces wel gebruik van zowel de geobserveerde waarden als van de eerder geïmputeerde waarden (Azur et al., 2011). Daarnaast kan de methode niet-lineaire verbanden behouden, zelfs wanneer het structurele deel van het imputatiemodel niet correct is (van Buuren & Groothuis-Oudshoorn, 2011).

Imputatiemethode ordinale vragenlijstitems met meer dan twee antwoordopties

Voor de imputatie van ordinale variabelen maakt het R-pakket *MICE* gebruik van het ordered logit model, dat ook bekendstaat als het proportional odds model. (van Buuren & Groothuis-

Oudshoorn, 2011). Dit model is een ordinale variant van logistische regressie (Liu, Bai & Koirala, 2018). In plaats van afzonderlijke kansen voor elke categorie te modelleren, schat het model de cumulatieve kans dat een observatie in of juist lager dan een categorie valt. *MICE* maakt voor het toepassen van deze methode gebruik van de *mice.impute.polr()* functie van een ander pakket in R genaamd *MASS* (van Buuren & Groothuis-Oudshoorn, 2011). Deze methode is ontworpen voor vragenlijstitems met meer dan twee antwoordopties met een ordinale volgorde.

Imputatiemethode categorische vragenlijstitems met meer dan twee antwoordopties

Voor categorische vragenlijstitems met meer dan twee antwoordopties wordt het multinomiale logit model geïmplementeerd (van Buuren & Groothuis-Oudshoorn, 2011). Het multinomiale logitmodel schat daarbij, ten opzichte van een referentiecategorie, de kans op elke antwoordoptie op basis van de andere variabelen in de dataset (Kwak & Clayton-Matthews, 2002). De schatting gebeurt via maximum likelihood. De uitkomsten zijn log-odds die worden omgerekend naar odds ratio's. Op basis daarvan kunnen de kansen per categorie worden berekend. *MICE* maakt voor de toepassing hiervan gebruik van de *multinom()* functie van een ander pakket in R genaamd *nnet* en gebruikt dit als basis voor de functie *mice.impute.polyreg()* (van Buuren & Groothuis-Oudshoorn, 2011). Deze methode is ontworpen voor vragenlijstitems met meer dan twee antwoordopties zonder logische volgorde.

3.4. Analyseplan

Zoals besproken worden de analyses voor dit onderzoek uitgevoerd in het R-pakket *MICE*. In dit analyseplan wordt toegelicht welke analyses er concreet worden toegepast en hoe deze zullen worden uitgevoerd. Hiervoor wordt eerst in 3.4.1 besproken hoe ontbrekende waarden concreet geïmputeerd zullen worden. In 3.4.2 wordt daarna besproken hoe de kwaliteit van imputaties zal worden geëvalueerd. Vervolgens worden in 3.4.3 (univariate statistieken) en 3.4.4 (bivariate statistieken) besproken welke beschrijvende analyses zullen worden uitgevoerd. Daarna worden in 3.4.5 per hypothese de modellen besproken. Tot slot wordt in 3.4.6 besproken hoe de invloed van imputaties wordt geëvalueerd en wordt in 3.4.7 het plan van aanpak voor de modevaluatie besproken.

3.4.1. Imputeren van ontbrekende waarden

De eerste stap is zoals besproken het imputeren van ontbrekende waarden (van Buuren & Groothuis-Oudshoorn, 2011). Hiervoor moet worden vastgesteld hoeveel iteraties er worden uitgevoerd en hoeveel geïmputeerde datasets er moeten worden gegenereerd. Volgens van Buuren & Groothuis-Oudshoorn (2011) is 10 tot 20 iteraties over het algemeen voldoende om tot convergentie te komen. In dit onderzoek wordt 20 iteraties aangehouden om meer zekerheid te hebben dat convergentie bereikt wordt.

Een vuistregel die wordt aangehouden voor het gewenste aantal geïmputeerde datasets is dat deze minimaal gelijk is aan het percentage cases dat minimaal één ontbrekende waarde heeft (White, Royston & Wood, 2010). In dit onderzoek hebben 356 respondenten deelgenomen. Bij 305 hiervan zijn alle vragenlijstitems die relevant zijn voor dit onderzoek ingevuld en bij 51 niet. Het percentage cases met minimaal één ontbrekende waarde is 14,33%. Er zijn dan minimaal 15 geïmputeerde datasets nodig. In dit onderzoek is ervoor gekozen om iets boven het minimum te zitten en worden 20 geïmputeerde datasets gegenereerd.

Bij kleine tot middelgrote datasets waarin 20 tot 30 variabelen zijn opgenomen, wordt aangeraden om al deze variabelen mee te nemen in de imputatiemodellen (van Buuren, 2018).

In dit onderzoek zijn 28 variabelen relevant. 18 zijn vragenlijstitems die ten grondslag liggen aan de eindvariabelen relatief cognitief sociaal kapitaal binnen dorp, relatief structureel sociaal kapitaal binnen dorp en verwachte verandering woonsituatie. De overige 10 zijn variabelen die direct in de analysemodellen worden opgenomen als afhankelijke variabele, predictor, interactieterm en/of controlevariabele. In lijn met van Buuren (2018) is ervoor gekozen om al deze variabelen te gebruiken als predictoren voor de variabelen waarbij ontbrekende waarden moeten worden geïmputeerd.

Niet alle vragenlijstitems bevatten ontbrekende waarden. De vragenlijstitems voor de variabelen leeftijd, vertrekintentie, geslacht en woonplaats zijn namelijk door alle respondenten beantwoord en bevatten dus geen ontbrekende waarden. De overige vragenlijstitems bevatten wel ontbrekende waarden. De ontbrekende waarden voor deze vragenlijstitems zullen worden geïmputeerd.

Imputatie variabelen sociaal kapitaal

Voor de variabelen over sociaal kapitaal geldt dat de imputaties zullen worden gedaan vóór het samenvoegen van de verschillende vragenlijstitems tot de schalen relatief cognitief sociaal kapitaal binnen dorp en relatief structureel sociaal kapitaal binnen dorp. De ontbrekende waarden op de individuele vragenlijstitems worden dus geïmputeerd. Hierna worden binnen de *MICE* omgeving met deze volledig ingevulde waarden de ontbrekende waarden van de eindschalen relatief cognitief sociaal kapitaal binnen dorp en relatief structureel sociaal kapitaal binnen dorp ingevuld.

De vragenlijstitems die bij de variabele relatief cognitief sociaal kapitaal horen, zijn gesteld met de antwoordopties helemaal niet mee eens tot en met helemaal mee eens. Dit zijn dus vragenlijstitems met ordinale antwoordopties. Imputatie vindt in deze gevallen zoals besproken plaats aan de hand van het ordered logit model dat is geïmplementeerd in het R-pakket *MICE* (van Buuren & Groothuis-Oudshoorn, 2011).

De vragenlijstitems die bij de variabele relatief structureel sociaal kapitaal horen, zijn gesteld met een open veld waarin men het aantal keer per maand kon invullen. Dit zijn dus vragenlijstitems met continue antwoordopties. Imputatie vindt in deze gevallen zoals besproken plaats aan de hand van de implementatie van predictive mean matching in het R-pakket *MICE* (van Buuren & Groothuis-Oudshoorn, 2011).

Imputatie variabele (verwacht) opleidingsniveau

De imputatie van het (verwachte) opleidingsniveau vindt plaats op de nieuwe samengestelde variabele waarin het hoogste niveau is vastgesteld: het reeds behaalde opleidingsniveau of, indien van toepassing, het niveau waarop respondenten nog studeren of verwachten te gaan studeren. Deze samengestelde variabele (verwacht) opleidingsniveau bevat alleen een ontbrekende waarde wanneer beide vragenlijstitems niet zijn beantwoord. De reden om te imputeren op (verwacht) opleidingsniveau is dat ontbrekende waarden voor het vragenlijstitem op welk niveau respondenten nog (gaan) studeren niet willekeurig zijn omdat dit vragenlijstitem enkel is getoond aan mensen die nog (gaan) studeren.

De imputatie vindt dus plaats op basis van het hoogste (verwachte) opleidingsniveau met als antwoordopties laagopgeleid, middelbaar opgeleid en hoogopgeleid. Dit zijn dus ordinale antwoordopties. Imputatie vindt daarom zoals besproken plaats aan de hand van het ordered logit model dat is geïmplementeerd in het R-pakket *MICE* (van Buuren & Groothuis-Oudshoorn, 2011).

Imputatie huidige woonsituatie en verwachte woonsituatie

Ook voor de variabele verwachte verandering woonsituatie is imputatie nodig. De twee vragenlijstitems die voor deze variabele zijn gesteld zijn, in tegenstelling tot (verwacht) opleidingsniveau, wel aan alle respondenten gesteld. Daarom vindt imputatie plaats bij beide vragenlijstitems. De ontbrekende waarden van de twee individuele vragenlijstitems worden dus geïmputeerd. Daarna worden binnen de *MICE* omgeving met deze volledig ingevulde waarden de ontbrekende waarden van de eindvariabele verwachte verandering woonsituatie ingevuld.

Beide vragenlijstitems voor deze variabele hebben zes categorische antwoordopties zonder logische volgorde. In deze gevallen vindt imputatie zoals besproken plaats aan de hand van het multinomiale logit model dat is geïmplementeerd in het R-pakket *MICE* (van Buuren & Groothuis-Oudshoorn, 2011).

3.4.2. Kwaliteit van imputaties

Om te beoordelen of de imputaties geschikt zijn voor verdere analyse wordt de kwaliteit geëvalueerd op basis van de convergentie en de plausibiliteit van de imputaties.

Convergentie

Volgens van Buuren & Groothuis-Oudshoorn (2011) is het allereerst belangrijk om convergentie van de imputaties te toetsen. Convergentie van de imputaties zal worden beoordeeld via visualisatie van de imputatiestromen met het *plot.mids* object. In deze plots wordt het gemiddelde en de standaarddeviatie van de geïmputeerde waarden per iteratie getoond waarbij elke keten staat voor de ontwikkeling over iteraties per geïmputeerde dataset. Er zijn dus in dit onderzoek 20 ketens waarvan de ontwikkeling over de 20 iteraties getoond wordt. Er is sprake van goede convergentie wanneer de variantie tussen de verschillende ketens niet groter is dan de variantie binnen elke individuele keten. Dit betekent dat de ketens door elkaar moeten lopen zonder duidelijke trend.

Plausibiliteit van de imputaties

Een andere belangrijke stap bij imputatie is het evalueren van de plausibiliteit van de imputaties. Imputaties zouden namelijk waarden moeten bevatten die ook zouden kunnen worden gevonden als de waarden niet ontbraken (van Buuren & Groothuis-Oudshoorn, 2011). Dit zal in dit onderzoek worden geëvalueerd via de *MICE*-functie *stripplot()*. In deze plots worden de geobserveerde en de geïmputeerde waarden getoond per geïmputeerde dataset nadat de 20 iteraties hebben plaatsgevonden. Dit geeft inzicht in de geobserveerde en geïmputeerde waarden van alle 20 geïmputeerde datasets. Er zal visueel worden geëvalueerd of er geen onmogelijke waarden zijn voortgekomen uit het imputatieproces. Negatieve waarden zijn bijvoorbeeld niet mogelijk. Daarnaast kan op deze manier worden geëvalueerd in hoeverre de univariate verdelingen van de geobserveerde en geïmputeerde waarden overeenkomen. Hiermee kan een eventuele schending van de assumptie ‘Missing at random’ worden blootgelegd.

3.4.3. Univariate statistieken voor en na imputatie

Om een inzicht te krijgen in de data zullen eerst de univariate statistieken worden geanalyseerd voor de variabelen die worden meegenomen in dit onderzoek. Allereerst gebeurt dit op basis van de dataset vóór de imputatie. Daarnaast worden de univariate statistieken getoond na imputatie. Voor het gemiddelde, de minimum- en maximumwaarden en de verdelingen wordt

het gemiddelde genomen over de 20 geïmputeerde datasets. Voor de standaarddeviatie wordt eerst de variantie per geïmputeerde dataset berekend. Hierna wordt het gemiddelde van deze varianties genomen. Hieruit wordt de wortel getrokken om de juiste standaarddeviatie te verkrijgen.

Door de univariate statistieken zowel voor als na imputatie te tonen, wordt inzicht verkregen in het patroon van ontbrekende waarden. Dit vergroot het inzicht in de mate waarin is voldaan aan de assumptie dat de ontbrekende data ‘Missing at random’ is. Ook geeft het inzicht in de impact die de imputaties hebben gehad op de univariate statistieken.

3.4.4. Bivariate statistieken voor en na imputatie

Ook worden bivariate statistieken onderzocht voor de variabelen in dit onderzoek. Dit gebeurt allereerst ook op basis van de dataset vóór imputatie. Ook de bivariate statistieken worden daarnaast getoond na imputatie. Dit gebeurt wederom door de gemiddelde toetsscores van de 20 geïmputeerde datasets te tonen. In dit geval worden de bivariate statistieken op basis van de geïmputeerde dataset gegenereerd en daarna samengevoegd tot één geaggregeerde gemiddelde uitkomst. De bivariate statistieken vóór en na imputatie vergroten het inzicht in de mate waarin is voldaan aan de assumptie van ‘Missing at random’. Daarnaast geeft het inzicht in de impact van de imputaties op de bivariate statistieken.

Voor relaties tussen twee continue of ordinale variabelen is de Spearman-correlatie (ρ) gebruikt. Hierbij zijn voor de statistieken na imputatie de correlaties per imputatie omgezet naar Fisher’s z . Daarvan is het gemiddelde genomen en is de correlatiemaat teruggetransformeerd naar ρ . Deze toetsscores worden gerapporteerd als ρ .

Voor relaties tussen een continue of ordinale variabele en een categorische variabele is de correlatie afgeleid uit een ANOVA F-toets. Hierbij is de wortel van de gemiddelde verklaarde variantie (R^2) genomen als maat voor de sterkte van de correlatie. Voor de statistieken na imputatie is de wortel van de gemiddelde R^2 van alle imputaties genomen. De toetsscores worden gerapporteerd als $\sqrt{R^2}$.

Voor relaties tussen twee categorische variabelen is Cramér’s V berekend als maat voor de correlatie. Voor de statistieken na imputatie wordt de gemiddelde V over de 20 imputaties berekend. Deze toetsscores worden gerapporteerd als V .

Voor de statistieken na imputatie kunnen geen p -waarden worden gerapporteerd omdat deze statistieken niet gepaard gaan met standaardfouten waardoor significantietoetsing via Rubin’s rules niet mogelijk is (Austin, Eekhout & van Buuren, 2024). Een mogelijke alternatieve methode, het gebruik van de mediane p -waarde, leidt tot een sterk verhoogde kans op type I-fouten en wordt om die reden door deze auteurs afgeraden. Voor de interpretatie zal daarom het verschil in sterkte van de samenhang voor en na imputatie worden besproken.

3.4.5. Modellen en plan van aanpak in MICE

Dan de daadwerkelijke modellen die worden gebruikt om de hypothesen te beantwoorden. In stap 2 en 3 van de analyse met het R-pakket *MICE* worden de uitkomsten van de modellen op basis van de geïmputeerde datasets afzonderlijk gegenereerd met de functie *with()* en daarna aan de hand van Rubin’s rules gepoold tot één geaggregeerde uitkomst met de functie *pool()* (van Buuren & Groothuis-Oudshoorn, 2011). In tabel 2 vind je de modellen die per hypothese zijn opgesteld. Bij elke hypothese is het doel van de modellen benoemd en wordt besproken welke variabelen er in de modellen zijn opgenomen. Bij de onafhankelijke variabelen zijn de relevante voorspellers **dikgedrukt** en de controlevariabelen niet.

Tabel 2*Modellen bij hypothesen, doel en variabelen*

Hypothese	Model	Doel	Afhankelijke variabele	Onafhankelijke variabelen
H1: Hogere leeftijd hangt samen met een lagere vertrekintentie uit dorpen in niet-stedelijke gebieden.	Model 1	Vaststellen of leeftijd een significant negatief effect heeft op vertrekintentie	Vertrekintentie	Leeftijd , (verwacht) opleidingsniveau, woonplaats, verwachte verandering woonsituatie en geslacht
H2: Hogere leeftijd hangt samen met relatief meer sociaal kapitaal binnen het eigen dorp en relatief minder sociaal kapitaal buiten het eigen dorp.	Model 2a en 2b	Vaststellen of leeftijd een significant positief effect heeft op relatief sociaal kapitaal binnen het eigen dorp en een significant negatief effect op relatief sociaal kapitaal buiten het eigen dorp.	Relatief cognitief sociaal kapitaal binnen dorp (2a) en relatief structureel sociaal kapitaal binnen dorp (2b)	Leeftijd , (verwacht) opleidingsniveau, woonplaats, verwachte verandering woonsituatie en geslacht (voor zowel 2a als 2b)
H3: Relatief meer sociaal kapitaal binnen het eigen dorp hangt samen met een lagere vertrekintentie, terwijl relatief meer sociaal kapitaal buiten het eigen dorp samenhangt met een hogere vertrekintentie.	Model 3	Vaststellen of relatief aandeel sociaal kapitaal binnen het eigen dorp een significant negatief effect heeft op vertrekintentie en of relatief sociaal kapitaal buiten het eigen dorp een significant positief effect heeft.	Vertrekintentie	Relatief cognitief sociaal kapitaal binnen dorp, relatief structureel sociaal kapitaal binnen dorp , (verwacht) opleidingsniveau, woonplaats, verwachte verandering woonsituatie en geslacht
H4: De hogere vertrekintentie van jongere mensen wordt (deels) verklaard doordat zij relatief minder sociaal kapitaal binnen en relatief meer sociaal kapitaal buiten het eigen dorp bezitten.	Model 4	Vaststellen of het negatieve effect van leeftijd op vertrekintentie zwakker wordt via mediatie door relatief aandeel sociaal kapitaal binnen en buiten het eigen dorp.	Vertrekintentie	Relatief cognitief sociaal kapitaal binnen dorp, relatief structureel sociaal kapitaal binnen dorp, leeftijd , (verwacht) opleidingsniveau, woonplaats, verwachte verandering woonsituatie en geslacht.
H5: De relatie tussen sociaal kapitaal en vertrekintentie wordt gemodereerd door leeftijd, waarbij de effecten van sociaal kapitaal binnen en buiten het eigen dorp op vertrekintentie sterker worden naarmate de leeftijd hoger is.	Model 5	Vaststellen of het effect van relatief aandeel sociaal kapitaal binnen en buiten het eigen dorp op vertrekintentie via moderatie sterker wordt bij een hogere leeftijd.	Vertrekintentie	Leeftijd, relatief cognitief sociaal kapitaal binnen dorp, relatief structureel sociaal kapitaal binnen dorp, relatief cognitief sociaal kapitaal binnen dorp * leeftijd, relatief structureel sociaal kapitaal binnen dorp * leeftijd , (verwacht) opleidingsniveau, woonplaats, verwachte verandering woonsituatie en geslacht.

3.4.6. Evaluatie invloed van geïmputeerde data

Om te controleren in hoeverre de gepoolde analyses beïnvloed worden door de geïmputeerde data wordt de ‘fraction of information about Q missing due to nonresponse (γ)’ geëvalueerd. Met de γ evalueer je hoeveel invloed de ontbrekende waarden hebben op de totale variantie, waarbij rekening wordt gehouden met het aantal imputaties dat is uitgevoerd (van Buuren, 2018). Volgens van Buuren (2018), gebaseerd op Li, Raghunathan en Rubin (1991), kan een γ tot 0,2 worden geïnterpreteerd als gering, 0,3 als matig groot en 0,5 als hoog.

3.4.7. Modelevaluatie

Om te beoordelen in welke mate de modellen de variantie in de data verklaren, worden de R^2 en de R^2 adjusted geëvalueerd. Hiervoor wordt gebruikgemaakt van de functie *pool.r.squared()* uit het R-pakket *MICE* (van Buuren & Groothuis-Oudshoorn, 2011). Deze functie maakt het mogelijk om de verklaarde variantie te schatten over meerdere geïmputeerde datasets door de R^2 -waarden van afzonderlijke analyses te aggregeren (poolen).

4. Voorbereidende analyses

In dit hoofdstuk worden de voorbereidende analyses gepresenteerd. Eerst wordt in paragraaf 4.1 de kwaliteit van de imputaties geanalyseerd. Vervolgens worden de univariate (paragraaf 4.2) en bivariate statistieken vóór en na imputatie besproken. Paragraaf 4.4 gaat in op de invloed van de geïmputeerde data waarna in 4.5 de modelevaluatie wordt geanalyseerd.

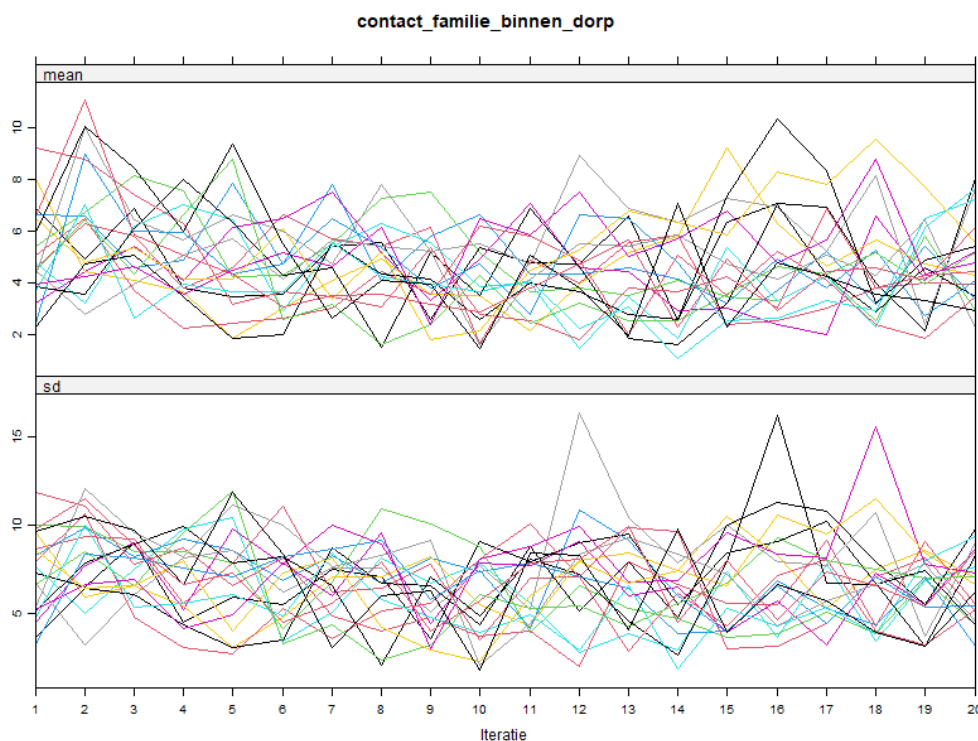
4.1. Analyse kwaliteit van imputaties

In bijlage C vind je het script waarmee de imputaties in R zijn uitgevoerd. Zoals besproken in paragraaf 3.4.2 wordt nu de kwaliteit van de imputaties geanalyseerd.

Convergentie

Voor convergentie is het uitgangspunt dat de ketens door elkaar moeten lopen zonder duidelijke trend. Om dit inzichtelijk te maken zijn convergentieplots gegenereerd. De iteraties staan op de x-as van de plots en het gemiddelde (bovenste grafiek) en de standaarddeviatie (onderste grafiek) van de geïmputeerde waarden staan op de y-as. Elke keten staat voor één geïmputeerde dataset.

De convergentieplots geven geen aanleiding om aan te nemen dat het uitgangspunt is geschonden. De ketens lopen door elkaar en er is geen sprake van duidelijke trends. Een voorbeeld hiervan is te vinden in figuur 2. In de plot is het vragenlijstitem aangeduid met de R-naam zoals gebruikt in de analyses. Zie bijlage D voor een overzicht met deze R-naam en het bijbehorende vragenlijstitem. Alle convergentieplots zijn opgenomen in bijlage A.



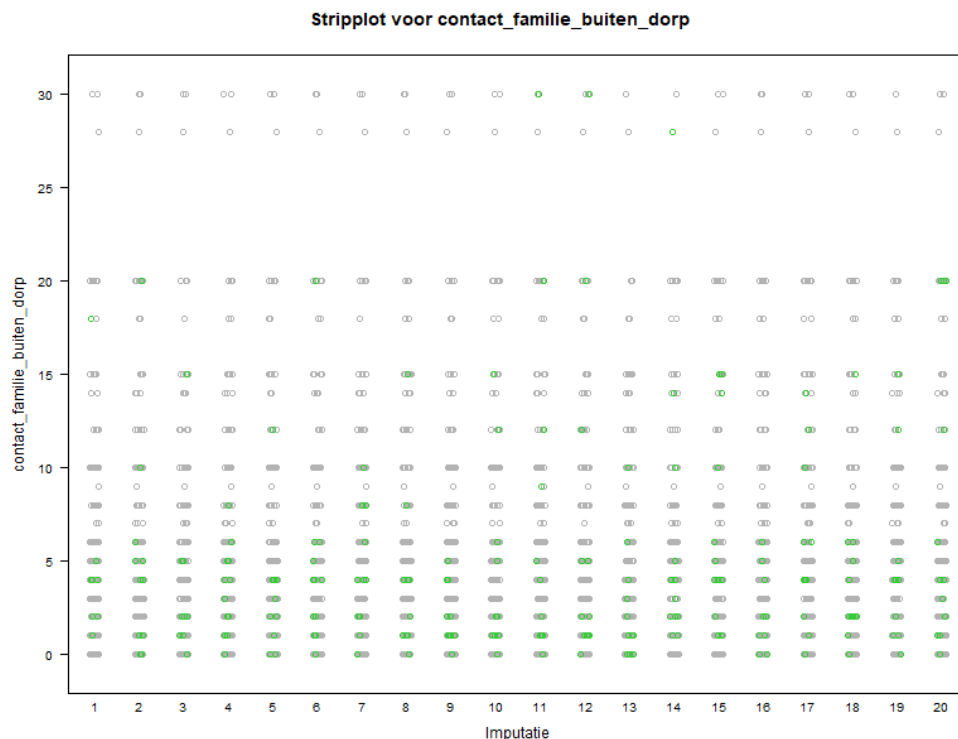
Figuur 2

Convergentieplot voor vragenlijstitem 'contact_familie_binnen_dorp' met gemiddelde (boven) en standaarddeviatie (onder) van geïmputeerde waarden per iteratie. Elke keten representeert één geïmputeerde dataset.

Plausibiliteit

Voor plausibiliteit is het uitgangspunt dat de geïmputeerde waarden niet onmogelijk of niet-plausibel zouden moeten zijn. Om dit inzichtelijk te maken zijn stripplots gegenereerd. De x-as van de stripplots toont bij welke geïmputeerde dataset de waarden horen en de y-as geeft de mogelijke waarden weer. In de plot geven grijze punten de geobserveerde waarden weer en groene punten de geïmputeerde waarden.

De stripplots tonen geen onmogelijke of niet-plausibele waarden. Zo zijn er bijvoorbeeld geen negatieve waarden geïmputeerd en zijn de geïmputeerde waarden nooit hoger dan het maximum van de geobserveerde waarden. Bij de meerkeuzevragenlijstitems vallen de imputaties altijd binnen de geldige antwoordmogelijkheden. Daarnaast laten de plots zien dat de verdelingen van de geïmputeerde waarden in lijn liggen met die van de geobserveerde waarden. Er zijn geen duidelijke afwijkingen zichtbaar die erop wijzen dat de data niet ‘Missing at random’ is. Een voorbeeld is opgenomen in figuur 3. In de plot is het vragenlijstitem aangeduid met de R-naam zoals gebruikt in de analyses. Zie bijlage D voor een overzicht met deze R-naam en het bijbehorende vragenlijstitem. Alle stripplots zijn terug te vinden in bijlage A.



Figuur 3

Stripplot voor het vragenlijstitem 'contact_familie_buiten_dorp' met geobserveerde (grijs) en geïmputeerde (groen) waarden per geïmputeerde dataset

4.2. Univariate statistieken voor en na imputatie

In tabel 3.1 tot en met 3.6 vind je de univariate statistieken van alle variabelen die in de onderzoeksmodellen zijn opgenomen. Voor de variabelen waarbij imputatie is toegepast, zijn ook de statistieken na imputatie opgenomen. Deze zijn berekend volgens de methode die is toegelicht in paragraaf 3.4.3. Leeftijd en vertrekintentie zijn niet geïmputeerd en worden daarom alleen in tabel 3.1 getoond.

Tabel 3.1*Univariate statistieken van continue en ordinale variabelen voor imputatie*

Variabele	N	Min	Max	Gem	Stdev
Leeftijd	356	16,00	92,00	52,23	19,85
(Verwacht) opleidingsniveau	355	0,00	2,00	1,24	0,81
Vertrekintentie	356	0,00	10,00	2,07	2,85
Relatief cognitief sociaal kapitaal binnen dorp	344	0,11	0,72	0,52	0,06
Relatief cognitief sociaal kapitaal binnen dorp *	344	2,43	54,67	27,04	10,93
Leeftijd					
Relatief structureel sociaal kapitaal binnen dorp	317	0,00	1,00	0,52	0,28
Relatief structureel sociaal kapitaal binnen dorp *	317	0,00	75,93	27,35	18,83
Leeftijd					

Tabel 3.2*Univariate statistieken van continue en ordinale variabelen na imputatie*

Variabele	N	Min	Max	Gem	Stdev
(Verwacht) opleidingsniveau	356	0,00	2,00	1,24	0,81
Relatief cognitief sociaal kapitaal binnen dorp	356	0,11	0,73	0,51	0,06
Relatief cognitief sociaal kapitaal binnen dorp *	356	2,43	54,67	26,99	10,90
Leeftijd					
Relatief structureel sociaal kapitaal binnen dorp	356	0,00	1,00	0,52	0,28
Relatief structureel sociaal kapitaal binnen dorp *	356	0,00	76,31	27,87	19,13
Leeftijd					

Tabel 3.3*Univariate statistieken verwachte verandering woonsituatie voor imputatie*

	N	Geen verwachte verandering	Wel verwachte verandering
Verwachte verandering woonsituatie	353	309	44

Tabel 3.4*Univariate statistieken verwachte verandering woonsituatie na imputatie*

	N	Geen verwachte verandering	Wel verwachte verandering
Verwachte verandering woonsituatie	356	311	45

Tabel 3.5*Univariate statistieken woonplaats (geen imputatie nodig)*

	N	Damwâld	De Westereen	Feanwâlden
Woonplaats	356	164	81	111

Tabel 3.6*Univariate statistieken geslacht (geen imputatie nodig)*

	N	Man	Vrouw	Anders	Zeg ik liever niet
Geslacht	356	154	200	0	2

Uit de univariate statistieken blijkt dat de gemiddelden en spreidingen vóór en na imputatie vrijwel gelijk zijn gebleven. Er is daarmee geen indicatie dat de imputatie een versturende werking heeft gehad. Deze statistieken geven ook geen aanwijzing dat de ontbrekende data niet ‘Missing at random’ zijn aangezien de univariate statistieken van de variabelen nauwelijks veranderen door de imputatie.

De respondenten zijn gemiddeld 52,23 jaar oud ($SD = 19,85$). De minimumleeftijd voor deelname was 16 jaar. De gemiddelde inwoner van 16 jaar of ouder in de gemeente Dantumadiel in 2024 was 51,16 jaar oud (Centraal Bureau voor de Statistiek, z.d.-a). De gemiddelde leeftijd van de respondenten wijkt dus amper af van inwoners van 16 jaar en ouder in de gehele gemeente Dantumadiel. Het gemiddelde (verwachte) opleidingsniveau ligt met 1,24 iets boven middelbaar opgeleid. Dit wijkt licht af van de situatie in de populatie van gemeente Dantumadiel waar deze gemiddelde score bij de meest recente data van 2022 op 1,04 uit zou komen (Centraal Bureau voor de Statistiek, z.d.-f).

Wat opvalt is dat de vertrekintentie met gemiddeld 2,07 op een schaal van 0 tot 10 erg laag ligt. Respondenten hebben over het algemeen dus niet de intentie om het dorp waar zij wonen binnen 2 jaar te verlaten.

Voor zowel cognitief als structureel sociaal kapitaal geldt dat respondenten gemiddeld iets meer van dat sociaal kapitaal bezitten binnen hun dorp dan buiten hun dorp. Ook valt op dat de standaarddeviatie van cognitief sociaal kapitaal veel lager ligt dan bij structureel sociaal kapitaal. Dit is te verklaren doordat cognitief sociaal kapitaal vragenlijstitems bevatte met een likert schaal terwijl structureel sociaal kapitaal ging over het aantal keren dat de respondent iets per maand deed.

Veruit de meeste respondenten verwachtten niet dat hun woonsituatie binnen twee jaar zou veranderen. De meeste respondenten (164) wonen daarnaast in Damwâld, 111 in Feanwâlden en de minste respondenten (81) wonen in De Westereen. In totaal wonen in 2024 in Damwâld 5690, in De Westereen 5170 en in Feanwâlden 3580 mensen (Centraal Bureau voor de Statistiek, 2024). Hieruit blijkt dat de Westereen ondervertegenwoordigd is in de steekproef en dat Damwâld en Feanwâlden juist licht oververtegenwoordigd zijn. Vrouwen zijn licht oververtegenwoordigd in deze steekproef. In de steekproef is 56,50 procent van de respondenten die hun geslacht wilden aangeven vrouw en in de gehele gemeente Dantumadiel was 49,68 procent van de inwoners in 2024 vrouw (Centraal Bureau voor de Statistiek, z.d.-a).

4.3. Bivariate statistieken voor en na imputatie

In tabel 4.1 staan de bivariate statistieken voor imputatie en in tabel 4.2 de gemiddelde statistieken na imputatie, zoals toegelicht in paragraaf 3.4.4.

Bij de statistieken voor imputatie staat er een * als er sprake is van een significant verband op 95% significantieniveau. Zoals toegelicht in paragraaf 3.4.4 kunnen er geen p-waarden worden gerapporteerd voor de bivariate statistieken na imputatie. Voor de interpretatie wordt daarom het verschil in sterkte van de samenhang voor en na imputatie besproken.

Tabel 4.1*Bivariate statistieken vóór imputatie*

	1	2	3	4	5	6	7	8	9	10
1: Woonplaats	—									
2: Leeftijd	$\sqrt{R^2}=,14$ $p=,033^*$ $n=356$	—								
3: (Verwacht) opleidingsniveau	$\sqrt{R^2}=,18$ $p=,002^*$ $n=355$	$\rho=-,33$ $p=<,001^*$ $n=355$	—							
4: Geslacht	$V=,05$ $p=,729$ $n=356$	$\sqrt{R^2}=,09$ $p=,233$ $n=356$	$\sqrt{R^2}=,02$ $p=,898$ $n=355$	—						
5: Verwachte verandering woonsituatie	$V=,06$ $p=,529$ $n=353$	$\sqrt{R^2}=,29$ $p=<,001^*$ $n=353$	$\sqrt{R^2}=,19$ $p=<,001^*$ $n=352$	$V=,09$ $p=,236$ $n=353$	—					
6: Vertrekintentie	$\sqrt{R^2}=,01$ $p=,983$ $n=356$	$\rho=-0,29$ $p=<,001^*$ $n=356$	$\rho=,16$ $p=,003^*$ $n=355$	$\sqrt{R^2}=,08$ $p=,296$ $n=356$	$\sqrt{R^2}=,32$ $p=<,001^*$ $n=353$	—				
7: Relatief cognitief sociaal kapitaal binnen dorp	$\sqrt{R^2}=,10$ $p=,159$ $n=344$	$\rho=,11$ $p=,050^*$ $n=344$	$\rho=-,04$ $p=,435$ $n=343$	$\sqrt{R^2}=,08$ $p=,325$ $n=344$	$\sqrt{R^2}=,06$ $p=,261$ $n=341$	$\rho=-,09$ $p=,111$ $n=344$	—			
8: Relatief cognitief sociaal kapitaal binnen dorp * leeftijd	$\sqrt{R^2}=,15$ $p=,017^*$ $n=344$	$\rho=,96$ $p=<,001^*$ $n=344$	$\rho=-,33$ $p=<,001^*$ $n=343$	$\sqrt{R^2}=,09$ $p=,233$ $n=344$	$\sqrt{R^2}=,31$ $p=<,001^*$ $n=341$	$\rho=-,31$ $p=<,001^*$ $n=344$	$\rho=,32$ $p=<,001^*$ $n=344$	—		
9: Relatief structureel sociaal kapitaal binnen dorp	$\sqrt{R^2}=,08$ $p=,391$ $n=317$	$\rho=,08$ $p=,168$ $n=317$	$\rho=-,13$ $p=,017^*$ $n=316$	$\sqrt{R^2}=,12$ $p=,039^*$ $n=317$	$\sqrt{R^2}=,10$ $p=,082$ $n=314$	$\rho=-,16$ $p=,004^*$ $n=317$	$\rho=,08$ $p=,167$ $n=309$	$\rho=,10$ $p=,087$ $n=309$	—	
10: Relatief structureel sociaal kapitaal binnen dorp * leeftijd	$\sqrt{R^2}=,11$ $p=,142$ $n=317$	$\rho=,57$ $p=<,001^*$ $n=317$	$\rho=-,29$ $p=<,001^*$ $n=316$	$\sqrt{R^2}=,11$ $p=,057$ $n=317$	$\sqrt{R^2}=,20$ $p=<,001^*$ $n=314$	$\rho=-,28$ $p=<,001^*$ $n=317$	$\rho=,11$ $p=,046^*$ $n=309$	$\rho=,57$ $p=<,001^*$ $n=309$	$\rho=,80$ $p=<,001^*$ $n=317$	—

Tabel 4.2*Bivariate statistieken na imputatie*

	1	2	3	4	5	6	7	8	9	10
1: Woonplaats	—									
2: Leeftijd	$\sqrt{R^2}=.14$ n=356	—								
3: (Verwacht) opleidingsniveau	$\sqrt{R^2}=.18$ n=356	$\rho=-.33$ n=356	—							
4: Geslacht	$V=.05$ n=356	$\sqrt{R^2}=.09$ n=356	$\sqrt{R^2}=.02$ n=356	—						
5: Verwachte verandering woonsituatie	$V=.07$ n=356	$\sqrt{R^2}=.29$ n=356	$\sqrt{R^2}=.18$ n=356	$V=.09$ n=356	—					
6: Vertrekintentie	$\sqrt{R^2}=.01$ n=356	$\rho=-.29$ n=356	$\rho=.16$ n=356	$\sqrt{R^2}=.08$ n=356	$\sqrt{R^2}=.31$ n=356	—				
7: Relatief cognitief sociaal kapitaal binnen dorp	$\sqrt{R^2}=.09$ n=356	$\rho=.10$ n=356	$\rho=-.02$ n=356	$\sqrt{R^2}=.07$ n=356	$\sqrt{R^2}=.06$ n=356	$\rho=-.09$ n=356	—			
8: Relatief cognitief sociaal kapitaal binnen dorp * leeftijd	$\sqrt{R^2}=.13$ n=356	$\rho=.96$ n=356	$\rho=-.31$ n=356	$\sqrt{R^2}=.10$ n=356	$\sqrt{R^2}=.28$ n=356	$\rho=-.30$ n=356	$\rho=.32$ n=356	—		
9: Relatief structureel sociaal kapitaal binnen dorp	$\sqrt{R^2}=.08$ n=356	$\rho=.12$ n=356	$\rho=-.14$ n=356	$\sqrt{R^2}=.10$ n=356	$\sqrt{R^2}=.11$ n=356	$\rho=-.17$ n=356	$\rho=.06$ n=356	$\rho=.12$ n=356	—	
10: Relatief structureel sociaal kapitaal binnen dorp * leeftijd	$\sqrt{R^2}=.13$ n=356	$\rho=.60$ n=356	$\rho=-.29$ n=356	$\sqrt{R^2}=.10$ n=356	$\sqrt{R^2}=.21$ n=356	$\rho=-.29$ n=356	$\rho=.10$ n=356	$\rho=.59$ n=356	$\rho=.81$ n=356	—

De waarden in de tabellen laten zien dat de sterkte van de verbanden tussen de variabelen vóór en na imputatie vrijwel gelijk is gebleven. Hiermee is er geen aanwijzing dat de onderlinge samenhang tussen variabelen verstoord is geraakt door de imputatieprocedure. Er is daarmee ook geen aanwijzing dat de ontbrekende data niet ‘Missing at Random’ zijn aangezien de statistieken vrijwel gelijk blijven het toepassen van imputatie. De korte duiding hieronder is gebaseerd op de p-waarden die enkel konden worden berekend bij de statistieken voor imputatie. Omdat het toepassen van imputatie weinig invloed heeft gehad, is het aannemelijk dat de verbanden ook na imputatie in stand blijven.

Relevante observaties zijn ten eerste dat vertrekintentie significant samenhangt met leeftijd en relatief structureel sociaal kapitaal binnen dorp. Voor relatief cognitief sociaal kapitaal binnen dorp is de bivariate samenhang wel aanwezig, maar niet significant. Ook de controlevariabelen (verwacht) opleidingsniveau en verwachte verandering woonsituatie hangen significant samen met de vertrekintentie.

Verder valt op dat relatief cognitief sociaal kapitaal binnen dorp en relatief structureel sociaal kapitaal binnen dorp niet significant met elkaar samenhangen.

Leeftijd vertoont een significante bivariate samenhang met relatief cognitief sociaal kapitaal binnen het dorp, met een p-waarde van 0,05 vóór imputatie. Na imputatie daalt de correlatiecoëfficiënt (ρ) licht van 0,11 naar 0,10. Aangezien p-waarden na imputatie niet worden gerapporteerd, kan niet met zekerheid worden vastgesteld of de samenhang nog significant is. Voor relatief structureel sociaal kapitaal binnen het dorp is er geen significante bivariate samenhang met leeftijd gevonden.

4.4. Analyse invloed van geïmputeerde data

Alle γ -waarden blijven in alle modellen ruim onder de grens voor geringe invloed (0,2) die in paragraaf 3.4.6 is besproken. Er is één uitzondering bij een controlevariabele: geslacht4 (“Zeg ik liever niet”) in model H2b. Hiervoor werd een waarde van 0,33 gevonden. Dit geval valt dus iets boven de grens van 0,3 bij matig groot. Aangezien dit de enige uitschieter betreft en alle overige waarden onder de grens voor geringe invloed blijven, is de impact van imputatie op de analyses als geheel zeer beperkt. De volledige lijst met γ -waarden is opgenomen in bijlage B.

4.5. Analyse modevaluatie

In tabel 5 zijn, op basis van analyses na imputatie en pooling, de R^2 - en R^2 Adjusted waarden weergegeven voor elk model. Deze modevaluatie volgt de aanpak beschreven in paragraaf 3.4.7.

Tabel 5

Modevaluatie per model

Model	Afhankelijke variabele	Onafhankelijke variabelen	R^2	R^2 Adjusted
Model 1	Vertrekintentie	Leeftijd, (verwacht) opleidingsniveau, woonplaats, verwachte verandering woonsituatie en geslacht	,17	,10
Model 2a	Relatief cognitief sociaal kapitaal binnen dorp	Leeftijd, (verwacht) opleidingsniveau, woonplaats, verwachte verandering woonsituatie en geslacht	,03	,00
Model 2b	Relatief structureel sociaal kapitaal binnen dorp	Leeftijd, (verwacht) opleidingsniveau, woonplaats, verwachte verandering woonsituatie en geslacht	,04	,01
Model 3	Vertrekintentie	Relatief cognitief sociaal kapitaal binnen dorp, relatief structureel sociaal kapitaal binnen dorp, (verwacht) opleidingsniveau, woonplaats, verwachte verandering woonsituatie en geslacht	,14	,08
Model 4	Vertrekintentie	Relatief cognitief sociaal kapitaal binnen dorp, relatief structureel sociaal kapitaal binnen dorp, leeftijd, (verwacht) opleidingsniveau, woonplaats, verwachte verandering woonsituatie en geslacht.	,18	,12
Model 5	Vertrekintentie	Leeftijd, relatief cognitief sociaal kapitaal binnen dorp, relatief structureel sociaal kapitaal binnen dorp, relatief cognitief sociaal kapitaal binnen dorp * leeftijd, relatief structureel sociaal kapitaal binnen dorp * leeftijd, (verwacht) opleidingsniveau, woonplaats, verwachte verandering woonsituatie en geslacht.	,21	,13

Zoals te zien in tabel 5 verklaart model 1 ($R^2 = ,17$; R^2 adjusted = ,10) met leeftijd en de controlevariabelen iets meer van de variantie dan model 3 ($R^2 = ,14$; R^2 adjusted = ,08) met relatief cognitief en structureel sociaal kapitaal binnen dorp. In model 4 worden zowel leeftijd als beide vormen van relatief sociaal kapitaal binnen dorp toegevoegd en stijgt de verklaarde variantie licht ($R^2 = ,18$; R^2 adjusted = ,12). Met het toevoegen van de interactietermen van relatief cognitief sociaal kapitaal binnen dorp met leeftijd en relatief structureel sociaal kapitaal binnen dorp met leeftijd stijgt de verklaarde variantie verder ($R^2 = ,21$; R^2 adjusted = ,13). Dit wijst erop dat het toevoegen van de interactietermen het model beter maakt in het voorspellen van vertrekintentie.

De modellen waarin de rol van leeftijd bij relatief cognitief (model 2a) en structureel (model 2b) sociaal kapitaal binnen het eigen dorp wordt onderzocht, verklaren weinig variantie ($R^2 = ,03$; R^2 adjusted = ,00 voor model 2a en $R^2 = ,04$; R^2 adjusted = ,01 voor model 2b). Dit suggereert dat leeftijd slechts een beperkte rol speelt bij de relatieve hoeveelheid sociaal kapitaal dat iemand binnen het eigen dorp bezit.

5. Resultaten

In dit hoofdstuk worden in paragraaf 5.1 tot en met 5.5 de resultaten getoond en de hypothesen beantwoord die horen bij deelvraag 1 tot en met 5. In de regressiemodellen worden enkele categorische variabelen opgenomen die als dummyvariabelen zijn gecodeerd. Dit betekent dat één categorie telkens als referentie wordt gebruikt. Voor woonplaats is de basiscategorie Damwâld, voor verwachte verandering in woonsituatie is dat ‘niet’ en voor geslacht ‘man’. De gepresenteerde coëfficiënten voor de overige categorieën worden dus telkens geïnterpreteerd ten opzichte van deze referentiecategorie.

Voor de ordinale variabele (verwacht) opleidingsniveau zijn orthogonale contrasten opgenomen: .L en .Q. Het .L-contrast laat zien of er sprake is van een lineair patroon: bijvoorbeeld of vertrekintentie toeneemt of afneemt naarmate het (verwachte) opleidingsniveau stijgt. Een positief en significant resultaat bij .L wijst dan op een toenemende vertrekintentie bij hoger opgeleiden, terwijl een negatief resultaat op het omgekeerde duidt. Het .Q-contrast geeft aan of het verband gebogen verloopt: bijvoorbeeld of middelbaar opgeleiden een hogere vertrekintentie hebben dan zowel laag- als hoogopgeleiden (U-vorm) of juist lager (omgekeerde U-vorm). Als .Q niet significant is, wijst dat op een overwegend rechte lijn zonder duidelijke buiging in het patroon.

Een * bij de p-waarde in de tabellen geeft aan dat het resultaat significant is bij een significantieniveau van 95%.

5.1. Hypothesetoetsing deelvraag 1: leeftijd en vertrek uit dorpen in niet-stedelijke gebieden

Hypothese 1 is: Hogere leeftijd hangt samen met een lagere vertrekintentie uit dorpen in niet-stedelijke gebieden. Om deze hypothese te toetsen is model 1 geschat. De resultaten staan in tabel 6.1.

Tabel 6.1

Resultaten model 1 met de afhankelijke variabele vertrekintentie (schaal 0 t/m 10)

Predictor	B	SE	t	p
Intercept	3,54	0,54	6,50	<,001*
Leeftijd	-0,04	0,01	-4,44	<,001*
(Verwacht) opleidingsniveau .L	0,06	0,27	0,23	,814
(Verwacht) opleidingsniveau .Q	0,16	0,26	0,63	,528
Woonplaats = De Westereen	-0,35	0,36	-0,97	,334
Woonplaats = Feanwâlden	-0,29	0,33	-0,87	,382
Verwachte verandering woonsituatie = wel	2,09	0,45	4,66	<,001*
Geslacht = vrouw	0,49	0,29	1,72	,086
Geslacht = zeg ik liever niet	1,09	1,88	0,58	,561

Uit tabel 6.1 blijkt dat leeftijd een significant negatief effect heeft op vertrekintentie ($B = -0,04$; $p <,001$). Bij gebruik van de nauwkeurigere regressiecoëfficiënt voor leeftijd ($B = -0,03566$) blijkt dat een persoon van 20 jaar, bij verder gelijke omstandigheden, een verschil in vertrekintentie van 2,14 punten heeft ten opzichte van een persoon van 80 jaar. Hypothese 1 wordt dus **bevestigd**.

5.2. Hypothesetoetsing deelvraag 2: locatie sociaal kapitaal en leeftijd

Hypothese 2 is: Hogere leeftijd hangt samen met relatief meer sociaal kapitaal binnen het eigen dorp en relatief minder sociaal kapitaal buiten het eigen dorp. Om deze hypothese te toetsen zijn twee modellen geschat met relatief cognitief sociaal kapitaal binnen dorp (model 2a) en relatief structureel sociaal kapitaal binnen dorp (model 2b) als afhankelijke variabelen. De resultaten staan in tabel 6.2 en 6.3.

Tabel 6.2

Resultaten model 2a met de afhankelijke variabele relatief cognitief sociaal kapitaal binnen dorp (schaal 0 t/m 1)

Predictor	B	SE	t	p
Intercept	0,51	0,01	39,64	<,001*
Leeftijd	0,00	0,00	1,42	,157
(Verwacht) opleidingsniveau .L	0,01	0,01	1,39	,164
(Verwacht) opleidingsniveau .Q	0,00	0,01	0,20	,843
Woonplaats = De Westereen	-0,00	0,01	-0,58	,562
Woonplaats = Feanwâlden	-0,01	0,01	-1,79	,074
Verwachte verandering woonsituatie = wel	-0,01	0,01	-0,85	,398
Geslacht = vrouw	-0,01	0,01	-1,32	,186
Geslacht = zeg ik liever niet	-0,01	0,04	-0,25	,805

Tabel 6.3

Resultaten model 2b met de afhankelijke variabele relatief structureel sociaal kapitaal binnen dorp (schaal 0 t/m 1)

Predictor	B	SE	t	p
Intercept	0,56	0,06	9,38	<,001*
Leeftijd	0,00	0,00	0,35	,726
(Verwacht) opleidingsniveau .L	-0,05	0,03	-1,73	,085
(Verwacht) opleidingsniveau .Q	0,02	0,03	0,58	,562
Woonplaats = De Westereen	0,00	0,04	0,12	,901
Woonplaats = Feanwâlden	-0,04	0,04	-1,05	,297
Verwachte verandering woonsituatie = wel	-0,08	0,05	-1,66	,098
Geslacht = vrouw	-0,06	0,03	-1,87	,063
Geslacht = zeg ik liever niet	-0,00	0,24	-0,01	,988

Uit tabel 6.2 en 6.3 blijkt dat leeftijd geen significant effect heeft op relatief cognitief sociaal kapitaal binnen dorp ($B = 0,00$; $p = ,157$) en ook niet op relatief structureel sociaal kapitaal binnen dorp ($B = 0,00$; $p = ,726$). Hypothese 2 wordt dus **niet bevestigd**.

5.3. Hypothesetoetsing deelvraag 3: sociaal kapitaal en vertrek uit dorpen in niet-stedelijke gebieden

Hypothese 3 is: Relatief meer sociaal kapitaal binnen het eigen dorp hangt samen met een lagere vertrekintentie, terwijl relatief meer sociaal kapitaal buiten het eigen dorp samenhangt met een hogere vertrekintentie. Om deze hypothese te toetsen is model 3 geschat. De resultaten staan in tabel 6.4.

Tabel 6.4*Resultaten model 3 met de afhankelijke variabele vertrekintentie (schaal 0 t/m 10)*

Predictor	B	SE	t	p
Intercept	4,30	1,28	3,37	,001*
Relatief cognitief sociaal kapitaal binnen dorp	-4,18	2,34	-1,79	,075
Relatief structureel sociaal kapitaal binnen dorp	-1,14	0,52	-2,18	,030*
(Verwacht) opleidingsniveau .L	0,46	0,26	1,74	,082
(Verwacht) opleidingsniveau .Q	-0,01	0,26	-0,04	,967
Woonplaats = De Westereen	-0,18	0,37	-0,49	,621
Woonplaats = Feanwâlden	-0,40	0,34	-1,20	,231
Verwachte verandering woonsituatie = wel	2,46	0,44	5,55	<,001*
Geslacht = vrouw	0,52	0,29	1,80	,073
Geslacht = zeg ik liever niet	0,59	1,92	0,31	,759

Uit tabel 6.4 blijkt dat relatief cognitief sociaal kapitaal binnen dorp negatief samenhangt met vertrekintentie ($B = -4,18$; $p = ,075$), dit effect is echter net niet significant.

Voor relatief structureel sociaal kapitaal binnen dorp is dit negatieve verband wel significant ($B = -1,14$; $p = ,030$). Iemand met 0% structureel sociaal kapitaal binnen het eigen dorp heeft dus gemiddeld een 1,14 punten hogere verwachte vertrekintentie dan iemand met 100% binnen het eigen dorp als de overige variabelewaarden constant zijn.

Deze bevindingen bieden grotendeels ondersteuning voor hypothese 3. Personen die relatief meer sociaal kapitaal binnen hun eigen dorp hebben, hebben naar verwachting een lagere vertrekintentie als de overige variabelewaarden constant zijn. Dit verband is significant voor structureel sociaal kapitaal, maar net niet voor cognitief sociaal kapitaal. Hypothese 3 wordt dus **grotendeels bevestigd**.

5.4. Hypothesetoetsing deelvraag 4: Mediatie van leeftijdsverschillen in vertrek uit dorpen in niet-stedelijke gebieden door locatie sociaal kapitaal

Hypothese 4 is: De hogere vertrekintentie van jongere mensen wordt (deels) verklaard doordat zij relatief minder sociaal kapitaal binnen en relatief meer sociaal kapitaal buiten het eigen dorp bezitten. Om deze hypothese te toetsen is model 4 geschat. De resultaten staan in tabel 6.5.

Tabel 6.5*Resultaten model 4 met de afhankelijke variabele vertrekintentie (schaal 0 t/m 10)*

Predictor	B	SE	t	p
Intercept	5,92	1,30	4,55	<,001*
Leeftijd	-0,03	0,01	-4,30	<,001*
Relatief cognitief sociaal kapitaal binnen dorp	-3,44	2,29	-1,51	,133
Relatief structureel sociaal kapitaal binnen dorp	-1,12	0,51	-2,17	,031*
(Verwacht) opleidingsniveau .L	0,04	0,27	0,14	,889
(Verwacht) opleidingsniveau .Q	0,19	0,26	0,73	,469
Woonplaats = De Westereen	-0,36	0,36	-1,01	,315
Woonplaats = Feanwâlden	-0,38	0,33	-1,15	,251
Verwachte verandering woonsituatie = wel	1,97	0,45	4,41	<,001*
Geslacht = vrouw	0,40	0,28	1,39	,166
Geslacht = zeg ik liever niet	1,05	1,87	0,56	,574

In model 4, waarin relatief cognitief en structureel sociaal kapitaal binnen dorp zijn toegevoegd als verklarende variabelen, blijft leeftijd een significante voorspeller van vertrekintentie ($B = -0,03$; $p < ,001$). Op basis van de nauwkeurigere regressiecoëfficiënt in dit model ($-0,03439$) blijkt dat een persoon van 20 jaar, bij constante waarden van sociaal kapitaal en de controlevariabelen, naar verwachting een vertrekintentie heeft die 2,06 punten hoger ligt dan die van een persoon van 80 jaar. In model 1 zonder sociaal kapitaal was dit 2,14.

De invloed van leeftijd op vertrekintentie neemt in model 4 dus zeer licht af ten opzichte van model 1. Deze afname is zo beperkt dat deze in de praktijk nauwelijks tot niet betekenisvol is. Bovendien blijkt uit model 2a en 2b dat leeftijd niet significant samenhangt met sociaal kapitaal binnen het eigen dorp. Hypothese 4 wordt daarom **niet bevestigd**.

5.5. Hypothesetoetsing deelvraag 5: Moderatie van verschillen in invloed sociaal kapitaal op vertrek uit dorpen in niet-stedelijke gebieden door leeftijd

Hypothese 5 is: De relatie tussen sociaal kapitaal en vertrekintentie wordt gemodereerd door leeftijd, waarbij de effecten van sociaal kapitaal binnen en buiten het eigen dorp op vertrekintentie sterker worden naarmate de leeftijd hoger is. Om deze hypothese te toetsen is model 5 geschat. De resultaten staan in tabel 6.6.

Tabel 6.6

Resultaten model 5 met de afhankelijke variabele vertrekintentie (schaal 0 t/m 10)

Predictor	B	SE	t	p
Intercept	12,31	2,82	4,37	<,001*
Leeftijd	-0,17	0,06	-3,13	,002*
Relatief cognitief sociaal kapitaal binnen dorp	-14,02	5,53	-2,53	,012*
Relatief structureel sociaal kapitaal binnen dorp	-3,09	1,49	-2,07	,040*
Relatief cognitief sociaal kapitaal binnen dorp * leeftijd	0,23	0,11	2,12	,035*
Relatief structureel sociaal kapitaal binnen dorp * leeftijd	0,04	0,03	1,43	,155
(Verwacht) opleidingsniveau .L	0,14	0,28	0,51	,612
(Verwacht) opleidingsniveau .Q	0,15	0,26	0,59	,554
Woonplaats = De Westereen	-0,40	0,36	-1,10	,272
Woonplaats = Feanwâlden	-0,31	0,33	-0,94	,348
Verwachte verandering woonsituatie = wel	1,86	0,45	4,13	<,001*
Geslacht = vrouw	0,36	0,28	1,28	,202
Geslacht = zeg ik liever niet	1,04	1,86	0,56	,576

Uit de resultaten blijkt dat de interactieterm tussen leeftijd en relatief cognitief sociaal kapitaal binnen dorp significant is ($B = 0,23$; $p = ,035$), maar dat de richting van dit effect tegengesteld is aan de verwachting. De positieve interactiecoëfficiënt geeft aan dat het negatieve verband tussen relatief cognitief sociaal kapitaal binnen dorp en vertrekintentie afneemt naarmate de leeftijd toeneemt.

Concreet betekent dit dat het verschil in vertrekintentie tussen jongere mensen met veel en weinig cognitief sociaal kapitaal binnen het eigen dorp groter is dan dat verschil bij oudere mensen. Op basis van de nauwkeurigere regressiecoëfficiënten in dit model ($B = -14,02229$ voor cognitief sociaal kapitaal en $B = 0,23277$ voor de interactie met leeftijd) blijkt dat een 20-jarige met 100% cognitief sociaal kapitaal binnen het eigen dorp naar verwachting een 9,37 punten

lagere vertrekintentie heeft dan een leeftijdsgenoot met 0% sociaal kapitaal binnen het eigen dorp, bij verder gelijke omstandigheden op alle andere variabelen. Voor een 80-jarige keert dit verschil om: iemand met 100% cognitief sociaal kapitaal binnen het eigen dorp heeft dan juist een 4,60 punten hogere verwachte vertrekintentie dan iemand met 0%. Jongere mensen zijn dus veel gevoeliger voor de locatie van hun cognitieve sociaal kapitaal bij hun beslissing om te vertrekken of te blijven dan oudere mensen. Vanaf 61 jaar wordt dit effect van relatief cognitief sociaal kapitaal binnen dorp op vertrekintentie zelfs positief.

Voor relatief structureel sociaal kapitaal binnen dorp werd geen significante interactie met leeftijd vastgesteld ($B = 0,04$; $p = ,155$). De invloed van dit type sociaal kapitaal op vertrekintentie verschilt dus niet significant tussen jongere en oudere respondenten.

De interactieterm van relatief cognitief sociaal kapitaal binnen dorp met leeftijd heeft een significant effect op vertrekintentie maar in de tegenovergestelde richting dan werd verwacht. Voor relatief structureel sociaal kapitaal binnen dorp met leeftijd is eveneens een effect in de omgekeerde richting gevonden, maar dit effect is niet significant. Hypothese 5 wordt dan ook **niet bevestigd**.

6. Conclusie en discussie

In dit hoofdstuk worden in 6.1 eerst de conclusies besproken. Deze conclusies worden in de discussie in paragraaf 6.2 in breder perspectief geplaatst.

6.1. Conclusie

Dit onderzoek richtte zich op de vraag welke rol de locatie van sociaal kapitaal speelt bij het relatief hoge vertrekpercentage van jongere mensen uit dorpen in niet-stedelijke gebieden. Vertrekintentie is hierbij gebruikt als voorspeller van vertrek uit het dorp.

De analyses bevestigen dat leeftijd een significante voorspeller is van vertrekintentie: hoe jonger iemand is, hoe groter de kans dat diegene het dorp wil verlaten.

Het relatieve aandeel van zowel cognitief als structureel sociaal kapitaal dat zich binnen het eigen dorp bevindt varieert niet significant met leeftijd. Verschillen in de locatie van sociaal kapitaal verklaren dus niet waarom jongere mensen vaker willen vertrekken.

De mediatieanalyse bevestigt dat het opnemen van sociaal kapitaal in het model nauwelijks invloed heeft op de relatie tussen leeftijd en vertrekintentie. Er is dus nauwelijks sprake van mediatie: sociaal kapitaal en leeftijd beïnvloeden vertrekintentie grotendeels onafhankelijk van elkaar.

Als alternatief voor hypothese 4 werd in hypothese 5 moderatie van de invloed van relatief sociaal kapitaal op vertrekintentie door leeftijd getoetst. Opvallend hiervoor is dat voor relatief cognitief sociaal kapitaal binnen het eigen dorp in de modellen zonder interactieterm geen sprake van een significant effect. In model 5 waarin interactie met leeftijd als alternatief is meegenomen, werd het negatieve effect wel significant. Dit gold ook voor de positieve interactieterm. Dit betekent dat het effect van relatief cognitief sociaal kapitaal binnen het eigen dorp op vertrekintentie sterker en significant is voor jongere personen, en minder sterk wordt naarmate de leeftijd hoger is. Vanaf 61 jaar wordt het verwachte effect van relatief cognitief sociaal kapitaal binnen het eigen dorp op vertrekintentie zelfs positief.

Relatief structureel sociaal kapitaal binnen het eigen dorp laat daarentegen in de verschillende modellen een consistent en significant negatief verband met vertrekintentie zien. De interactieterm was niet significant. Dit wijst op een breed bestaand effect dat niet significant afhankelijk is van leeftijd.

Als antwoord op de onderzoeksvraag blijkt dat de locatie van sociaal kapitaal de hogere vertrekans van jongere mensen niet verklaart: zij beschikken niet over significant minder relatief sociaal kapitaal binnen het eigen dorp dan oudere mensen. Wel hangt relatief meer structureel sociaal kapitaal binnen het eigen dorp consistent samen met een lagere vertrekintentie, ongeacht leeftijd. Voor relatief cognitief sociaal kapitaal is dit effect alleen bij jongere mensen significant. De locatie van sociaal kapitaal beïnvloedt vertrekintentie dus wel, maar verklaart niet waarom jongere mensen vaker willen vertrekken dan oudere mensen.

6.2. Discussie

In dit discussiehoofdstuk worden eerst in paragraaf 6.2.1 de resultaten geïnterpreteerd. Vervolgens worden in paragraaf 6.2.2 de validiteit en betrouwbaarheid van het onderzoek besproken. In 6.2.3 worden de belangrijkste beperkingen van het onderzoek benoemd. Op basis van de bevindingen worden in paragraaf 6.2.4 aanbevelingen gedaan voor beleid en in paragraaf 6.2.5 worden suggesties voor toekomstig wetenschappelijk onderzoek besproken.

6.2.1. Interpretatie van resultaten

De resultaten van dit onderzoek sluiten in algemene zin aan bij de toepassing van het aspirations-capabilities framework (ACF) van de Haas (2021). Deze theorie werd aangehaald om te onderbouwen dat meer sociaal kapitaal binnen het eigen dorp de vertrekintentie vermindert. Uit de resultaten van hypothese 3 bleek dit inderdaad het geval te zijn. Wel was dit negatieve verband net niet significant bij relatief cognitief sociaal kapitaal binnen dorp.

Er is geen bewijs gevonden dat de relatieve locatie van het sociaal kapitaal een verklarende rol speelt bij leeftijdsverschillen in de vertrekintentie. Dit wordt verklaard doordat in dit onderzoek geen leeftijdsverschillen in de relatieve locatie van sociaal kapitaal konden worden bewezen en dus niet doordat leeftijd of sociaal kapitaal geen rol speelt bij de vertrekintentie.

Om als alternatief op de mediatiehypothese eventuele moderatie door leeftijd te onderbouwen is de sociaalemotionele selectiviteitstheorie (Carstensen, 1995) toegepast. Er werd onderbouwd dat sociaal kapitaal een belangrijkere rol gaat spelen bij oudere mensen. Als dit inderdaad het geval is, is het in lijn met het ACF aannemelijk dat de relatieve locatie van sociaal kapitaal een sterkere rol speelt in de intentie om uit het dorp te vertrekken. Deze verwachting werd echter niet bevestigd. Uit de resultaten blijkt dat de invloed van cognitief sociaal kapitaal binnen het eigen dorp juist sterker is bij jongere mensen. Deze invloed wordt zwakker naarmate de leeftijd hoger is en vanaf 61 jaar wordt de invloed van cognitief sociaal kapitaal binnen dorp op vertrekintentie zelf positief. Voor structureel sociaal kapitaal is de invloed van de interactieterm met leeftijd op vertrekintentie in dezelfde richting, maar dit effect is niet significant. Er zijn twee mogelijke redenen waardoor de hypothese niet is uitgekomen.

Ten eerste is het mogelijk dat sociaal kapitaal in de context van vertrekintentie juist een belangrijkere rol speelt voor jongere mensen dan voor oudere mensen. In dat geval wordt de centrale aanname van de sociaalemotionele selectiviteitstheorie in deze context niet ondersteund.

Ten tweede is het mogelijk dat de sociaalemotionele selectiviteitstheorie op zichzelf wel klopt, maar niet goed toepasbaar is op vertrekintentie als specifieke gedraging. De aanname dat sociaal kapitaal bij ouderen sterker doorwerkt in vertrekintentie is gebaseerd op de toepassing van de sociaalemotionele selectiviteitstheorie bij het ACF. De verwachting die hierop volgde wordt in dit onderzoek echter niet ondersteund. Dat hoeft niet te betekenen dat de afzonderlijke theorieën onjuist zijn maar wijst erop dat de belangrijkere rol die sociaal kapitaal op oudere leeftijd gaat spelen wellicht geen verband houdt met wel of niet vertrekken uit de woonplaats.

6.2.2. Validiteit en betrouwbaarheid

De externe validiteit in dit onderzoek wordt versterkt doordat er gebruik is gemaakt van een volledig willekeurige steekproef onder mensen van 16 jaar of ouder uit de betreffende drie dorpen. De gemiddelde leeftijd van deelnemers van 52,23 jaar komt goed overeen met de populatie. Het gemiddelde (verwachte) opleidingsniveau in de steekproef is licht hoger dan in de populatie. Dit verschil is echter minimaal en zou mogelijk verklaard kunnen worden doordat in dit onderzoek rekening is gehouden met opleidingen die de respondenten naar verwachting nog gaan afronden. Vrouwen zijn daarnaast licht oververtegenwoordigd ten opzichte van de populatie. De respondenten zijn verder ongelijk verdeeld over de drie dorpen: De Westereen is ondervertegenwoordigd terwijl Damwâld en Feanwâlden licht oververtegenwoordigd zijn. Deze controlevariabelen (verwacht) opleidingsniveau, woonplaats en geslacht bleken in de modellen voor hypothesetoetsing geen significante voorspellers te zijn van vertrekintentie of sociaal

kapitaal. De invloed van deze scheve vertegenwoordiging op de bevindingen is daardoor beperkt.

Ook is de betrouwbaarheid van de imputatieprocedure geëvalueerd. Hierbij werd vastgesteld dat de imputaties convergeerden, plausibel waren en slechts beperkte invloed hadden op de schattingsresultaten.

De interne validiteit en betrouwbaarheid van de operationalisaties zijn zorgvuldig besproken in paragraaf 3.2. De belangrijkste onderdelen worden nu kort samengevat.

Vertrekintentie is valide gemeten via een duidelijk vragenlijstitem met een expliciete tijdshorizon met duidelijke antwoordopties. De betrouwbaarheid van de meting is redelijk. De formulering van het vragenlijstitem is gebaseerd op eerder onderzoek, wat consistente interpretatie bevordert. Wel is de meting gevoelig voor contextuele invloeden.

Leeftijd is als objectief kenmerk zowel valide als betrouwbaar gemeten: het vragenlijstitem sluit direct aan op het concept en is niet gevoelig voor interpretatie of contextuele invloeden.

De validiteit van cognitief sociaal kapitaal is hoog: de vragenlijstitems sluiten goed aan op het theoretische concept. Wel wordt slechts één dimensie wordt gemeten. De betrouwbaarheid is hoog dankzij de sterke interne consistentie. Wel kunnen antwoorden worden beïnvloed door contextuele invloeden.

Structureel sociaal kapitaal is eveneens valide gemeten doordat ook deze vragenlijstitems goed aansluiten op het theoretische concept. Ook hier wordt echter slechts één dimensie van het concept gemeten: de frequentie per maand. Dit beperkt de validiteit. De betrouwbaarheid is over het algemeen hoog doordat het om feitelijke gedragingen gaat maar kan negatief worden beïnvloed door inschattingfouten.

De interne validiteit van het onderzoek is daarnaast versterkt door de controlevariabelen die zijn opgenomen in de analyses. De operationalisaties van deze controlevariabelen zijn toegelicht in paragraaf 3.2.

6.2.3. Beperkingen

Ondanks de relevante inzichten die dit onderzoek biedt, zijn er drie belangrijke beperkingen die in acht moeten worden genomen bij de interpretatie van de resultaten. Allereerst is in dit onderzoek gebruikgemaakt van vertrekintentie en niet van feitelijk vertrekgedrag. Hoewel vertrekintentie een belangrijke indicator van daadwerkelijk vertrek is, vertrekt niet iedereen die de intentie heeft daadwerkelijk. Ook blijft niet iedereen die de intentie heeft om te blijven.

Een tweede beperking betreft de meting van sociaal kapitaal. Hoewel in dit onderzoek een onderscheid is gemaakt tussen cognitief en structureel sociaal kapitaal, ontbreekt een concreet onderscheid in de kwaliteit van en het gevoel bij het sociaal kapitaal. Daardoor wordt invloed van sociaal kapitaal op vertrekintentie niet volledig blootgelegd.

De derde beperking is dat de verklaarde variantie van de modellen relatief beperkt is. Het model met de meeste verklarende variabelen verklaarde 21% van de variantie in vertrekintentie. Dit betekent dat vertrekintentie ook sterk wordt beïnvloed door andere, niet onderzochte factoren.

6.2.4. Implicaties voor beleid

Dit onderzoek biedt handvatten voor beleidsmakers die beleid maken voor dorpen in niet-stedelijke gebieden. Sociaal kapitaal binnen de woonplaats vermindert de vertrekintentie. Om de veroudering in dorpen in niet-stedelijke gebieden tegen te gaan, is het van belang om jonge mensen voor het dorp te behouden. Uit dit onderzoek blijkt dat zowel relatief meer cognitief

sociaal kapitaal binnen het eigen dorp als relatief meer structureel sociaal kapitaal binnen het eigen dorp voor jonge mensen samenhangen met een lagere vertrekintentie. De bevindingen suggereren daarmee dat beleid gericht op het vergroten van sociaal kapitaal binnen dorpen kan bijdragen aan het behoud van jonge inwoners. Als hun sociaal kapitaal binnen het eigen dorp namelijk toeneemt en de hoeveelheid buiten het eigen dorp gelijk blijft, stijgt de relatieve hoeveelheid sociaal kapitaal binnen het eigen dorp. Dit verlaagt vervolgens de verwachte vertrekintentie.

Gemeenten met dorpen in niet-stedelijke gebieden wordt daarom aangeraden om te onderzoeken welke initiatieven er al zijn die jonge inwoners verbinden met andere dorpsbewoners en waar nog kansen liggen om het sociaal kapitaal van jonge mensen binnen het eigen dorp te vergroten. In combinatie hiermee dient er aandacht te zijn voor het evalueren van de effectiviteit van deze initiatieven. Initiatieven dragen immers alleen bij aan het behouden van jonge mensen als ze jonge mensen bereiken én als ze daadwerkelijk de hoeveelheid sociaal kapitaal binnen het eigen dorp vergroten.

Naast de evaluatie van individuele initiatieven wordt aangeraden om de relatieve hoeveelheid sociaal kapitaal binnen het eigen dorp in het algemeen te monitoren. Deze monitoring moet zo worden ingericht dat deze inzicht geeft in de mate waarin het beleid daadwerkelijk bijdraagt aan het terugdringen van het vertrek van jonge inwoners. Op deze manier kan doorlopend gemonitord worden of het beleid effectief is om het doel, het verminderen van het relatief hoge vertrekpercentage van jonge mensen, te bereiken.

6.2.5. Suggesties voor wetenschappelijk vervolgonderzoek

Dit onderzoek biedt aanknopingspunten voor wetenschappelijk vervolgonderzoek. Een eerste logische stap is om toekomstig onderzoek te richten op feitelijk vertrekgedrag in plaats van vertrekintentie. Door gebruik te maken van longitudinale data kan beter worden vastgesteld hoe leeftijd en sociaal kapitaal samenhangen met daadwerkelijk vertrek. Daarmee kunnen de inzichten uit dit onderzoek nauwkeuriger worden getoetst en nog steviger worden onderbouwd.

Ten tweede is het relevant om in toekomstig onderzoek niet alleen aandacht te hebben voor de hoeveelheid sociaal kapitaal, maar ook voor de ervaren kwaliteit ervan en het gevoel erbij. Daarmee kan beter worden vastgesteld welke aspecten van sociaal kapitaal daadwerkelijk doorslaggevend zijn voor de beslissing om wel of niet te vertrekken uit de woonplaats.

Ten slotte is vervolgonderzoek naar de toepasbaarheid van de sociaalemotionele selectiviteitstheorie op het vertrek uit de woonplaats wenselijk. De bevinding dat cognitief sociaal kapitaal juist sterker samenhangt met vertrekintentie bij jongere mensen suggereert dat de theorie niet goed aansluit op deze specifieke gedraging. Het is relevant om verder te onderzoeken waarom de resultaten van dit onderzoek afwijken van de theoretische verwachtingen.

7. Literatuur

- Alguezaui, S., & Filieri, R. (2010). Investigating the role of social capital in innovation: sparse versus dense network. *J. Knowl. Manag.*, 14, 891-909.
- Alwin, D. F. (1997). Feeling Thermometers versus 7-Point Scales. *Sociological Methods & Research*, 25(3), 318–340. <https://doi.org/10.1177/0049124197025003003>
- Austin, P. C., Eekhout, I., & Van Buuren, S. (2024). Evaluating the median p -value method for assessing the statistical significance of tests when using multiple imputation. *Journal Of Applied Statistics*, 52(6), 1161–1176. <https://doi.org/10.1080/02664763.2024.2418473>
- Azur, M. J., Stuart, E. A., Frangakis, C., & Leaf, P. J. (2011). Multiple imputation by chained equations: what is it and how does it work? *International Journal Of Methods in Psychiatric Research*, 20(1), 40–49. <https://doi.org/10.1002/mpr.329>
- Bain, K., & Hicks, N. (1998). Building social capital and reaching out to excluded groups: the challenge of partnerships. *World Bank*, 15(1), 13-25.
- Bourdieu, P. 1986. "The Forms of Capital." Pp. 241–58 in *Handbook of theory and research for the sociology of education*, edited by J. G. Richardson. New York: Greenwood Press.
- Büchel, K., Ehrlich, M. V., Puga, D., & Viladecans-Marsal, E. (2020). Calling from the outside: The role of networks in residential mobility. *Journal Of Urban Economics*, 119, 103277. <https://doi.org/10.1016/j.jue.2020.103277>
- Carifio, J., & Perla, R. (2008). Resolving the 50-year debate around using and misusing Likert scales. *Medical Education*, 42(12), 1150–1152. <https://doi.org/10.1111/j.1365-2923.2008.03172.x>
- Carstensen, L. L. (1995). Evidence for a Life-Span Theory of Socioemotional Selectivity. *Current Directions in Psychological Science*, 4(5), 151–156. <https://doi.org/10.1111/1467-8721.ep11512261>
- Centraal Bureau voor de Statistiek. (z.d.-a). *Bevolking op 1 januari en gemiddeld; geslacht, leeftijd en regio*. <https://opendata.cbs.nl/statline/#/CBS/nl/dataset/03759ned/table?dl=C430B>
- Centraal Bureau voor de Statistiek. (z.d.-b). *Gebieden in Nederland 2024*. <https://opendata.cbs.nl/#/CBS/nl/dataset/85755NED/table?dl=63525>
- Centraal Bureau voor de Statistiek. (z.d.-c). *Jongeren en ouderen per gemeente*. Centraal Bureau Voor de Statistiek. <https://www.cbs.nl/nl-nl/visualisaties/dashboard-bevolking/regionaal/jongeren-en-ouderen>
- Centraal Bureau voor de Statistiek. (z.d.-d). *Leeftijdsverdeling*. Centraal Bureau Voor de Statistiek. <https://www.cbs.nl/nl-nl/visualisaties/dashboard-bevolking/leeftijd/bevolking>
- Centraal Bureau voor de Statistiek. (z.d.-e). *Overlijden*. Centraal Bureau Voor de Statistiek. <https://www.cbs.nl/nl-nl/visualisaties/dashboard-bevolking/levensloop/verweduwen>
- Centraal Bureau voor de Statistiek. (z.d.-f). *Sociaal-economische status; scores per wijk en buurt, regio-indeling 2023*. <https://opendata.cbs.nl/#/CBS/nl/dataset/85900NED/table?dl=C42FE>
- Centraal Bureau voor de Statistiek. (z.d.-g). *Standaard Onderwijsindeling 2021*. Centraal Bureau Voor de Statistiek. <https://www.cbs.nl/nl-nl/onze-diensten/methoden/classificaties/onderwijs-en-beroepen/standaard-onderwijsindeling--soi--/standaard-onderwijsindeling-2021>

- Centraal Bureau voor de Statistiek. (2024). Kerncijfers wijken en buurten 2024. <https://opendata.cbs.nl/#/CBS/nl/dataset/85984NED/table?dl=C4312>
- Civelek, M. E. (2018). Essentials of Structural Equation Modeling. *Zea Books*. <https://doi.org/10.13014/k2sj1hr5>
- Cohen, S. A., & Greaney, M. L. (2023). Aging in Rural Communities. *Current epidemiology reports*, 10(1), 1–16. <https://doi.org/10.1007/s40471-022-00313-9>
- Coleman, J. S. (1988). Social Capital in the Creation of Human Capital. *American Journal Of Sociology*, 94, S95–S120. <https://doi.org/10.1086/228943>
- David, Q., Janiak, A., & Wasmer, E. (2010). Local social capital and geographical mobility. *Journal Of Urban Economics*, 68(2), 191–204. <https://doi.org/10.1016/j.jue.2010.04.003>
- de Groot, C., Manting, D., & Mulder, C. H. (2011). Intentions to Move and Actual Moving Behaviour in The Netherlands. *Housing Studies*, 26(3), 307–328. <https://doi.org/10.1080/02673037.2011.542094>
- de Haas, H. (2021). A theory of migration: the aspirations-capabilities framework. *Comparative Migration Studies*, 9(1). <https://doi.org/10.1186/s40878-020-00210-4>
- Ejlskov, L., Mortensen, R. N., Overgaard, C., Christensen, L. B., Vardinghus-Nielsen, H., Kræmer, S. R. J., Wissenberg, M., Hansen, S. M., Torp-Pedersen, C., & Hansen, C. D. (2014). Individual social capital and survival: a population study with 5-year follow-up. *BMC Public Health*, 14(1). <https://doi.org/10.1186/1471-2458-14-1025>
- Eriksson, M. (2011). Social capital and health – implications for health promotion. *Global Health Action*, 4(1), 5611. <https://doi.org/10.3402/gha.v4i0.5611>
- Fukuyama, F. (2001). Social capital, civil society and development. *Third World Quarterly*, 22(1), 7–20. <https://doi.org/10.1080/713701144>
- Giannone, E., Miyauchi, Y., Paixão, N., Pang, X., & Suzuki, Y. (2024). Living in a Ghost Town: The Geography of Depopulation and Aging. *CREI Working Paper*. <https://crei.cat/wp-content/uploads/2025/03/TheGeographyofAgingandDepopulation-6-2.pdf>
- Haartsen, T., & Venhorst, V. (2010). Planning for decline: Anticipating on population decline in the Netherlands. *Tijdschrift voor Economische en Sociale Geografie*, 101(2), 218–227. <https://doi.org/10.1111/j.1467-9663.2010.00597.x>
- Halpern, D. (2005). *Social Capital*. Polity.
- Hodge, D. R., & Gillespie, D. F. (2007). Phrase completion scales. *Journal Of Social Service Research*, 33(4), 1–12. https://doi.org/10.1300/j079v33n04_01
- Kwak, C., & Clayton-Matthews, A. (2002). Multinomial logistic regression. *Nursing Research*, 51(6), 404–410. <https://doi.org/10.1097/00006199-200211000-00009>
- Krishna, A., & Shrader, E. (2000). Cross-cultural measures of social capital: a tool and results from India and Panama. *Social capital initiative working paper*, 21, 1-118.
- Li, K. H., Raghunathan, T. E., & Rubin, D. B. (1991). Large-Sample Significance Levels from Multiply Imputed Data Using Moment-Based Statistics and an F Reference Distribution. *Journal Of The American Statistical Association*, 86(416), 1065. <https://doi.org/10.2307/2290525>
- Liu, X., Bai, H., & Koirala, H. (2018). Multiple Imputation for Missing Data Analysis in Proportional Odds Models for Ordinal Response Variables. *General Linear Model Journal*, 44(1), 1–10. <https://doi.org/10.31523/glmj.044001.001>

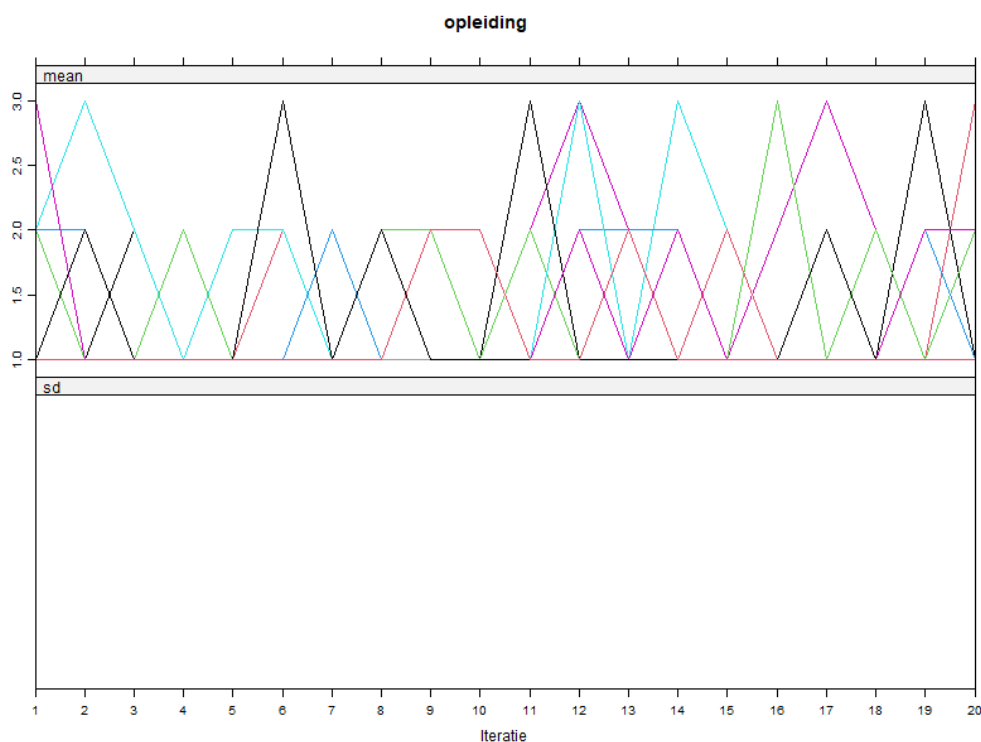
- Long, L. (1992). Changing Residence: Comparative Perspectives on its Relationship to Age, Sex, and Marital Status. *Population Studies*, 46(1), 141–158.
<https://doi.org/10.1080/0032472031000146056>
- Ministerie van Sociale Zaken en Werkgelegenheid. (2024). *Gematigde groei – rapport van de Staatscommissie Demografische Ontwikkelingen 2050*. Rapport | Rijksoverheid.nl.
<https://www.rijksoverheid.nl/documenten/rapporten/2024/01/15/gematigde-groei-rapport-van-de-staatscommissie-demografische-ontwikkelingen-2050>
- Ohnmacht, T. (2009). Social-Activity Travel: Do the ‘Strong-Tie Relationships’ of a Person Exist in the Same Community? The Case of Switzerland. *Environment And Planning A Economy And Space*, 41(12), 3003–3022. <https://doi.org/10.1068/a41396>
- Putnam, R. D. (1995). Bowling Alone: America’s declining social capital. *Journal Of Democracy*, 6(1), 65–78. <https://doi.org/10.1353/jod.1995.0002>
- Putnam, R. D. (2000). *Bowling alone: The collapse and revival of American community*. Touchstone Books/Simon & Schuster. <https://doi.org/10.1145/358916.361990>
- Sørensen, J. F. L. (2014). Rural–Urban Differences in Bonding and Bridging Social Capital. *Regional Studies*, 50(3), 391–410. <https://doi.org/10.1080/00343404.2014.918945>
- Thiede, B. C., Brown, D. L., Sanders, S. R., Glasgow, N., & Kulcsar, L. J. (2016). A Demographic Deficit? Local Population Aging and Access to Services in Rural America, 1990–2010. *Rural Sociology*, 82(1), 44–74. <https://doi.org/10.1111/ruso.12117>
- van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate Imputation by Chained Equations in R. *Journal Of Statistical Software*, 45(3). <https://doi.org/10.18637/jss.v045.i03>
- van Buuren, S. (2018). Flexible Imputation of Missing Data, Second Edition. In Chapman and Hall/CRC eBooks. <https://doi.org/10.1201/9780429492259>
- Verwey-Jonker Instituut. (2022, 14 december). *Het platteland in beeld*.
<https://www.verwey-jonker.nl/publicatie/het-platteland-in-beeld/>
- White, I. R., Royston, P., & Wood, A. M. (2010). Multiple imputation using chained equations: Issues and guidance for practice. *Statistics in Medicine*, 30(4), 377–399.
<https://doi.org/10.1002/sim.4067>
- Wu, H., & Leung, S. A. (2017). Can Likert Scales be Treated as Interval Scales?—A Simulation Study. *Journal Of Social Service Research*, 43(4), 527–532.
<https://doi.org/10.1080/01488376.2017.1329775>

Bijlage A: grafieken kwaliteit van imputaties

In deze bijlage vind je alle grafieken die zijn gebruikt om de kwaliteit van de imputaties te evalueren. Eerst worden de convergentieplots getoond en daarna de stripplots.

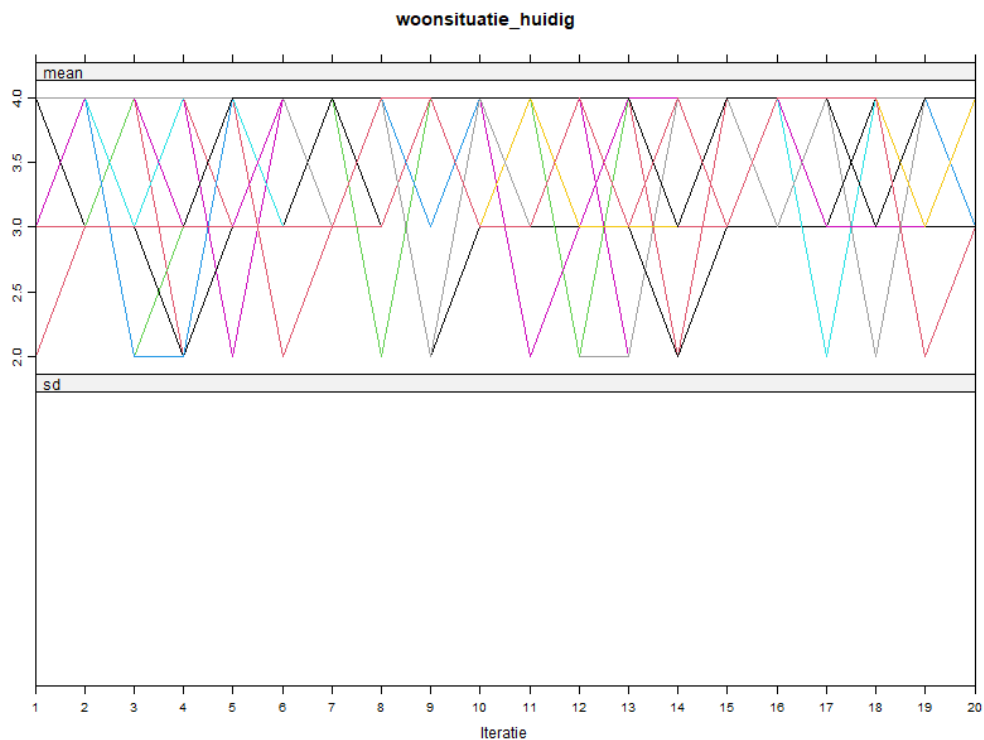
Convergentietoetsing van geïmputeerde samengestelde variabele en vragenlijstitems

Hieronder vind je de convergentieplots per samengestelde variabele of vragenlijstitem waarvoor is geïmputeerd. Bij sommige plots zie je geen standaarddeviatie. Dat komt doordat er maar één ontbrekende waarde was. Per iteratie wordt dan maar één waarde geïmputeerd. Omdat er geen verschil tussen waarden is, kun je geen standaarddeviatie berekenen. In de plots zijn de variabele en de vragenlijstitems aangeduid met hun R-namen zoals gebruikt in de analyses. Zie bijlage D voor een overzicht van deze R-namen en de bijbehorende variabele en vragenlijstitems.



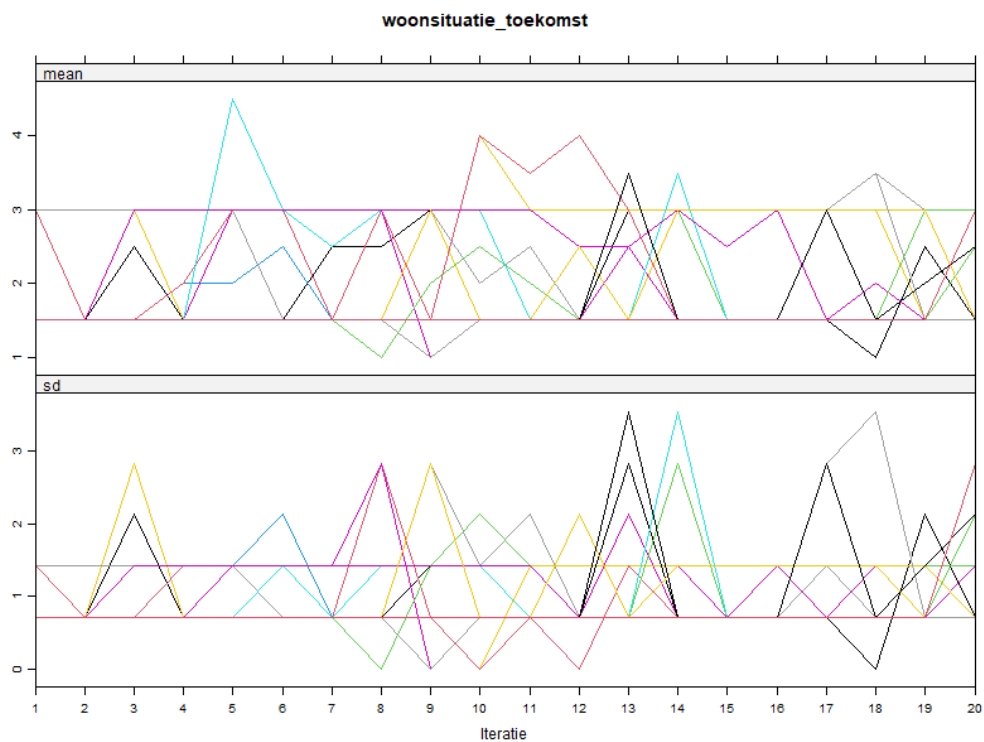
Figuur A1

Convergentieplot voor de samengestelde variabele 'opleiding' met gemiddelde (boven) en standaarddeviatie (onder) van geïmputeerde waarden per iteratie. Elke keten representeert één geïmputeerde dataset.



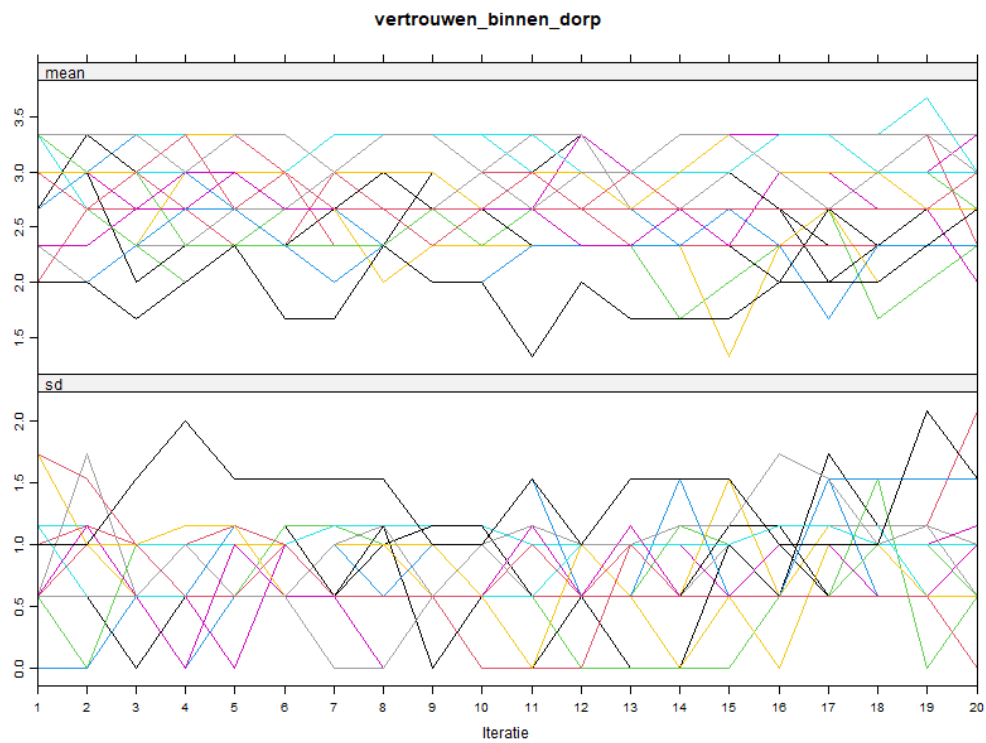
Figuur A2

Convergentieplot voor het vragenlijstitem 'woonsituatie_huidig' met gemiddelde (boven) en standaarddeviatie (onder) van geïmputeerde waarden per iteratie. Elke keten representeert één geïmputeerde dataset.



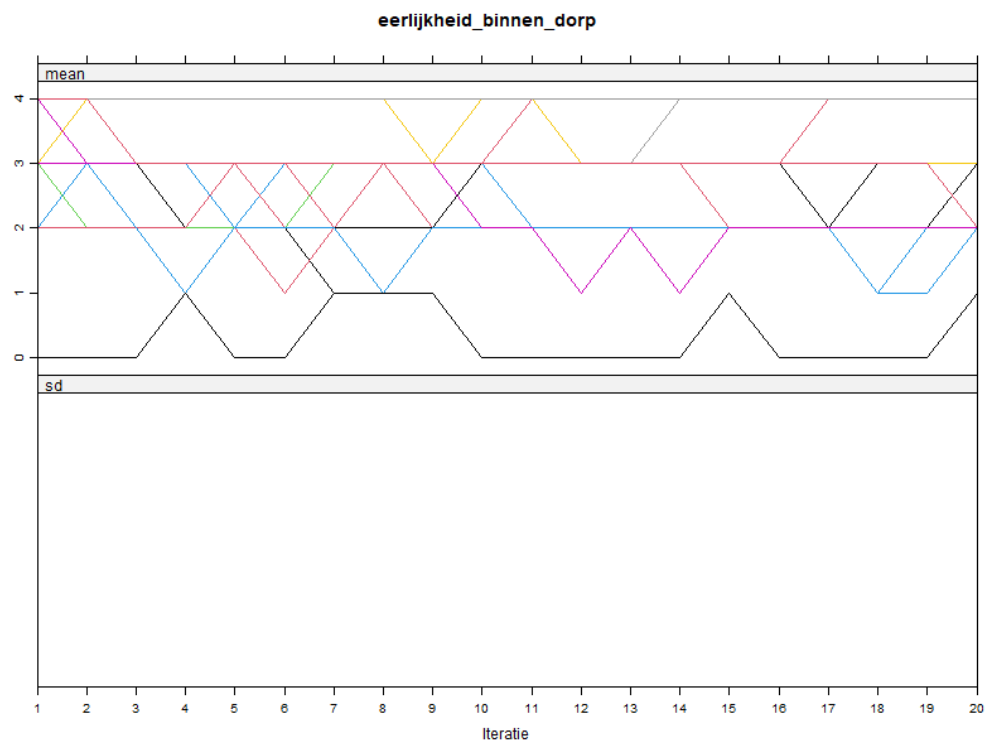
Figuur A3

Convergentieplot voor het vragenlijstitem 'woonsituatie_toekomst' met gemiddelde (boven) en standaarddeviatie (onder) van geïmputeerde waarden per iteratie. Elke keten representeert één geïmputeerde dataset.



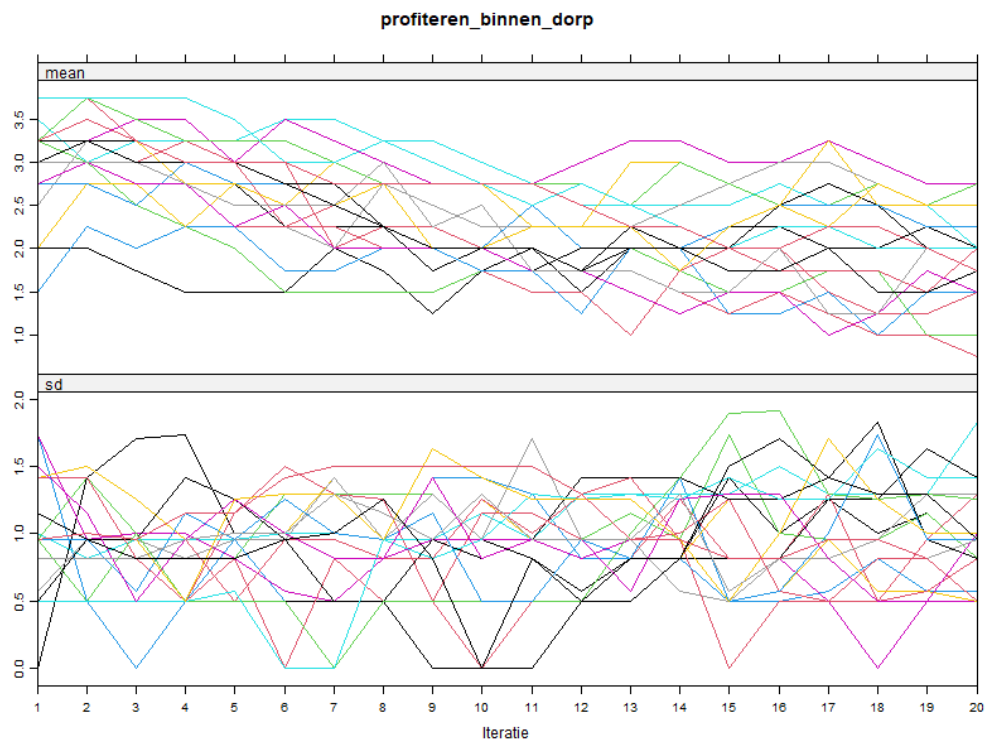
Figuur A4

Convergentieplot voor het vragenlijstitem 'vertrouwen_binnen_dorp' met gemiddelde (boven) en standaarddeviatie (onder) van geïmputeerde waarden per iteratie. Elke keten representeert één geïmputeerde dataset.



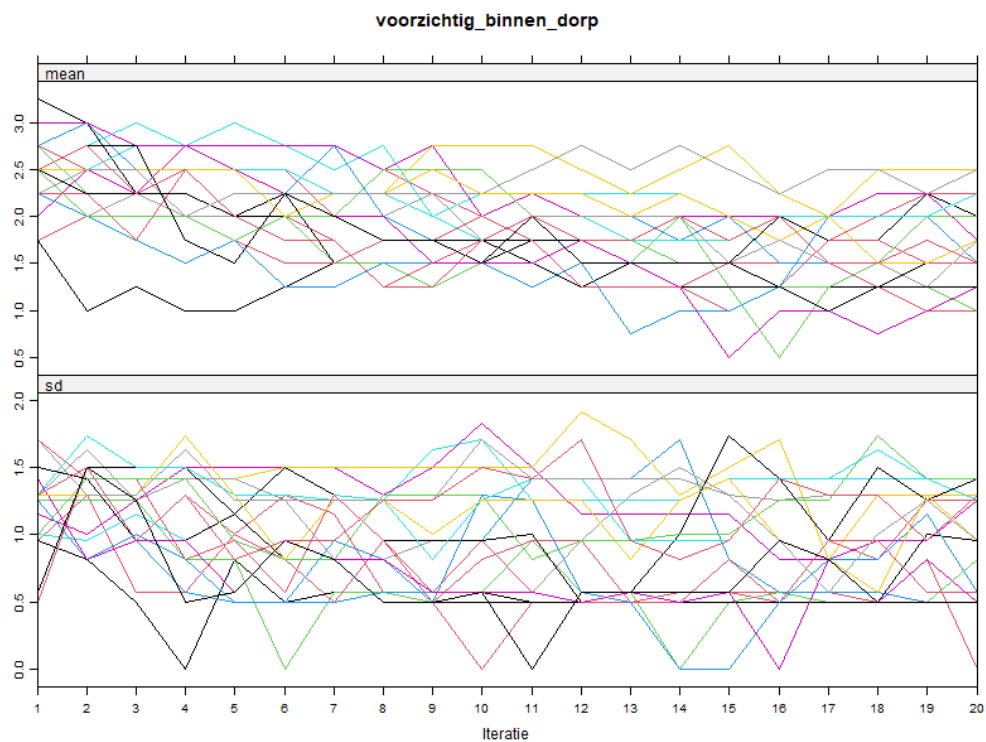
Figuur A5

Convergentieplot voor het vragenlijstitem 'eerlijkheid_binnen_dorp' met gemiddelde (boven) en standaarddeviatie (onder) van geïmputeerde waarden per iteratie. Elke keten representeert één geïmputeerde dataset.



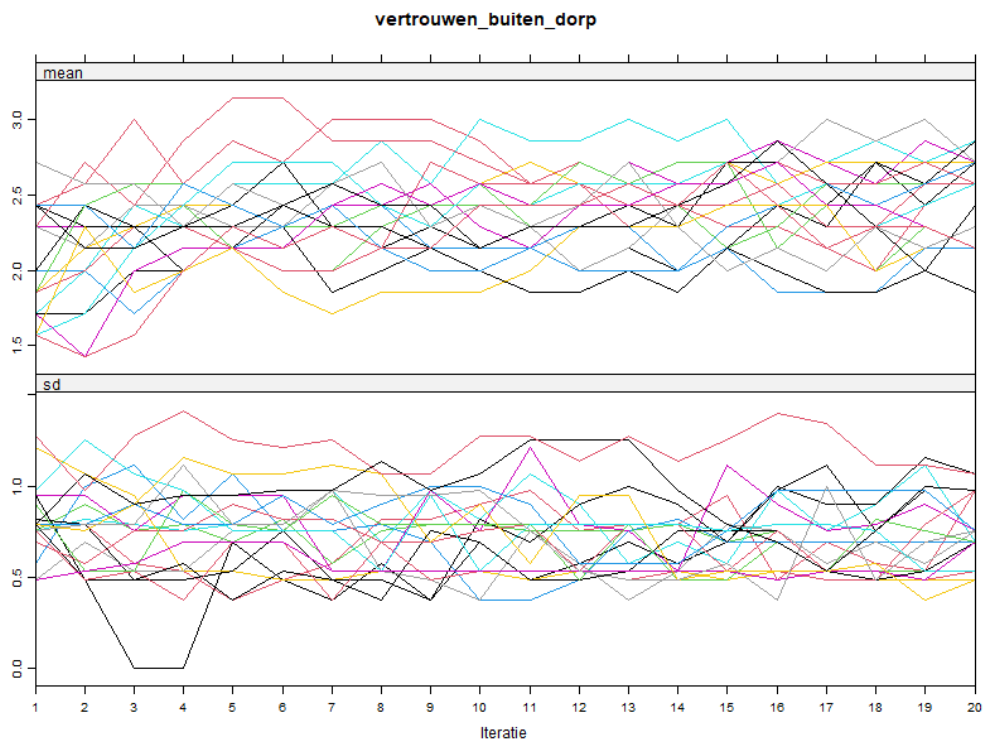
Figuur A6

Convergentieplot voor het vragenlijstitem 'profiteren_binnen_dorp' met gemiddelde (boven) en standaarddeviatie (onder) van geïmputeerde waarden per iteratie. Elke keten representeert één geïmputeerde dataset.



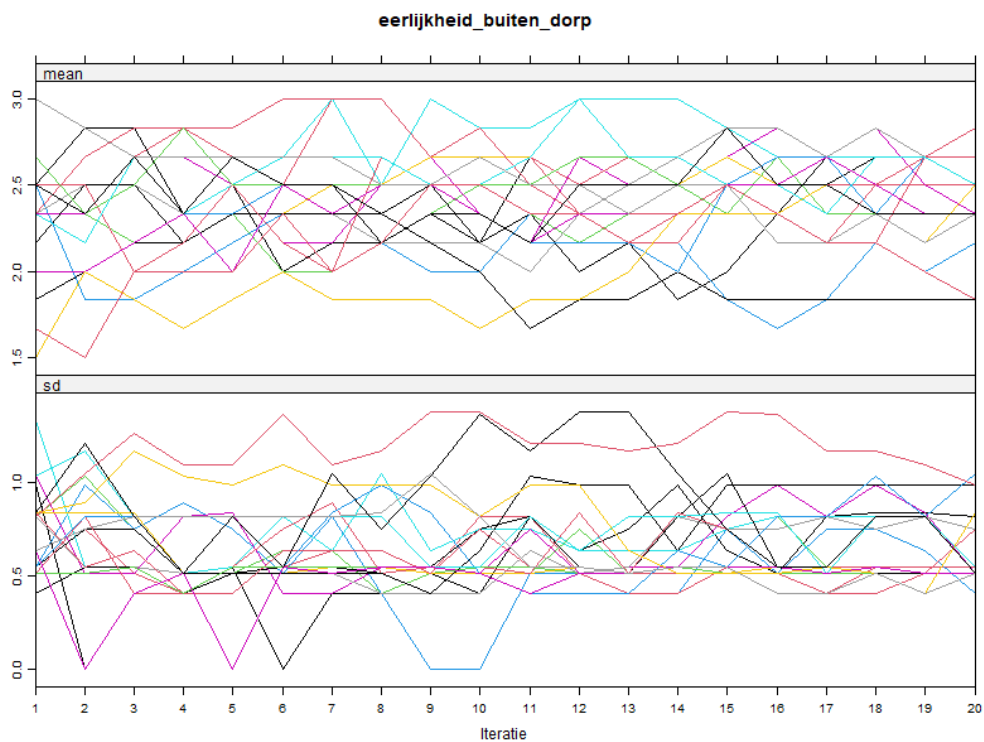
Figuur A7

Convergentieplot voor het vragenlijstitem 'voorzichtig_binnen_dorp' met gemiddelde (boven) en standaarddeviatie (onder) van geïmputeerde waarden per iteratie. Elke keten representeert één geïmputeerde dataset.



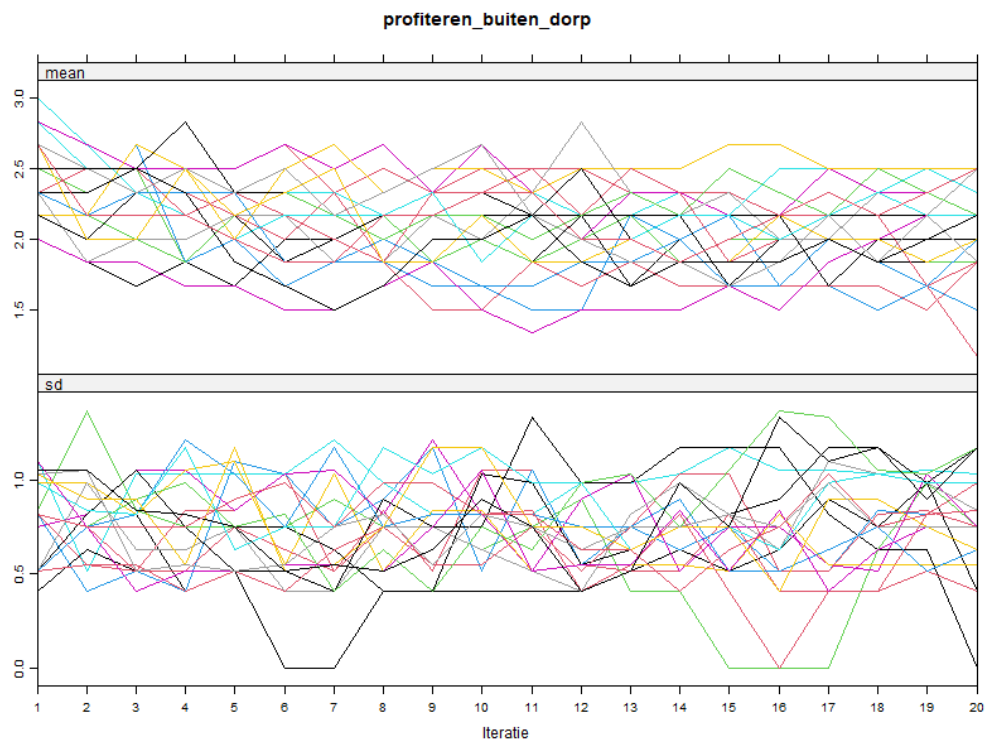
Figuur A8

Convergentieplot voor het vragenlijstitem 'vertrouwen_buiten_dorp' met gemiddelde (boven) en standaarddeviatie (onder) van geïmputeerde waarden per iteratie. Elke keten representeert één geïmputeerde dataset.



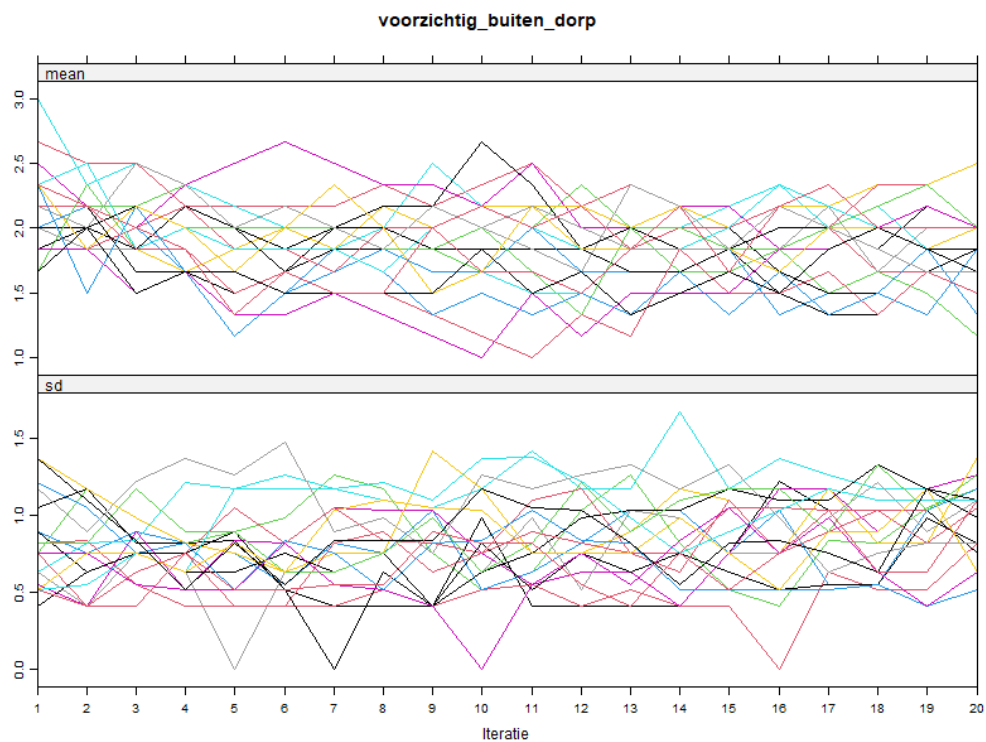
Figuur A9

Convergentieplot voor het vragenlijstitem 'eerlijkheid_buiten_dorp' met gemiddelde (boven) en standaarddeviatie (onder) van geïmputeerde waarden per iteratie. Elke keten representeert één geïmputeerde dataset.



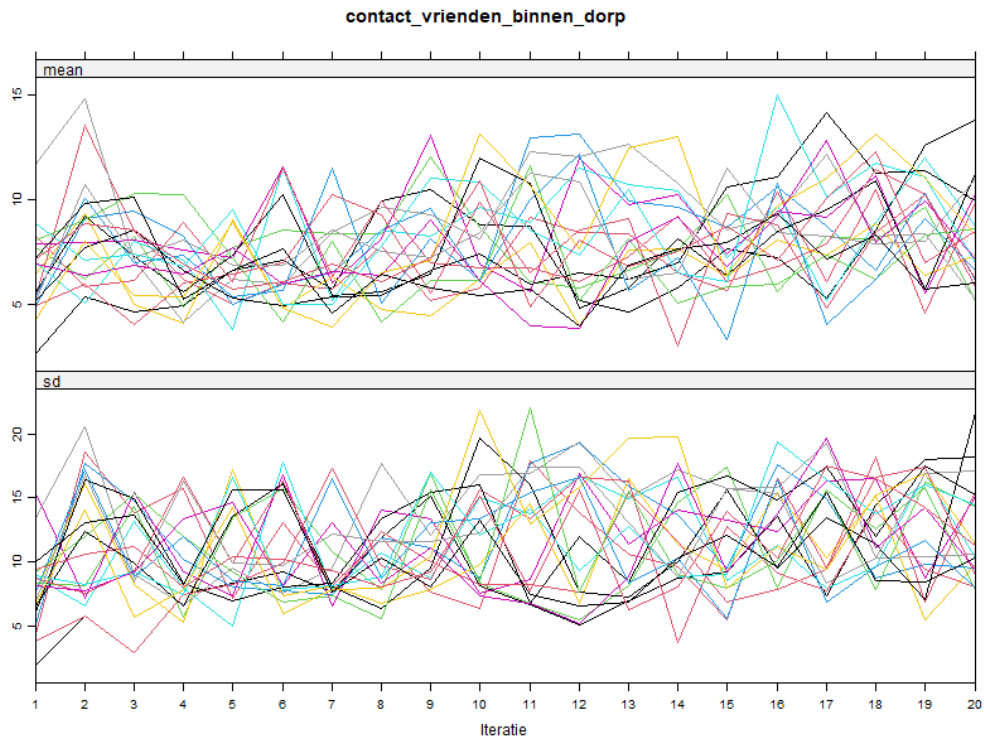
Figuur A10

Convergentieplot voor het vragenlijstitem 'profiteren_buiten_dorp' met gemiddelde (boven) en standaarddeviatie (onder) van geïmputeerde waarden per iteratie. Elke keten representeert één geïmputeerde dataset.



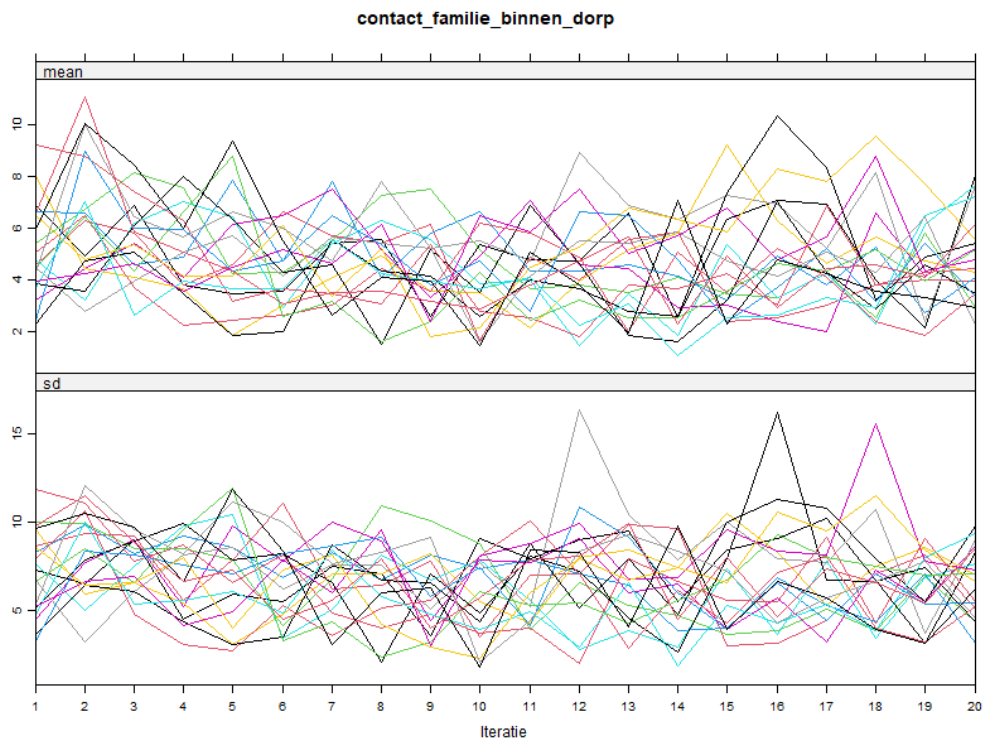
Figuur A11

Convergentieplot voor het vragenlijstitem 'voorzichtig_buiten_dorp' met gemiddelde (boven) en standaarddeviatie (onder) van geïmputeerde waarden per iteratie. Elke keten representeert één geïmputeerde dataset.



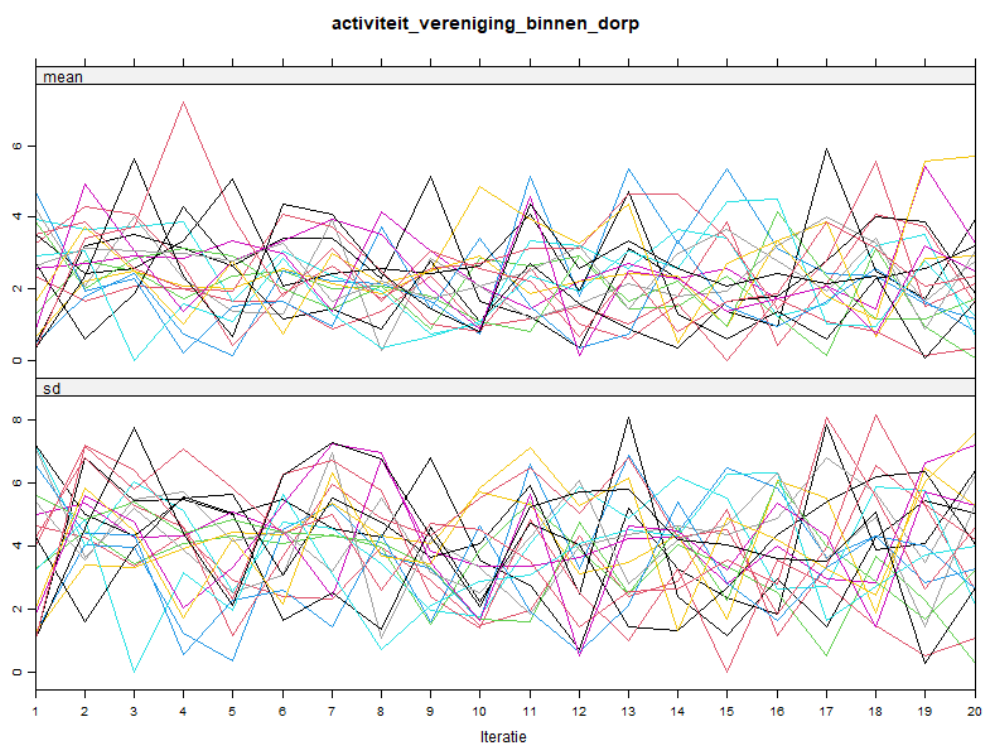
Figuur A12

Convergentieplot voor het vragenlijstitem 'contact_vrienden_binnen_dorp' met gemiddelde (boven) en standaarddeviatie (onder) van geïmputeerde waarden per iteratie. Elke keten representeert één geïmputeerde dataset.



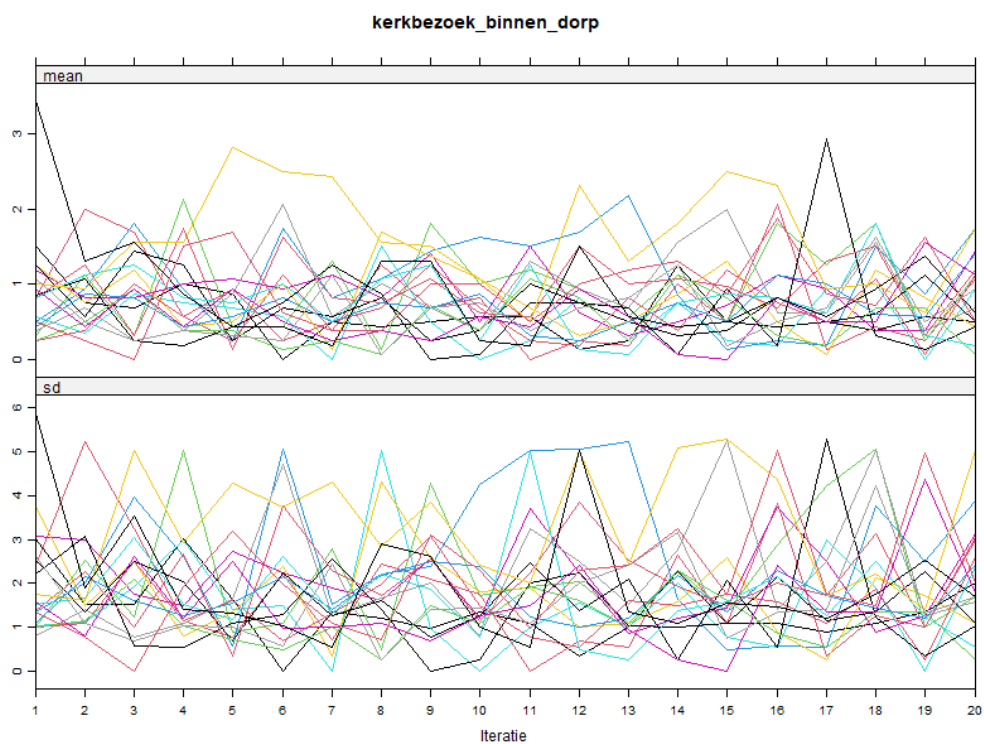
Figuur A13

Convergentieplot voor het vragenlijstitem 'contact_familie_binnen_dorp' met gemiddelde (boven) en standaarddeviatie (onder) van geïmputeerde waarden per iteratie. Elke keten representeert één geïmputeerde dataset.



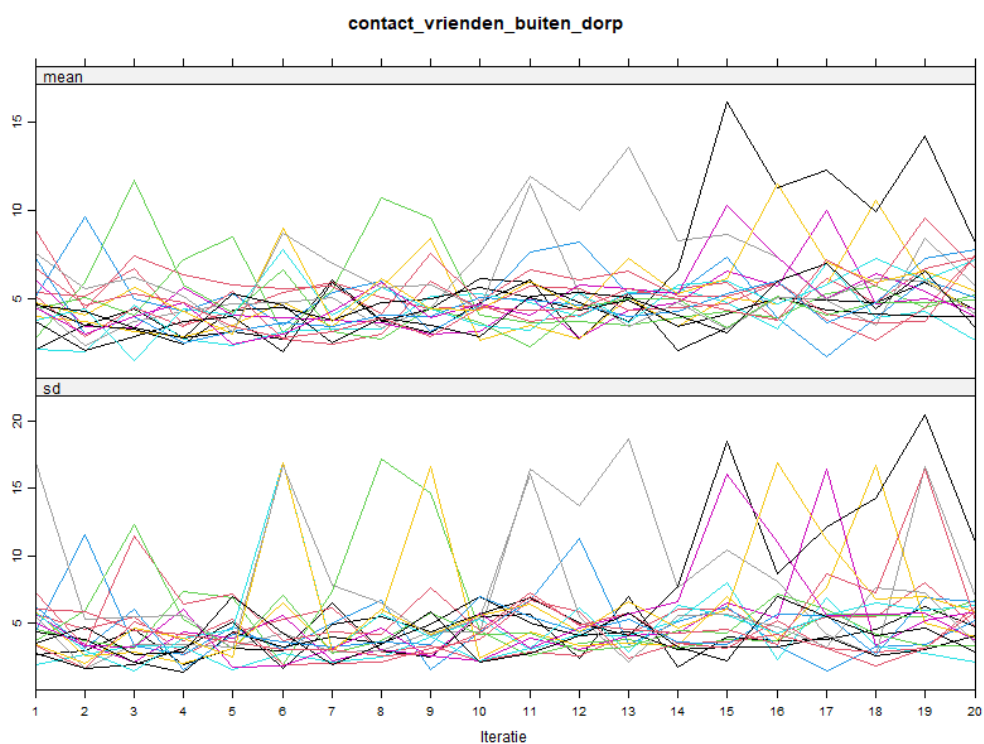
Figuur A14

Convergentieplot voor het vragenlijstitem 'activiteit_vereniging_binnen_dorp' met gemiddelde (boven) en standaarddeviatie (onder) van geïmputeerde waarden per iteratie. Elke keten representeert één geïmputeerde dataset.



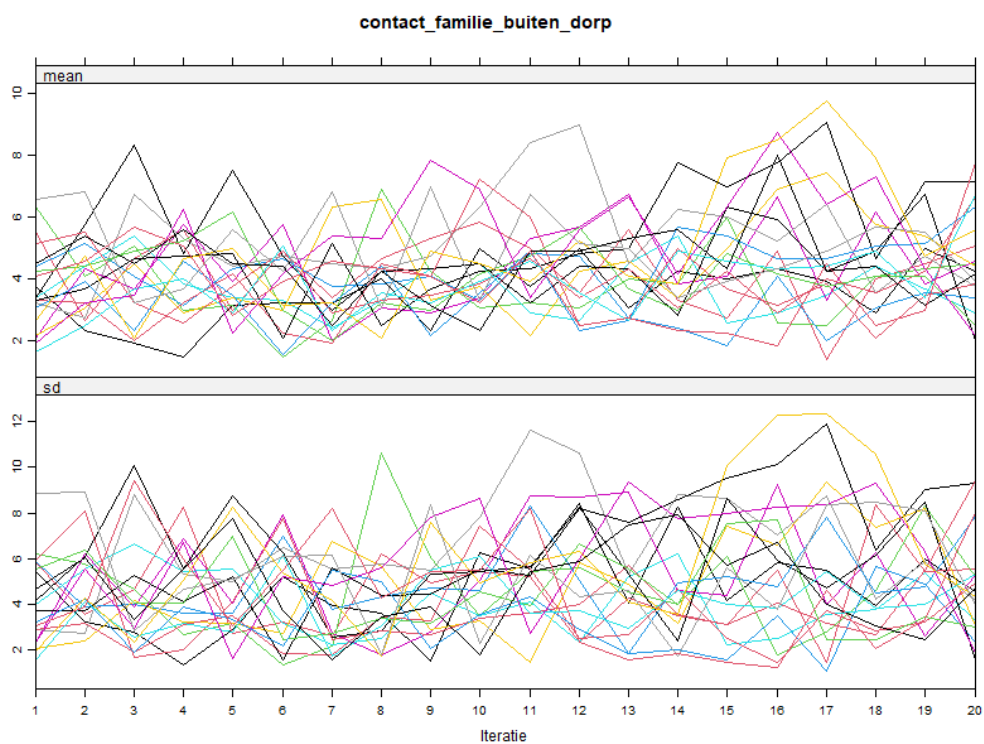
Figuur A15

Convergentieplot voor het vragenlijstitem 'kerkbezoek_binnen_dorp' met gemiddelde (boven) en standaarddeviatie (onder) van geïmputeerde waarden per iteratie. Elke keten representeert één geïmputeerde dataset.



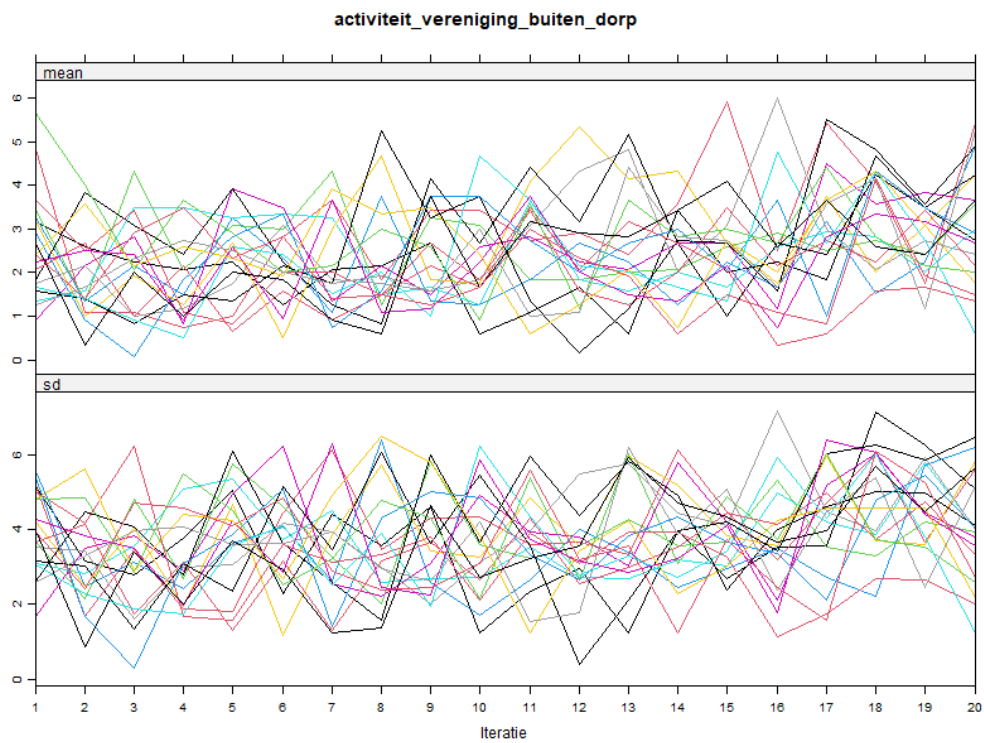
Figuur A16

Convergentieplot voor het vragenlijstitem 'contact_vrienden_buiten_dorp' met gemiddelde (boven) en standaarddeviatie (onder) van geïmputeerde waarden per iteratie. Elke keten representeert één geïmputeerde dataset.



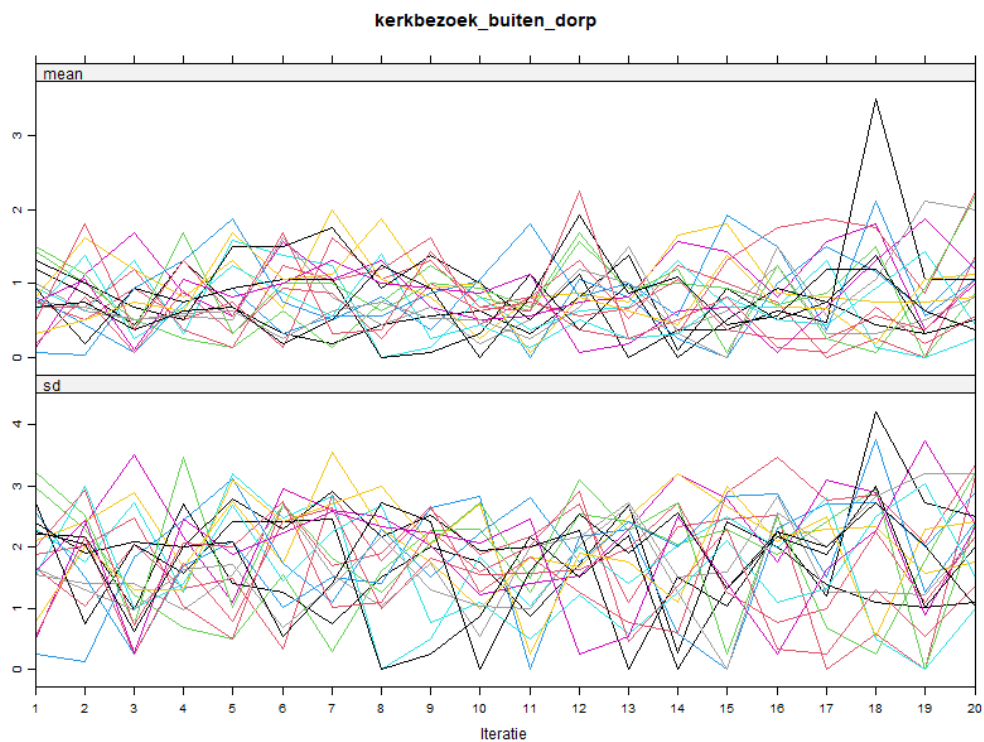
Figuur A17

Convergentieplot voor het vragenlijstitem 'contact_familie_buiten_dorp' met gemiddelde (boven) en standaarddeviatie (onder) van geïmputeerde waarden per iteratie. Elke keten representeert één geïmputeerde dataset.



Figuur A18

Convergentieplot voor het vragenlijstitem 'activiteit_vereniging_buiten_dorp' met gemiddelde (boven) en standaarddeviatie (onder) van geïmputeerde waarden per iteratie. Elke keten representeert één geïmputeerde dataset.

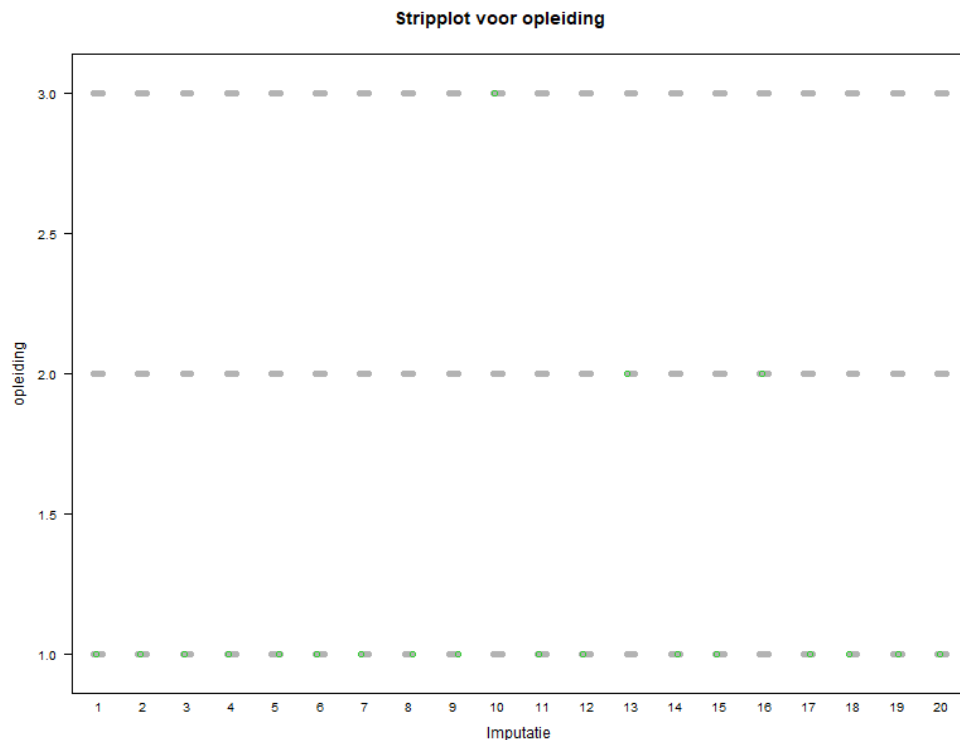


Figuur A19

Convergentieplot voor het vragenlijstitem 'kerkbezoek_buiten_dorp' met gemiddelde (boven) en standaarddeviatie (onder) van geïmputeerde waarden per iteratie. Elke keten representeert één geïmputeerde dataset.

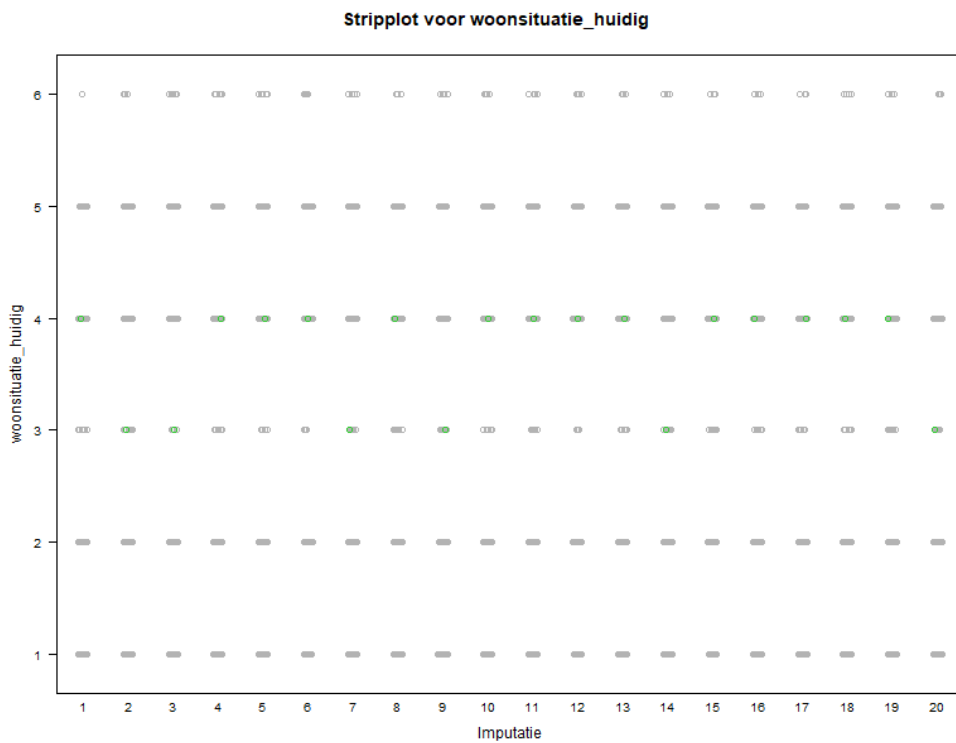
Plausibiliteitstoetsing van geïmputeerde samengestelde variabele en vragenlijstitems

Hieronder vind je de stripplots die zijn gegenereerd per samengestelde variabele of vragenlijstitem waarvoor is geïmputeerd. In de plots zijn de variabele en de vragenlijstitems aangeduid met hun R-namen zoals gebruikt in de analyses. Zie bijlage D voor een overzicht van deze R-namen en de bijbehorende variabele en vragenlijstitems.



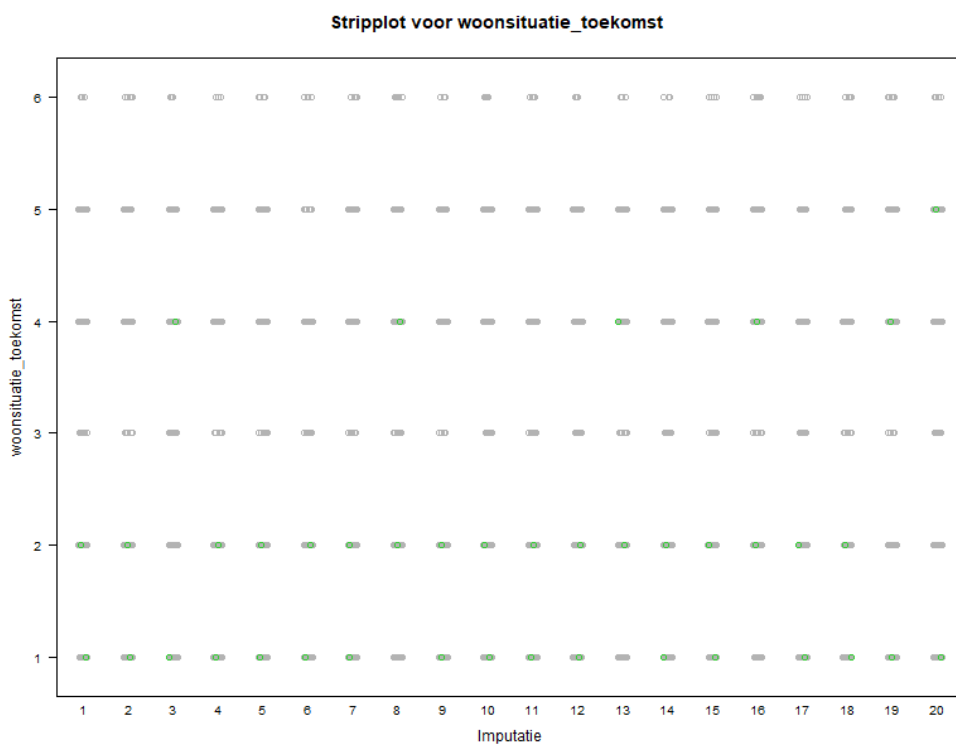
Figuur A20

Stripplot voor de samengestelde variabele 'opleiding' met geobserveerde (grijs) en geïmputeerde (groen) waarden per geïmputeerde dataset



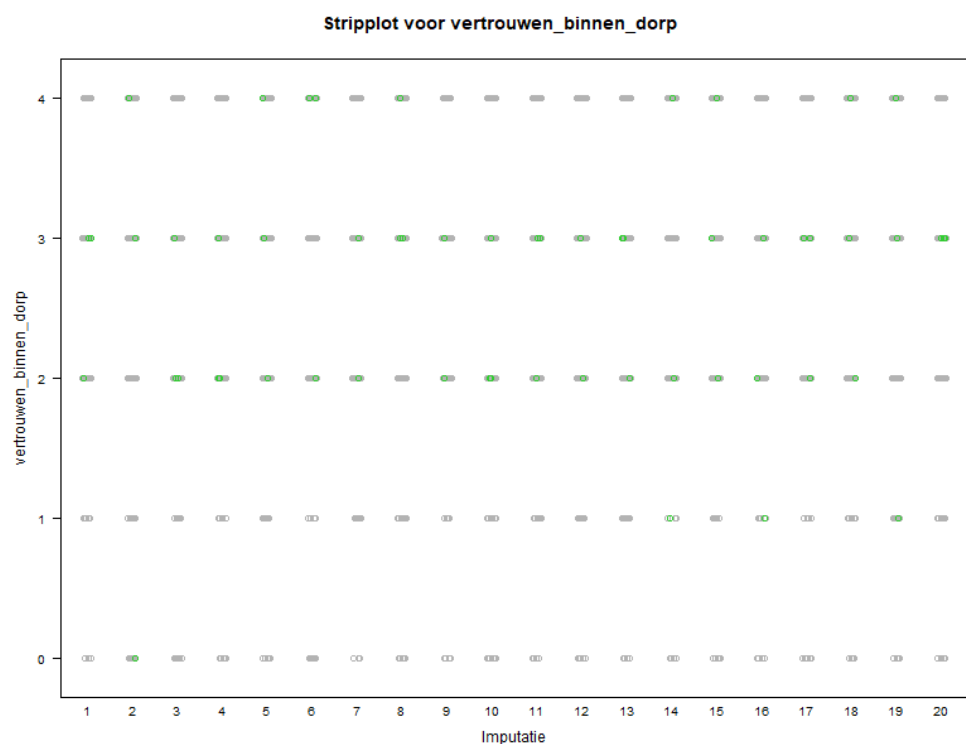
Figuur A21

Stripplot voor het vragenlijstitem 'woonsituatie_huidig' met geobserveerde (grijs) en geïmputeerde (groen) waarden per geïmputeerde dataset



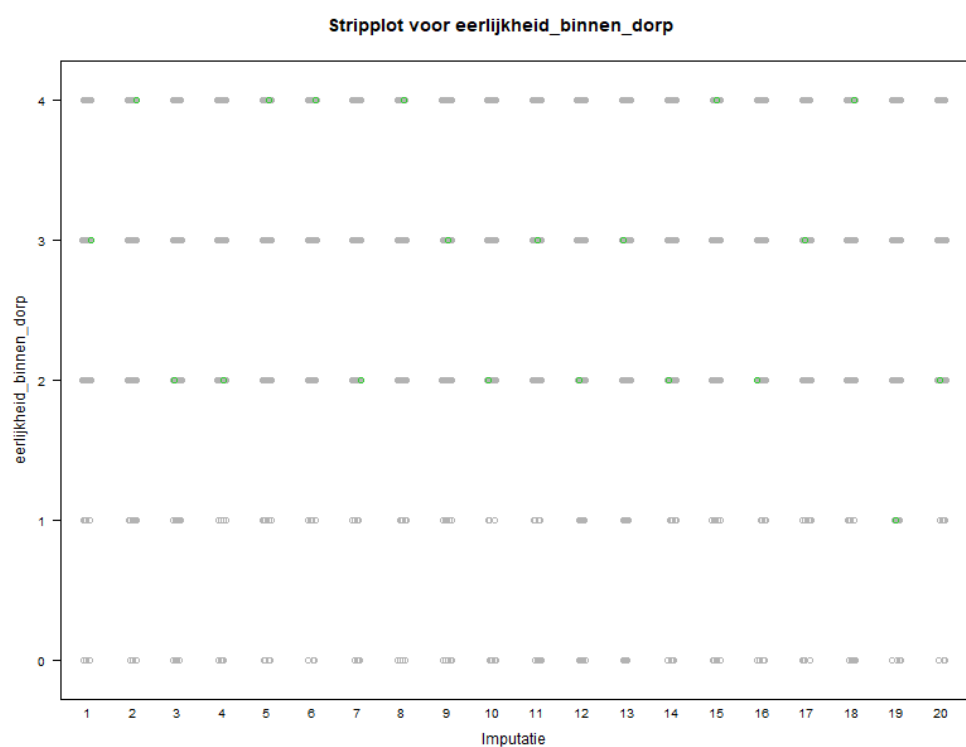
Figuur A22

Stripplot voor het vragenlijstitem 'woonsituatie_toekomst' met geobserveerde (grijs) en geïmputeerde (groen) waarden per geïmputeerde dataset



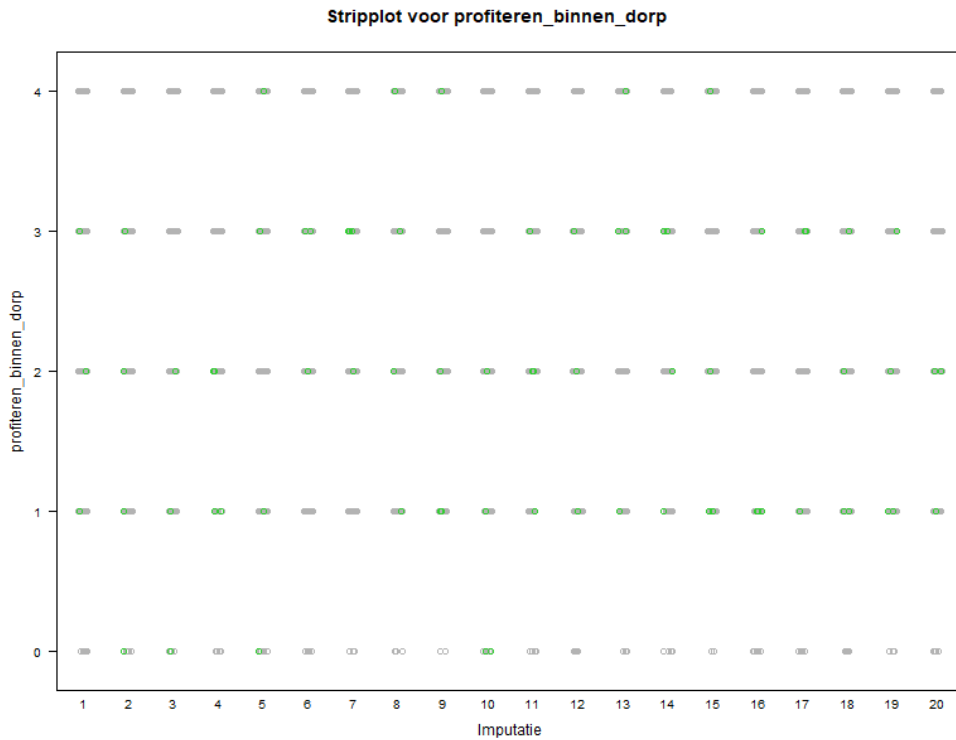
Figuur A23

Stripplot voor het vragenlijstitem 'vertrouwen_binnen_dorp' met geobserveerde (grijs) en geïmputeerde (groen) waarden per geïmputeerde dataset



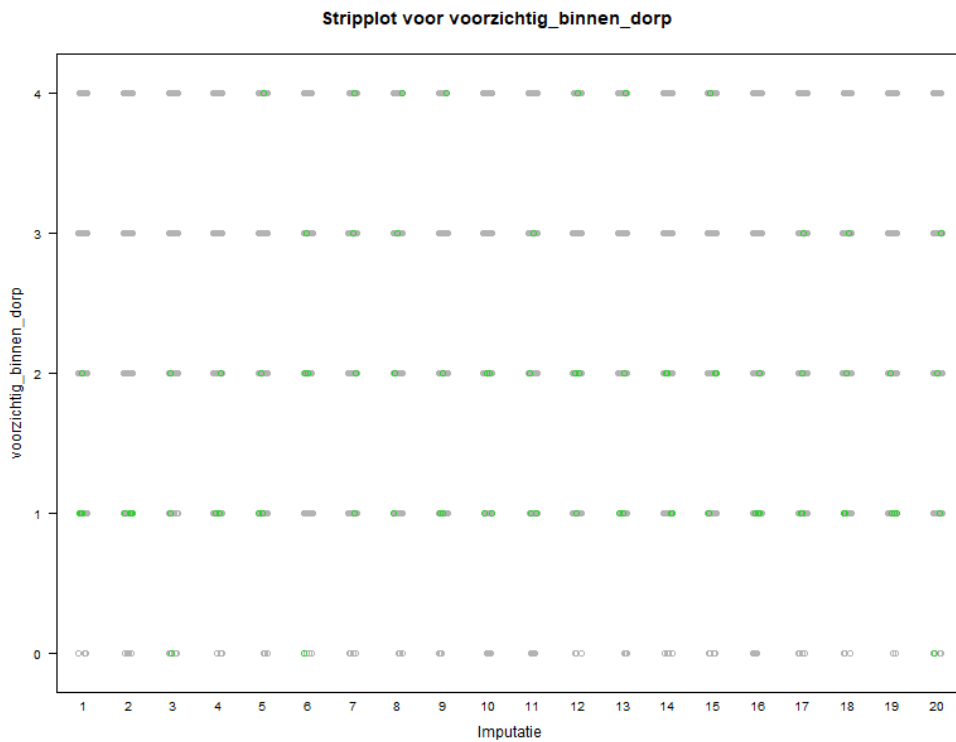
Figuur A24

Stripplot voor het vragenlijstitem 'eerlijkheid_binnen_dorp' met geobserveerde (grijs) en geïmputeerde (groen) waarden per geïmputeerde dataset



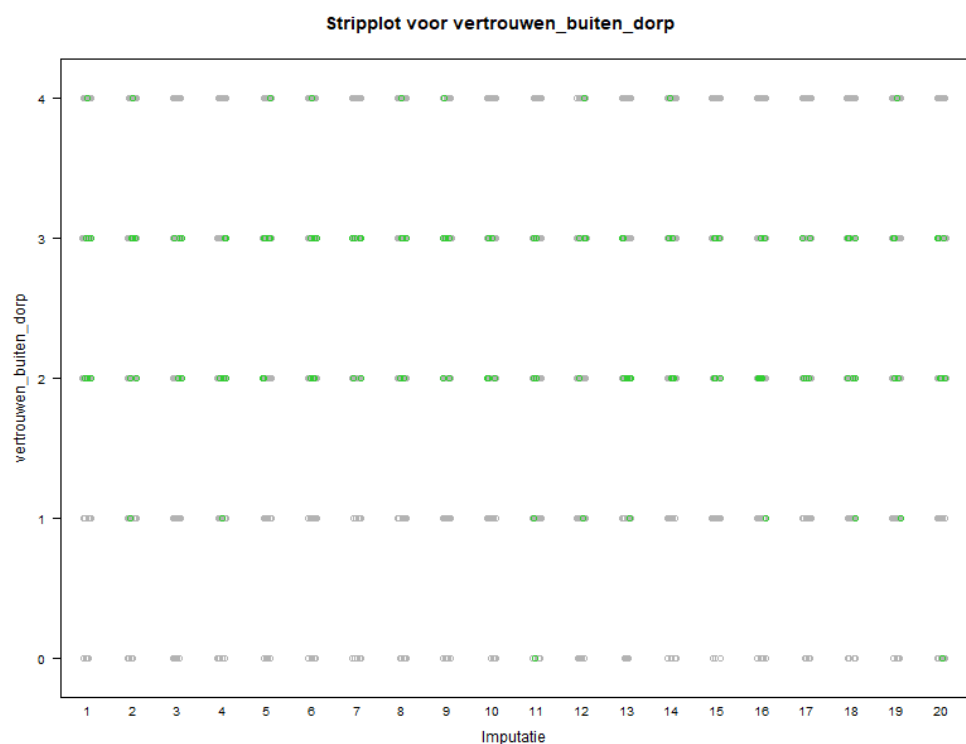
Figuur A25

Stripplot voor het vragenlijstitem 'profiteren_binnen_dorp' met geobserveerde (grijs) en geïmputeerde (groen) waarden per geïmputeerde dataset



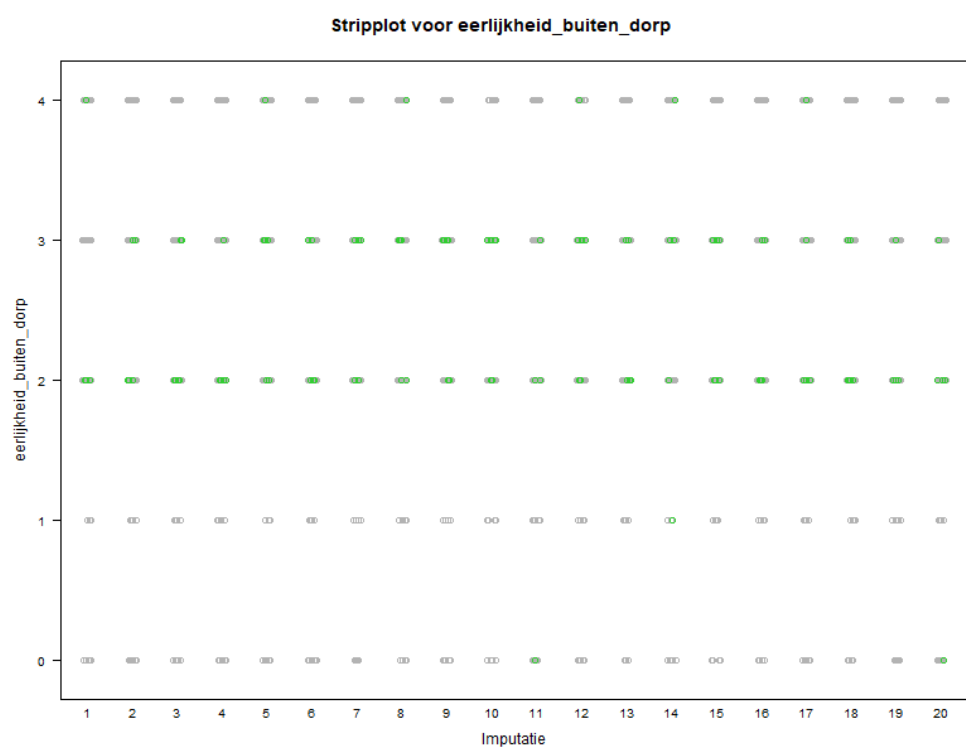
Figuur A26

Stripplot voor het vragenlijstitem 'voorzichtig_binnen_dorp' met geobserveerde (grijs) en geïmputeerde (groen) waarden per geïmputeerde dataset



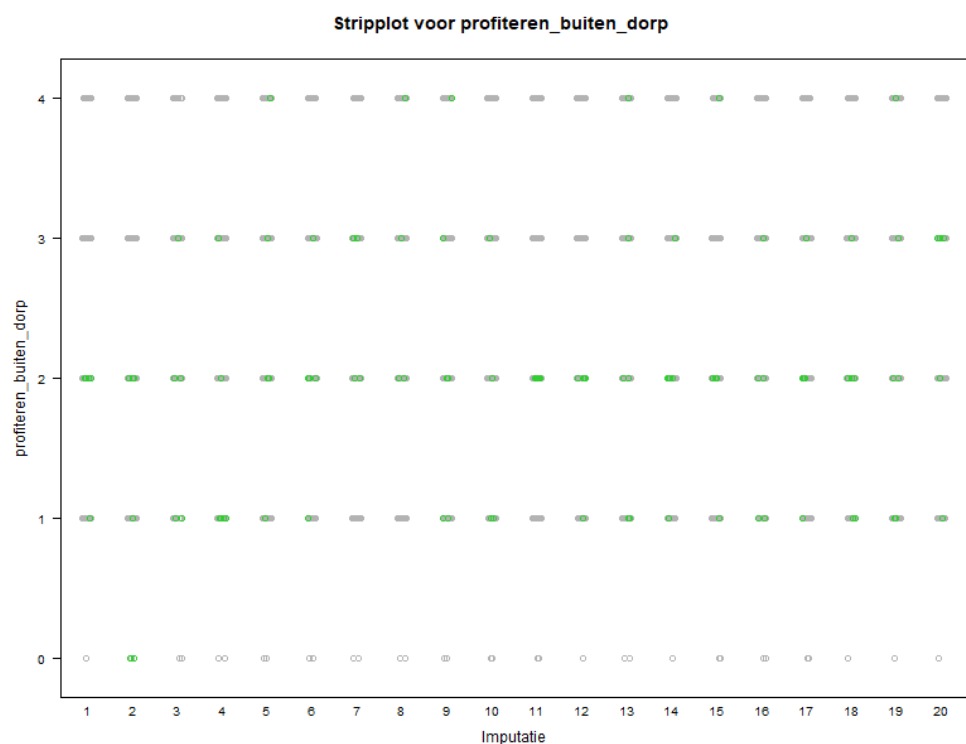
Figuur A27

Stripplot voor het vragenlijstitem 'vertrouwen_buiten_dorp' met geobserveerde (grijs) en geïmputeerde (groen) waarden per geïmputeerde dataset



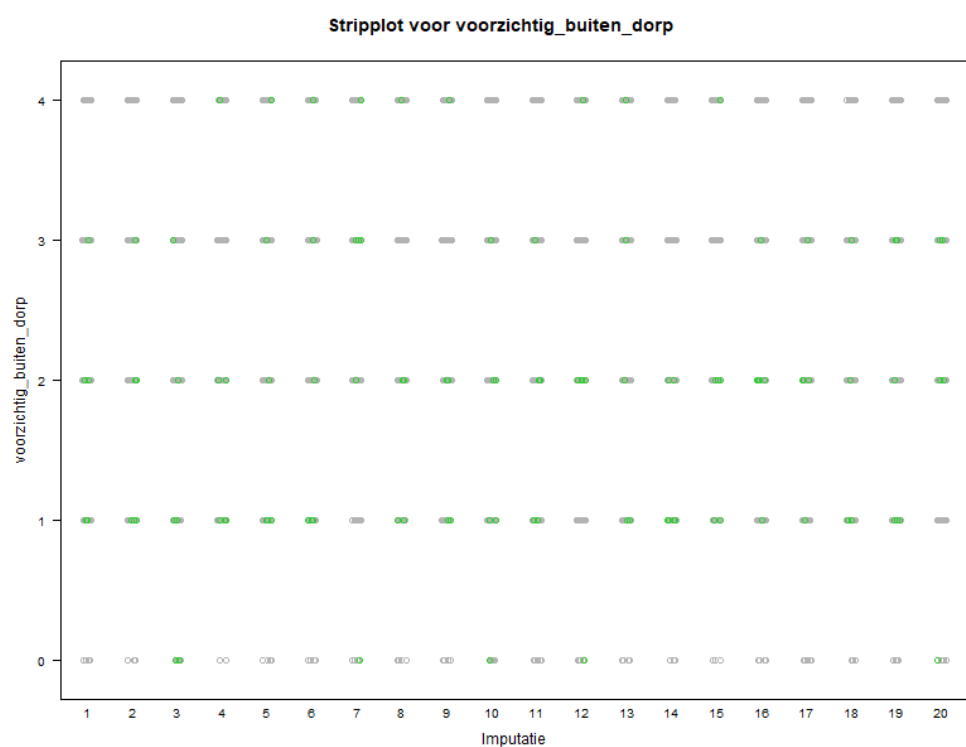
Figuur A28

Stripplot voor het vragenlijstitem 'eerlijkheid_buiten_dorp' met geobserveerde (grijs) en geïmputeerde (groen) waarden per geïmputeerde dataset



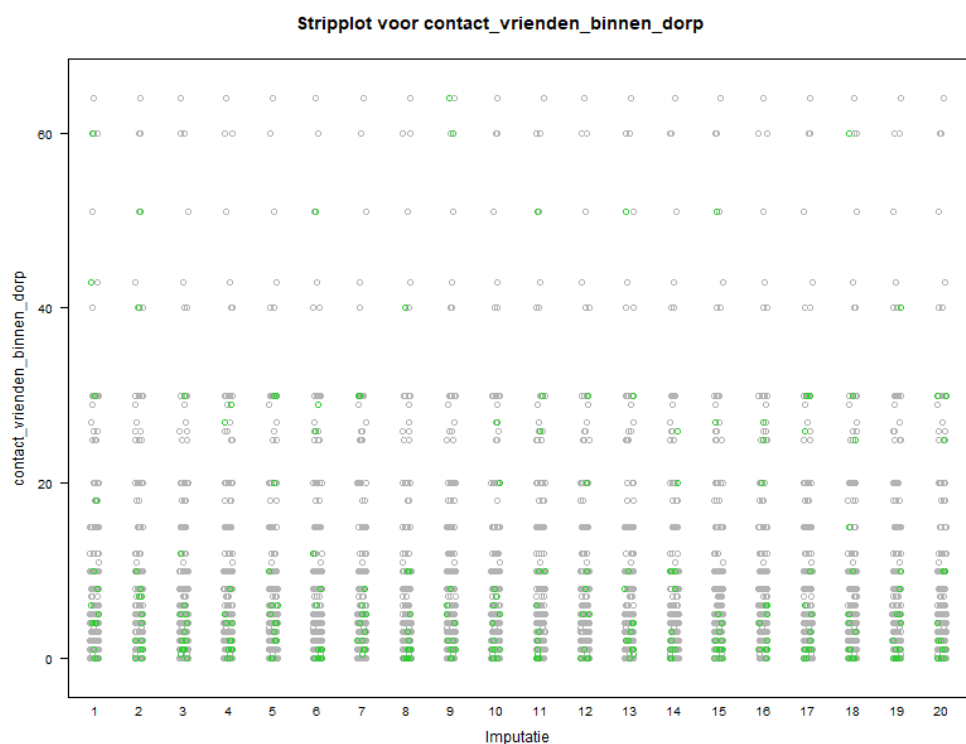
Figuur A29

Stripplot voor het vragenlijstitem 'profiteren_buiten_dorp' met geobserveerde (grijs) en geïmputeerde (groen) waarden per geïmputeerde dataset



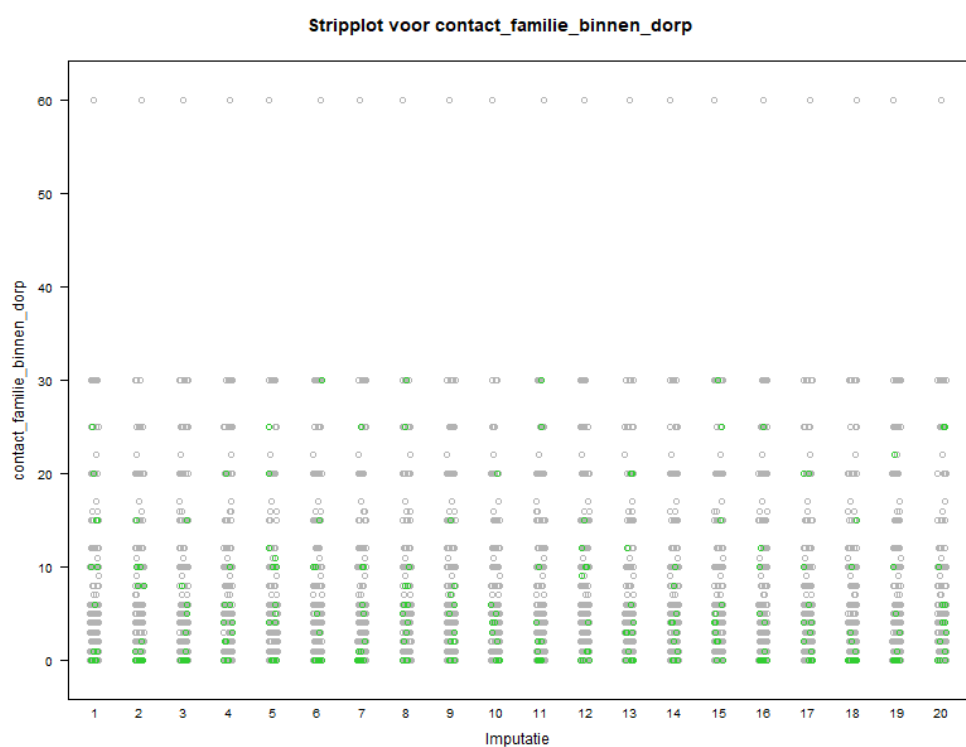
Figuur A30

Stripplot voor het vragenlijstitem 'voorzichtig_buiten_dorp' met geobserveerde (grijs) en geïmputeerde (groen) waarden per geïmputeerde dataset



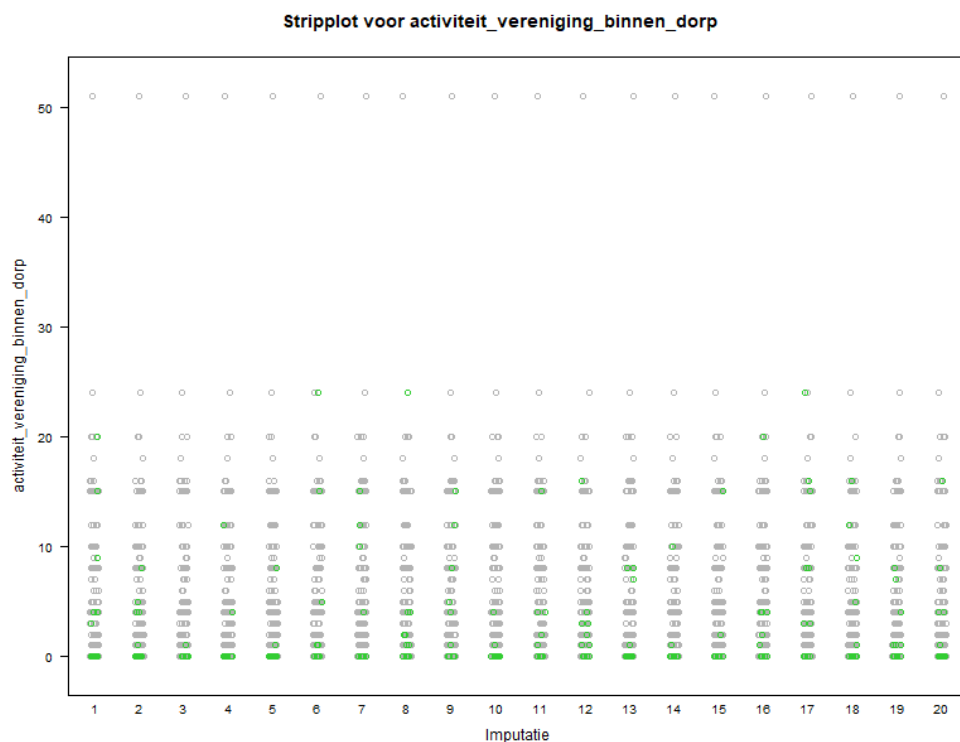
Figuur A31

Stripplot voor het vragenlijstitem 'contact_vrienden_binnen_dorp' met geobserveerde (grijs) en geïmputeerde (groen) waarden per geïmputeerde dataset



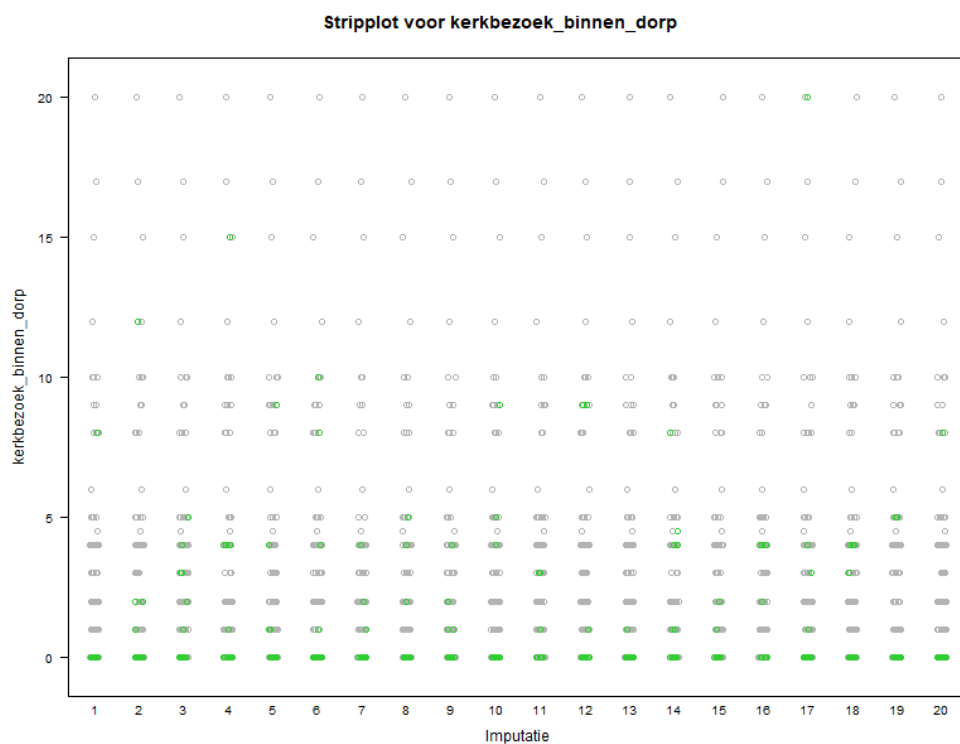
Figuur A32

Stripplot voor het vragenlijstitem 'contact_familie_binnen_dorp' met geobserveerde (grijs) en geïmputeerde (groen) waarden per geïmputeerde dataset



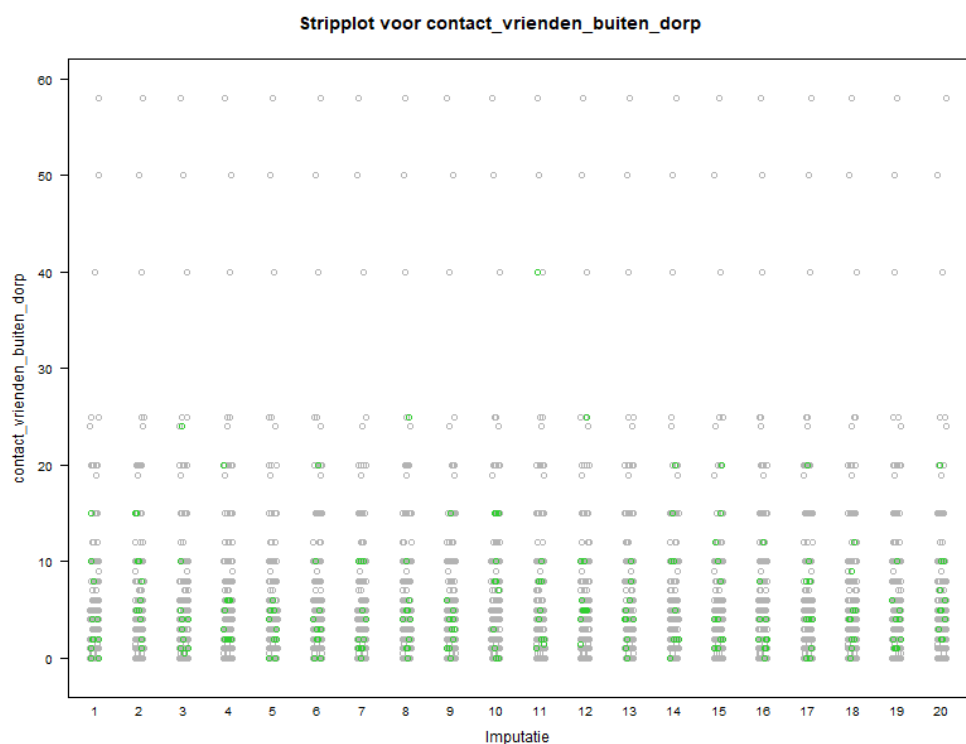
Figuur A33

Stripplot voor het vragenlijstitem 'activiteit_vereniging_binnen_dorp' met geobserveerde (grijs) en geïmputeerde (groen) waarden per geïmputeerde dataset



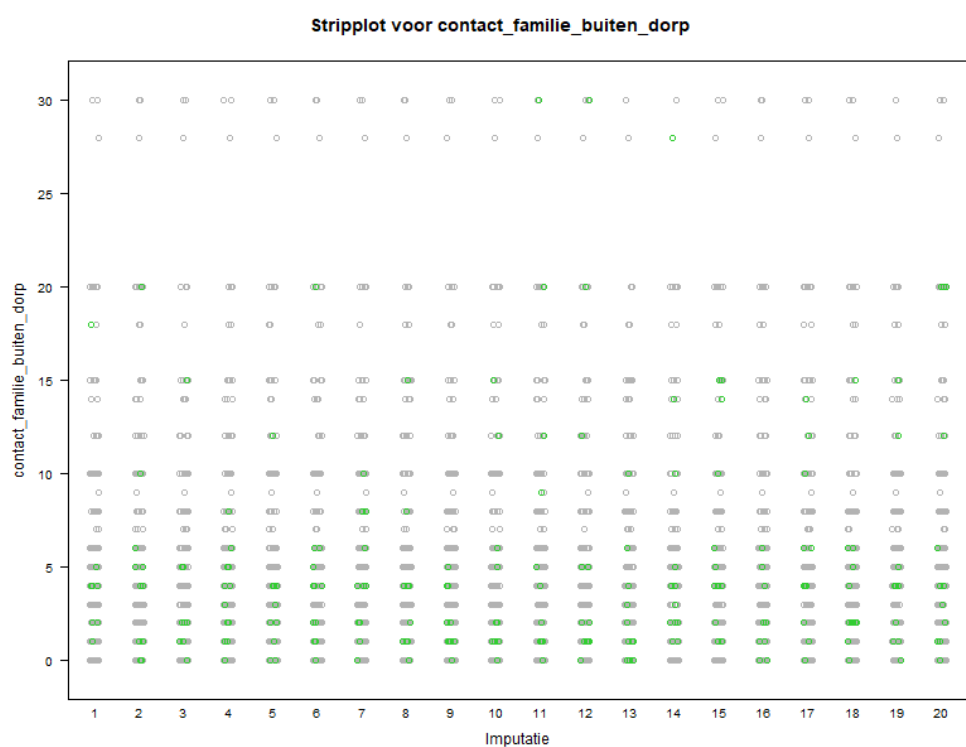
Figuur A34

Stripplot voor het vragenlijstitem 'kerkbezoek_binnen_dorp' met geobserveerde (grijs) en geïmputeerde (groen) waarden per geïmputeerde dataset



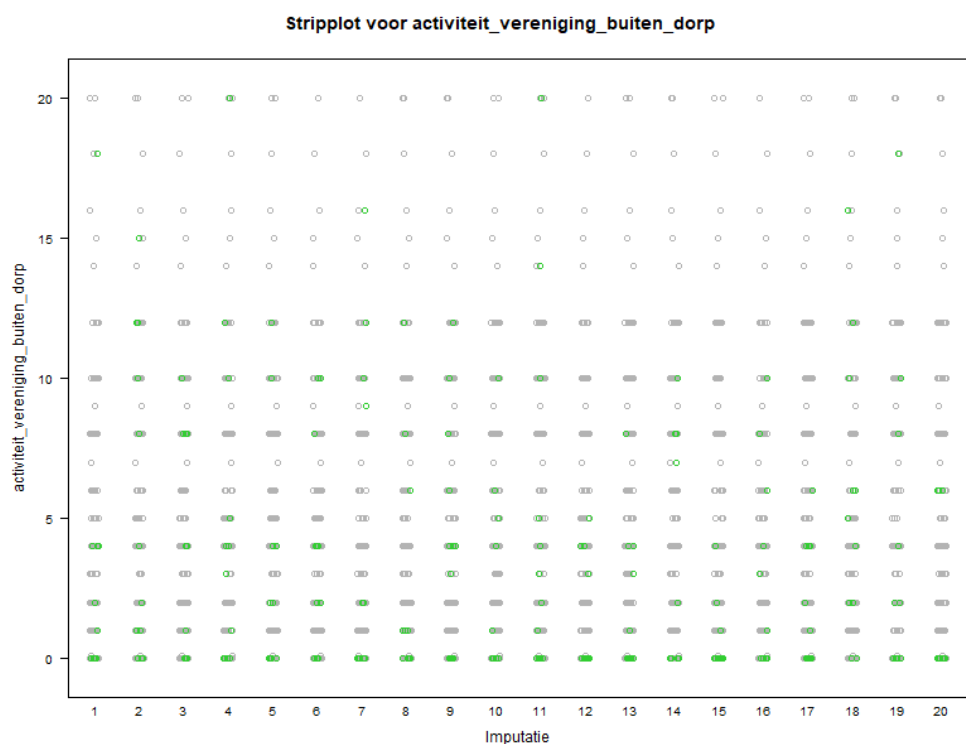
Figuur A35

Stripplot voor het vragenlijstitem 'contact_vrienden_buiten_dorp' met geobserveerde (grijs) en geïmputeerde (groen) waarden per geïmputeerde dataset



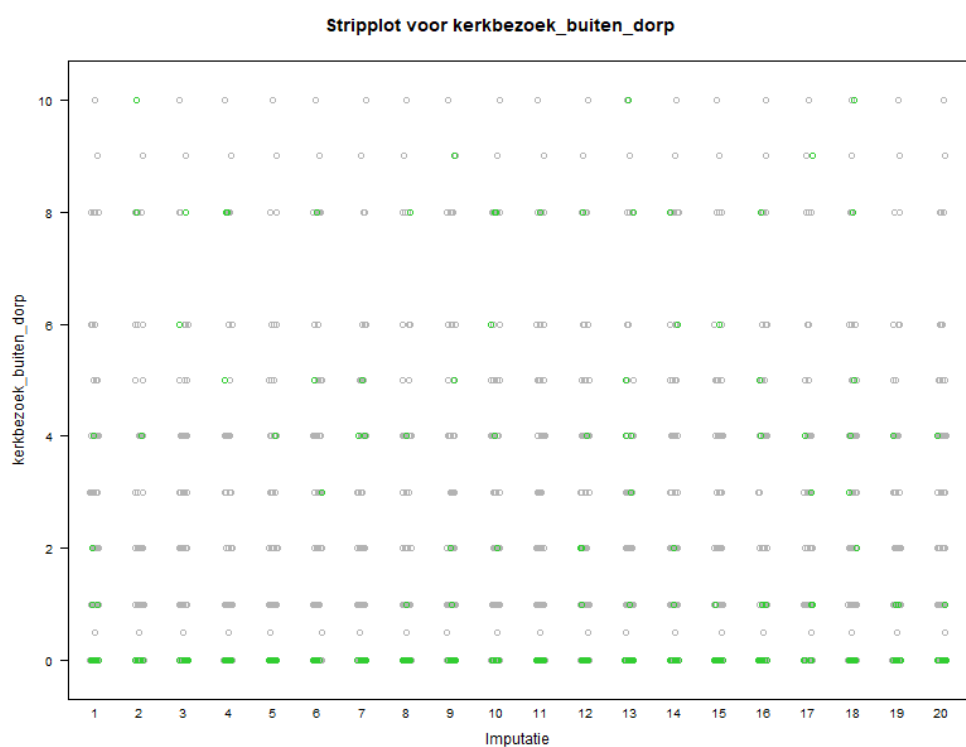
Figuur A36

Stripplot voor het vragenlijstitem 'contact_familie_buiten_dorp' met geobserveerde (grijs) en geïmputeerde (groen) waarden per geïmputeerde dataset



Figuur A37

Stripplot voor het vragenlijstitem 'activiteit_vereniging_buiten_dorp' met geobserveerde (grijs) en geïmputeerde (groen) waarden per geïmputeerde dataset



Figuur A38

Stripplot voor het vragenlijstitem 'kerkbezoek_buiten_dorp' met geobserveerde (grijs) en geïmputeerde (groen) waarden per geïmputeerde dataset

Bijlage B: invloed van ontbrekende waarden

In deze tabel zijn de variabelen aangeduid met hun R-namen zoals gebruikt in de analyses. Zie bijlage D voor een overzicht van deze R-namen en de bijbehorende variabelen.

Tabel B1

Fraction of missing information (y) waarden per variabele en per model

Model	Variabele	y
H1	(Intercept)	0,01
H1	leeftijd	0,01
H1	opleiding.L	0,01
H1	opleiding.Q	0,01
H1	woonplaats2	0,01
H1	woonplaats3	0,01
H1	woonsituatie_verandering2	0,02
H1	geslacht2	0,01
H1	geslacht4	0,01
H2a	(Intercept)	0,02
H2a	leeftijd	0,02
H2a	opleiding.L	0,01
H2a	opleiding.Q	0,01
H2a	woonplaats2	0,01
H2a	woonplaats3	0,02
H2a	woonsituatie_verandering2	0,02
H2a	geslacht2	0,02
H2a	geslacht4	0,01
H2b	(Intercept)	0,07
H2b	leeftijd	0,06
H2b	opleiding.L	0,04
H2b	opleiding.Q	0,07
H2b	woonplaats2	0,04
H2b	woonplaats3	0,03
H2b	woonsituatie_verandering2	0,03
H2b	geslacht2	0,05
H2b	geslacht4	0,33
H3	(Intercept)	0,02
H3	relatief_cognitief_kapitaal_binnen_dorp	0,02
H3	relatief_structureel_kapitaal_binnen_dorp	0,06
H3	opleiding.L	0,01
H3	opleiding.Q	0,01
H3	woonplaats2	0,01
H3	woonplaats3	0,01
H3	woonsituatie_verandering2	0,02
H3	geslacht2	0,01
H3	geslacht4	0,01
H4	(Intercept)	0,02
H4	leeftijd	0,01
H4	relatief_cognitief_kapitaal_binnen_dorp	0,02

H4	relatief_structureel_kapitaal_binnen_dorp	0,06
H4	opleiding.L	0,01
H4	opleiding.Q	0,01
H4	woonplaats2	0,01
H4	woonplaats3	0,01
H4	woonsituatie_verandering2	0,02
H4	geslacht2	0,01
H4	geslacht4	0,01
H5	(Intercept)	0,05
H5	leeftijd	0,05
H5	relatief_cognitief_kapitaal_binnen_dorp	0,06
H5	relatief_structureel_kapitaal_binnen_dorp	0,13
H5	relatief_cognitief_kapitaal_binnen_dorp_leeftijd	0,06
H5	relatief_structureel_kapitaal_binnen_dorp_leeftijd	0,14
H5	opleiding.L	0,01
H5	opleiding.Q	0,01
H5	woonplaats2	0,01
H5	woonplaats3	0,01
H5	woonsituatie_verandering2	0,03
H5	geslacht2	0,01
H5	geslacht4	0,01

Bijlage C: R-script

In deze bijlage is het script te vinden dat in R is toegepast.

```
#Inladen Excelbestand:
```

```
View(Bestand_voor_import_R)
```

```
#Interne consistentie via Cronbach's alpha: cognitief sociaal kapitaal binnen eigen dorp
```

```
library(psych)
```

```
alpha_binnen <- alpha(Bestand_voor_import_R[, c(
  "vertrouwen_binnen_dorp",
  "eerlijkheid_binnen_dorp",
  "profiteren_binnen_dorp",
  "voorzichtig_binnen_dorp"
)])
```

```
print(alpha_binnen)
```

```
#Interne consistentie via Cronbach's alpha: cognitief sociaal kapitaal buiten eigen dorp
```

```
library(psych)
```

```
alpha_buiten <- alpha(Bestand_voor_import_R[, c(
  "vertrouwen_buiten_dorp",
  "eerlijkheid_buiten_dorp",
  "profiteren_buiten_dorp",
  "voorzichtig_buiten_dorp"
)])
```

```
print(alpha_buiten)
```

```
#Univariate statistieken voor continue en ordinale variabelen: N, Minimum, Maximum,
Gemiddelde, Standaarddeviatie voor imputatie
```

```
data_orig <- Bestand_voor_import_R
```

```
vars_cont_ord <- c(
  "leeftijd",
  "opleiding",
  "vertrekintentie",
  "relatief_cognitief_kapitaal_binnen_dorp",
  "relatief_cognitief_kapitaal_binnen_dorp_leeftijd",
  "relatief_structureel_kapitaal_binnen_dorp",
  "relatief_structureel_kapitaal_binnen_dorp_leeftijd"
)
```

```
univariaat_voor_imputatie_cont_ord <- do.call(rbind, lapply(vars_cont_ord, function(v) {
  x <- as.numeric(as.character(data_orig[[v]]))
}))
```

```

x_clean <- x[!is.na(x)]
data.frame(
  Variabele = v,
  N = length(x_clean),
  Min = round(min(x_clean), 2),
  Max = round(max(x_clean), 2),
  Gem = round(mean(x_clean), 2),
  SD = round(sd(x_clean), 2),
  stringsAsFactors = FALSE
)
}))

print(univariaat_voor_imputatie_cont_ord)

```

```
#Frequentieverdeling categorische variabelen voor imputatie
```

```

invisible(lapply(c("woonplaats", "woonsituatie_verandering", "geslacht"), \(var) {
  cat("\nFrequentieverdeling voor:", var, "\n")
  print(table(Bestand_voor_import_R[[var]], useNA = "always"))
}))

```

```
#Bivariate statistieken voor imputatie (alleen eindvariabelen)
```

```

library(lsr)

cont_ordinaal <- c(
  "leeftijd", "opleiding", "vertrekintentie",
  "relatief_cognitief_kapitaal_binnen_dorp",
  "relatief_cognitief_kapitaal_binnen_dorp_leeftijd",
  "relatief_structureel_kapitaal_binnen_dorp",
  "relatief_structureel_kapitaal_binnen_dorp_leeftijd"
)

categorisch <- c("woonsituatie_verandering", "woonplaats", "geslacht")

df <- Bestand_voor_import_R # dataset

# p-waarde <0,001
pval_str <- function(p) {
  if (is.na(p)) "NA" else if (p < 0.001) "< 0.001" else sprintf("%.3f", p)
}

# 1. Spearman-correlaties
cat("### Spearman-correlaties (continu vs continu) ###\n\n")
for (i in 1:(length(cont_ordinaal) - 1)) {
  for (j in (i + 1):length(cont_ordinaal)) {
    x <- df[[cont_ordinaal[i]]]
    y <- df[[cont_ordinaal[j]]]
    complete <- complete.cases(x, y)
    x <- x[complete]

```



```

y <- y[complete]
n <- length(x)
test <- cor.test(x, y, method = "spearman")
cat(cont_ordinaal[i], "|", cont_ordinaal[j],
    "- rho =", round(test$estimate, 2),
    "p =", pval_str(test$p.value),
    "n =", n, "\n")
}
}

# 2. ANOVA  $\sqrt{R^2}$ 
cat("\n### ANOVA  $\sqrt{R^2}$  (continu vs categorisch) ###\n\n")
for (x in cont_ordinaal) {
  for (g in categorisch) {
    complete <- complete.cases(df[[x]], df[[g]])
    y <- df[[x]][complete]
    group <- as.factor(df[[g]][complete])
    n <- length(y)
    aov_result <- summary(aov(y ~ group))
    ss <- aov_result[[1]][["Sum Sq"]]
    r2 <- ss[1] / sum(ss)
    pval <- aov_result[[1]][["Pr(>F)"]][1]
    cat(x, "|", g,
        "-  $\sqrt{R^2}$  =", round(sqrt(r2), 2),
        "p =", pval_str(pval),
        "n =", n, "\n")
  }
}

# 3. Cramér's V
cat("\n### Cramér's V (categorisch vs categorisch) ###\n\n")
for (i in 1:(length(categorisch) - 1)) {
  for (j in (i + 1):length(categorisch)) {
    x <- df[[categorisch[i]]]
    y <- df[[categorisch[j]]]
    complete <- complete.cases(x, y)
    x <- as.factor(x[complete])
    y <- as.factor(y[complete])
    tbl <- table(x, y)
    n <- sum(tbl)
    chi <- chisq.test(tbl)
    cv <- cramersV(tbl)
    cat(categorisch[i], "|", categorisch[j],
        "- Cramér's V =", round(cv, 2),
        "p =", pval_str(chi$p.value),
        "n =", n, "\n")
  }
}

```

```

#Imputatiescript
# Stap 0: Laad benodigde packages

library(mice)
library(dplyr)
library(openxlsx)

# Stap 1: Laad dataset

data <- Bestand_voor_import_R # reeds ingelezen in RStudio

# Stap 2: Definieer variabelen

cognitief_vars <- c(
  "vertrouwen_binnen_dorp", "eerlijkheid_binnen_dorp", "profiteren_binnen_dorp",
  "voorzichtig_binnen_dorp",
  "vertrouwen_buiten_dorp", "eerlijkheid_buiten_dorp", "profiteren_buiten_dorp",
  "voorzichtig_buiten_dorp"
)

structureel_vars <- c(
  "contact_vrienden_binnen_dorp", "contact_familie_binnen_dorp",
  "activiteit_vereniging_binnen_dorp", "kerkbezoek_binnen_dorp",
  "contact_vrienden_buiten_dorp", "contact_familie_buiten_dorp",
  "activiteit_vereniging_buiten_dorp", "kerkbezoek_buiten_dorp"
)

categorical_vars_impute <- c("woonsituatie_huidig", "woonsituatie_toekomst")
opleidings_var <- "opleiding"
volledig_beantwoorde_vars <- c("leeftijd", "vertrekintentie", "woonplaats", "geslacht")

# Stap 3: Zet kolommen om naar juiste type

for (v in cognitief_vars) {
  if (v %in% names(data)) {
    data[[v]] <- as.numeric(as.character(data[[v]]))
  }
}

for (v in structureel_vars) {
  if (v %in% names(data)) {
    data[[v]] <- as.numeric(as.character(data[[v]]))
  }
}

for (v in categorical_vars_impute) {
  if (v %in% names(data)) {

```

```

    data[[v]] <- as.factor(data[[v]])
  }
}

if (opleidings_var %in% names(data)) {
  data[[opleidings_var]] <- as.numeric(as.character(data[[opleidings_var]]))
  data[[opleidings_var]] <- factor(data[[opleidings_var]], levels = c(0, 1, 2), ordered = TRUE)
}

# Stap 4: Imputatiemethoden instellen

method <- make.method(data)
method[cognitief_vars] <- "pmm"
method[structureel_vars] <- "pmm"
method[categorical_vars_impute] <- "polyreg"
method[opleidings_var] <- "polr"
method[volledig_beantwoorde_vars] <- ""

# Stap 5: Voer imputatie uit

set.seed(23032024)
pred <- make.predictorMatrix(data)

imp <- mice(data, method = method, predictorMatrix = pred, m = 20, maxit = 20, printFlag =
TRUE)

# Stap 6: Voeg samengestelde variabelen toe + woonsituatie_verandering

cog_binnen <- cognitief_vars[1:4]
cog_buiten <- cognitief_vars[5:8]
struct_binnen <- structureel_vars[1:4]
struct_buiten <- structureel_vars[5:8]

imputed_list <- complete(imp, "all")

imputed_list <- lapply(imputed_list, function(df) {
  # Samengestelde variabelen
  df$relatief_cognitief_kapitaal_binnen_dorp <-
    rowSums(df[cog_binnen], na.rm = TRUE) /
    rowSums(df[c(cog_binnen, cog_buiten)], na.rm = TRUE)

  df$relatief_structureel_kapitaal_binnen_dorp <-
    rowSums(df[struct_binnen], na.rm = TRUE) /
    rowSums(df[c(struct_binnen, struct_buiten)], na.rm = TRUE)

  df$relatief_cognitief_kapitaal_binnen_dorp_leeftijd <-
    df$relatief_cognitief_kapitaal_binnen_dorp * df$leeftijd

```

```

df$relatief_structueel_kapitaal_binnen_dorp_leeftijd <-
  df$relatief_structueel_kapitaal_binnen_dorp * df$leeftijd

# Woonsituatie_verandering: 1 = gelijk, 2 = verandering
df$woonsituatie_verandering <- ifelse(
  df$woonsituatie_huidig == df$woonsituatie_toekomst,
  1,
  2
)

return(df)
})

```

#Exporteer datasets naar Excel

```

# Stel de werkdirectory in
setwd("C:/Users/lenovo/Documents/Master Soc PMB/Stage en scriptie/Scriptie/Uitwerken
scriptie/R project resultaten")

# Sla elk element van de lijst op als een afzonderlijk Excel-bestand
for (i in 1:20) {
  write.xlsx(imputed_list[[i]], file = paste0("imputed_data", i, ".xlsx"), overwrite = TRUE)
}

```

#Convergentie script

```

library(mice)
library(lattice)

# Maak map aan indien nodig
if (!dir.exists("Convergentietoetsing")) {
  dir.create("Convergentietoetsing")
}

# Selecteer alleen geïmputeerde variabelen
plot_vars <- intersect(c(
  "vertrouwen_binnen_dorp", "eerlijkheid_binnen_dorp", "profiteren_binnen_dorp",
  "voorzichtig_binnen_dorp",
  "vertrouwen_buiten_dorp", "eerlijkheid_buiten_dorp", "profiteren_buiten_dorp",
  "voorzichtig_buiten_dorp",
  "contact_vrienden_binnen_dorp", "contact_familie_binnen_dorp",
  "activiteit_vereniging_binnen_dorp", "kerkbezoek_binnen_dorp",
  "contact_vrienden_buiten_dorp", "contact_familie_buiten_dorp",
  "activiteit_vereniging_buiten_dorp", "kerkbezoek_buiten_dorp",
  "woonsituatie_huidig", "woonsituatie_toekomst", "opleiding"
), names(imp$imp))

# Loop door alle variabelen en maak een plot
for (v in plot_vars) {
  cat("Plot:", v, "\n")
}

```

```

png(file.path("Convergentietoetsing", paste0("Convergentietoetsing_", v, ".png")), width =
800, height = 600)

# Genereer standaard plot en pas layout en x-as aan
p <- update(
  plot(imp, v, main = v),
  layout = c(1, 2),
  xlab = "Iteratie",
  xlim = c(1, 20),
  scales = list(
    x = list(draw = TRUE, alternating = 1, at = 1:20, limits = c(1, 20)),
    y = list(draw = TRUE)
  ),
  par.settings = list(
    strip.background = list(col = "white"),
    layout.heights = list(strip = 2)
  )
)

print(p)
dev.off()
}

```

#Plausibiliteit script

```

library(mice)
library(lattice)

# Maak map aan indien nodig
if (!dir.exists("Plausibiliteitstoetsing")) {
  dir.create("Plausibiliteitstoetsing")
}

# Selecteer alleen geïmputeerde variabelen
plot_vars <- intersect(c(
  "vertrouwen_binnen_dorp", "eerlijkheid_binnen_dorp", "profiteren_binnen_dorp",
  "voorzichtig_binnen_dorp",
  "vertrouwen_buiten_dorp", "eerlijkheid_buiten_dorp", "profiteren_buiten_dorp",
  "voorzichtig_buiten_dorp",
  "contact_vrienden_binnen_dorp", "contact_familie_binnen_dorp",
  "activiteit_vereniging_binnen_dorp", "kerkbezoek_binnen_dorp",
  "contact_vrienden_buiten_dorp", "contact_familie_buiten_dorp",
  "activiteit_vereniging_buiten_dorp", "kerkbezoek_buiten_dorp",
  "woonsituatie_huidig", "woonsituatie_toekomst", "opleiding"
), names(imp$imp))

# Maak een stripplot
for (v in plot_vars) {
  cat("Plot:", v, "\n")
}

```

```

png(file.path("Plausibiliteitstoetsing", paste0("Plausibiliteitstoetsing_", v, ".png")), width =
800, height = 600)

# Stripplot maken met aangepaste x-as titel en contrasterende kleuren
print(stripplot(
  imp, formula(paste0(v, "~ .imp")),
  subset = .imp > 0,
  main = paste("Stripplot voor", v),
  xlab = "Imputatie",
  col = c("grey70", "limegreen")
))

dev.off()
}

```

#Univariate statistieken voor continue en ordinale variabelen: N, Minimum, Maximum, Gemiddelde, Standaarddeviatie na imputatie

```

library(mice)

vars <- c(
  "leeftijd",
  "opleiding",
  "vertrekintentie",
  "relatief_cognitief_kapitaal_binnen_dorp",
  "relatief_cognitief_kapitaal_binnen_dorp_leeftijd",
  "relatief_structureel_kapitaal_binnen_dorp",
  "relatief_structureel_kapitaal_binnen_dorp_leeftijd"
)

# Aantal imputaties
m <- length(imputed_list)

# Resultaten initialiseren
results <- data.frame(
  Variable = character(),
  N = numeric(),
  Min = numeric(),
  Max = numeric(),
  Mean = numeric(),
  SD = numeric(),
  stringsAsFactors = FALSE
)

for (v in vars) {
  # Bereken univariate statistieken per imputatie
  stats <- lapply(imputed_list, function(df) {
    x <- as.numeric(as.character(df[[v]]))
    x_clean <- x[!is.na(x)]
  })
}

```

```

c(
  N = length(x_clean),
  Min = min(x_clean),
  Max = max(x_clean),
  Mean = mean(x_clean),
  Var = var(x_clean)
)
})
stats_df <- as.data.frame(do.call(rbind, stats))

# Bereken standaarddeviatie: wortel van gemiddelde variantie
SD <- sqrt(mean(stats_df$Var, na.rm = TRUE))

# Voeg resultaten toe
results <- rbind(results, data.frame(
  Variable = v,
  N = round(mean(stats_df$N), 0),
  Min = round(mean(stats_df$Min), 2),
  Max = round(mean(stats_df$Max), 2),
  Mean = round(mean(stats_df$Mean), 2),
  SD = round(SD, 2)
))
}

print(results)

```

#Frequentieverdeling categorische variabelen na imputatie

```

freqs <- lapply(imputed_list, function(df) {
  table(df$woonsituatie_verandering)
})

# Zet tabellen om naar numerieke rijen
freq_matrix <- do.call(rbind, lapply(freqs, as.numeric))

# Bereken gemiddelde frequentie per categorie
mean_freq <- colMeans(freq_matrix)

# Zet in overzichtelijke tabel
categories <- names(freqs[[1]])
result <- data.frame(
  Categorie = categories,
  GemiddeldeFrequentie = round(mean_freq, 2)
)

print(result)

```

```
#Bivariate statistieken na imputatie (alleen eindvariabelen)
```

```
library(lsr)
```

```
# Variabelen
```

```
continu_en_ordinale_variabelen <- c(  
  "leeftijd",  
  "opleiding",  
  "vertrekintentie",  
  "relatief_cognitief_kapitaal_binnen_dorp",  
  "relatief_cognitief_kapitaal_binnen_dorp_leeftijd",  
  "relatief_structureel_kapitaal_binnen_dorp",  
  "relatief_structureel_kapitaal_binnen_dorp_leeftijd"  
)
```

```
categorisch <- c("woonsituatie_verandering", "woonplaats", "geslacht")
```

```
# Fisher's z-transformatie functies
```

```
fisher_z <- function(r) {  
  0.5 * log((1 + r) / (1 - r))  
}  
inv_fisher_z <- function(z) {  
  (exp(2 * z) - 1) / (exp(2 * z) + 1)  
}
```

```
cat("### Gemiddelde Spearman-correlaties (continu vs continu) ###\n\n")
```

```
for (i in 1:(length(continu_en_ordinale_variabelen) - 1)) {  
  for (j in (i + 1):length(continu_en_ordinale_variabelen)) {  
    var1 <- continu_en_ordinale_variabelen[i]  
    var2 <- continu_en_ordinale_variabelen[j]  
  
    z_scores <- c()  
    ns <- c()  
  
    for (d in seq_along(imputed_list)) {  
      data <- imputed_list[[d]]  
      complete_cases <- complete.cases(data[[var1]], data[[var2]])  
      x <- as.numeric(data[[var1]][complete_cases])  
      y <- as.numeric(data[[var2]][complete_cases])  
      rho <- cor(x, y, method = "spearman")  
      if (!is.na(rho) && abs(rho) < 1) {  
        z <- fisher_z(rho)  
        z_scores <- c(z_scores, z)  
        ns <- c(ns, length(x))  
      }  
    }  
  }  
}
```

```
if (length(z_scores) > 0) {
```



```

mean_z <- mean(z_scores)
mean_rho <- inv_fisher_z(mean_z)
mean_n <- round(mean(ns))
cat(var1, "|", var2, "- Gemiddelde Spearman rho (via Fisher's z):", round(mean_rho, 2),
    " (gemiddelde n =", mean_n, ")\n\n")
}
}
}

cat("### Gemiddelde correlatie continu vs categorisch (ANOVA F-toets, via R2) ###\n\n")

for (cont in continu_en_ordinale_variabelen) {
  for (cat_var in categorisch) {
    r_squares <- c()
    ns <- c()

    for (d in seq_along(imputed_list)) {
      data <- imputed_list[[d]]
      complete_cases <- complete.cases(data[[cont]], data[[cat_var]])
      x <- as.numeric(data[[cont]][complete_cases])
      g <- as.factor(data[[cat_var]][complete_cases])
      aov_res <- aov(x ~ g)
      aov_sum <- summary(aov_res)[[1]]
      ss_between <- aov_sum["g", "Sum Sq"]
      ss_total <- sum(aov_sum[["Sum Sq"]])
      r2 <- ss_between / ss_total
      r_squares <- c(r_squares, r2)
      ns <- c(ns, length(x))
    }

    if (length(r_squares) > 0) {
      mean_r2 <- mean(r_squares)
      mean_corr <- sqrt(mean_r2)
      mean_n <- round(mean(ns))
      cat(cont, "|", cat_var, "- Gemiddelde correlatie (via wortel R2):", round(mean_corr, 2),
          " (gemiddelde n =", mean_n, ")\n\n")
    }
  }
}

cat("### Gemiddelde Cramér's V (categorisch vs categorisch) ###\n\n")

for (i in 1:(length(categorisch) - 1)) {
  for (j in (i + 1):length(categorisch)) {
    var1 <- categorisch[i]
    var2 <- categorisch[j]
    cramers_vs <- c()
    ns <- c()
  }
}

```

```

for (d in seq_along(imputed_list)) {
  data <- imputed_list[[d]]
  complete_cases <- complete.cases(data[[var1]], data[[var2]])
  tbl <- table(as.factor(data[[var1]][complete_cases]),
as.factor(data[[var2]][complete_cases]))
  cv <- tryCatch({
    cramersV(tbl)
  }, error = function(e) NA)
  if (!is.na(cv)) {
    cramers_vs <- c(cramers_vs, cv)
    ns <- c(ns, sum(tbl))
  }
}

if (length(cramers_vs) > 0) {
  mean_cv <- mean(cramers_vs)
  mean_n <- round(mean(ns))
  cat(var1, "|", var2, "- Gemiddelde Cramér's V:", round(mean_cv, 2),
    " (gemiddelde n =", mean_n, ")\n\n")
}
}
}

```

#Hypothesetoetsing:

```

library(mice)
library(dplyr)
library(broom)

# Stap 1: Zet categorische variabelen om als factor (dummy)
imp$data$opleiding <- factor(imp$data$opleiding, ordered = TRUE)
imp$data$woonplaats <- factor(imp$data$woonplaats)
imp$data$woonsituatie_verandering <- factor(imp$data$woonsituatie_verandering)
imp$data$geslacht <- factor(imp$data$geslacht)

# Stap 2: Definieer modellen per hypothese

# H1
model1 <- with(imp, lm(vertrekintentie ~ leeftijd + opleiding + woonplaats +
woonsituatie_verandering + geslacht))
pooled1 <- pool(model1)

# H2a
model2a <- with(imp, lm(relatief_cognitief_kapitaal_binnen_dorp ~ leeftijd + opleiding +
woonplaats + woonsituatie_verandering + geslacht))
pooled2a <- pool(model2a)

# H2b

```

```

model2b <- with(imp, lm(relatief_structureel_kapitaal_binnen_dorp ~ leeftijd + opleiding +
  woonplaats + woonsituatie_verandering + geslacht))
pooled2b <- pool(model2b)

# H3
model3 <- with(imp, lm(vertrekintentie ~ relatief_cognitief_kapitaal_binnen_dorp +
  relatief_structureel_kapitaal_binnen_dorp +
  opleiding + woonplaats + woonsituatie_verandering + geslacht))
pooled3 <- pool(model3)

# H4
model4 <- with(imp, lm(vertrekintentie ~ leeftijd + relatief_cognitief_kapitaal_binnen_dorp +
  relatief_structureel_kapitaal_binnen_dorp +
  opleiding + woonplaats + woonsituatie_verandering + geslacht))
pooled4 <- pool(model4)

# H5
model5 <- with(imp, lm(vertrekintentie ~ leeftijd +
  relatief_cognitief_kapitaal_binnen_dorp +
  relatief_structureel_kapitaal_binnen_dorp +
  relatief_cognitief_kapitaal_binnen_dorp_leeftijd +
  relatief_structureel_kapitaal_binnen_dorp_leeftijd +
  opleiding + woonplaats + woonsituatie_verandering + geslacht))
pooled5 <- pool(model5)

# Stap 3: printfunctie met 5 decimalen + n

print_model <- function(pooled_model) {
  model_summary <- tidy(pooled_model) %>%
    mutate(across(c(estimate, std.error, statistic), ~ round(., 5)),
      p.value = formatC(p.value, format = "f", digits = 5)) %>%
    select(term, estimate, std.error, statistic, p.value)

  # Haal het gemiddeld aantal observaties op via glance()
  model_n <- glance(pooled_model)$nobs
  cat("Aantal observaties (gemiddeld n over imputaties):", round(model_n), "\n\n")

  return(model_summary)
}

# Stap 4: print resultaten

cat("\n--- H1 ---\n"); print(print_model(pooled1))
cat("\n--- H2a ---\n"); print(print_model(pooled2a))
cat("\n--- H2b ---\n"); print(print_model(pooled2b))
cat("\n--- H3 ---\n"); print(print_model(pooled3))
cat("\n--- H4 ---\n"); print(print_model(pooled4))
cat("\n--- H5 ---\n"); print(print_model(pooled5))

```

```
#FMI
```

```
library(mice)

get_fmi_values <- function(pooled_model, hypothese) {
  model_summary <- summary(pooled_model, type = "all")

  # Zorg dat fmi-waarden numeriek zijn
  fmi_vals <- as.numeric(model_summary$fmi)

  # Resultatentabel
  gamma_df <- data.frame(
    Variabele = model_summary$term,
    Gamma = round(fmi_vals, 2)
  )

  cat("\n--- γ (FMI) ---", hypothese, "---\n")
  print(gamma_df)
}

# γ-waarden printen voor alle modellen
get_fmi_values(pooled1, "H1")
get_fmi_values(pooled2a, "H2a")
get_fmi_values(pooled2b, "H2b")
get_fmi_values(pooled3, "H3")
get_fmi_values(pooled4, "H4")
get_fmi_values(pooled5, "H5")
```

```
#r2 en r2a
```

```
#Functie om R2 en Adjusted R2 te printen
print_r_squared <- function(pooled_model, modelnaam) {
  rsq <- pool.r.squared(pooled_model)

  cat(paste0("\n--- ", modelnaam, " ---\n"))
  cat("R2: ", round(rsq[1], 5), "\n", sep = "")
  cat("Adjusted R2: ", round(rsq[2], 5), "\n", sep = "")
}

# Resultaten printen per model
print_r_squared(pooled1, "H1")
print_r_squared(pooled2a, "H2a")
print_r_squared(pooled2b, "H2b")
print_r_squared(pooled3, "H3")
print_r_squared(pooled4, "H4")
print_r_squared(pooled5, "H5")
```

Bijlage D: R-namen bij vragenlijstitems en variabelen

In deze bijlage staan de namen die in R zijn gebruikt voor vragenlijstitems en variabelen. In de bijlage zijn ten eerste individuele vragenlijstitems waarvan ontbrekende waarden zijn geïmputeerd opgenomen. Ook zijn alle variabelen getoond die in de uiteindelijke modellen zijn opgenomen.

Tabel D1

Vragenlijstitems waarvan ontbrekende waarden zijn geïmputeerd met bijbehorende R-namen

Vragenlijstitem	R-naam
Wat is uw woonsituatie?	woonsituatie_huidig
Welke woonsituatie verwacht u dat over twee jaar voor u geldt?	woonsituatie_toekomst
De meeste van mijn dorpsgenoten zijn te vertrouwen	vertrouwen_binnen_dorp
De meeste van mijn dorpsgenoten proberen meestal eerlijk te zijn	eerlijkheid_binnen_dorp
De meeste van mijn dorpsgenoten zouden van je profiteren als ze de kans kregen	profiteren_binnen_dorp
Je kunt niet te voorzichtig zijn bij het omgaan met mijn dorpsgenoten	voorzichtig_binnen_dorp
De meeste mensen die geen dorpsgenoten van mij zijn, zijn te vertrouwen	vertrouwen_buiten_dorp
De meeste mensen die geen dorpsgenoten van mij zijn, proberen meestal eerlijk te zijn	eerlijkheid_buiten_dorp
De meeste mensen die geen dorpsgenoten van mij zijn, zouden van je profiteren als ze de kans kregen	profiteren_buiten_dorp
Je kunt niet te voorzichtig zijn bij het omgaan met mensen die geen dorpsgenoten van mij zijn	voorzichtig_buiten_dorp
Hoe vaak spreekt u in een gemiddelde maand van 30 dagen af met vrienden of kennissen die in hetzelfde dorp als uzelf wonen? (<i>geen huisgenoten</i>)	contact_vrienden_binnen_dorp
Hoe vaak spreekt u in een gemiddelde maand van 30 dagen af met familieleden die in hetzelfde dorp als uzelf wonen? (<i>geen huisgenoten</i>)	contact_familie_binnen_dorp
Hoe vaak bent u in een gemiddelde maand van 30 dagen actief bij verenigingen in het dorp waar u woont (bijvoorbeeld sportverenigingen, hobbyclubs etc.)?	activiteit_vereniging_binnen_dorp
Hoe vaak bezoekt u in een gemiddelde maand van 30 dagen een kerk en/of andere religieuze activiteiten in het dorp waar u woont?	kerkbezoek_binnen_dorp
Hoe vaak spreekt u in een gemiddelde maand van 30 dagen af met vrienden of kennissen die niet in hetzelfde dorp als uzelf wonen? (<i>geen huisgenoten</i>)	contact_vrienden_buiten_dorp
Hoe vaak spreekt u in een gemiddelde maand van 30 dagen af met familieleden die niet in hetzelfde dorp als uzelf wonen? (<i>geen huisgenoten</i>)	contact_familie_buiten_dorp
Hoe vaak bent u in een gemiddelde maand van 30 dagen actief bij verenigingen buiten het dorp waar u woont (bijvoorbeeld sportverenigingen, hobbyclubs etc.)?	activiteit_vereniging_buiten_dorp
Hoe vaak bezoekt u in een gemiddelde maand van 30 dagen een kerk en/of andere religieuze activiteiten buiten het dorp waar u woont?	kerkbezoek_buiten_dorp

Tabel D2

Variabelen die zijn opgenomen in uiteindelijke modellen met bijbehorende R-namen

Variabele	R-naam
Vertrekintentie	vertrekintentie
Leeftijd	leeftijd
Relatief cognitief sociaal kapitaal binnen dorp	relatief_cognitief_kapitaal_binnen_dorp
Relatief structureel sociaal kapitaal binnen dorp	relatief_structureel_kapitaal_binnen_dorp
Relatief cognitief sociaal kapitaal binnen dorp * leeftijd	relatief_cognitief_kapitaal_binnen_dorp_leeftijd
Relatief structureel sociaal kapitaal binnen dorp * leeftijd	relatief_structureel_kapitaal_binnen_dorp_leeftijd
(Verwacht) opleidingsniveau	opleiding
Woonplaats	woonplaats
Verwachte verandering woonsituatie	woonsituatie_verandering
Geslacht	geslacht