



Vertrouwensexperiment Wageningen: werkt het en waarom wel of niet?

Voorlopig eindverslag

*Ruud Muffels (projectleider), Kirsten Blom-Stam (onderzoeker)
& Stefan van Wanrooij (student assistent)*

Wageningen, April 2020

© Tilburg University/Reflect

Inhoudsopgave

1. Inleiding	3
1.1 Aanleiding voor het onderzoek	3
1.2 Theoretisch kader	7
1.3 Beleid -en onderzoeksvragen: hypothesen	14
2. Opzet onderzoek	16
2.1 Experimentgroepen	16
2.2 Procedure van random selectie, toewijzing en werving	19
2.3 Onderzoekspopulatie: deelnemers en afmeldingen	20
2.4 Product- en procesevaluatie	24
2.5 Dataverzameling product- en procesevaluatie	25
3 Resultaten productevaluatie	28
3.1 Productevaluatie en uitkomstmaten	28
3.2 Resultaten voltijdse uitstroom naar werk	30
3.3 De invloed van duur nader onderzocht: matching modellen	38
3.4 Conclusies uitstroom naar werk	42
3.5 Secundaire uitkomstmaten	44
3.5.1 Gezondheid en welbevinden	45
3.5.2 Keuzemogelijkheden (capabilities)	49
3.5.3 Sociale participatie	51
3.5.4 Zelfredzaamheid	55
3.5.5. Sociaal en institutioneel vertrouwen	57
3.5.6. Financiële situatie en armoede	59
3.5.7. Samenvatting treatment effecten	63
4. Resultaten Procesevaluatie	
4.1 Profiel van de deelnemer	66
4.2 Analyse vragenlijsten	66
4.2.1 Gehanteerde uitkomstmaten	66
4.2.2 Analyse methode	67
4.2.3 Resultaten analyse	68
4.2.4 Conclusies	84
4.3 Focusgroep gesprekken met klantmanagers	86
4.3.1 Inleiding	88
4.3.2 Het proces van experimenteren	88

Inhoudsopgave (vervolg)

5. Bevindingen en conclusies	95
4.1 Opzet van het experiment en werving	95
4.2 Uitvoering van de treatments	95
4.3 Uitstroom naar werk en matching	97
4.4 Niet-werk gerelateerde uitkomstmaten	98
4.5 Tot besluit	101
Literatuur	107
Bijlagen	110

1. Inleiding

1.1 Aanleiding voor het onderzoek

Per 1 oktober 2017 is de gemeente Wageningen gestart met een experiment in het kader van artikel 83 van de Participatiewet met als titel en adagium “Vertrouwen werkt!”. In het onderzoeksvoorstel (Muffels & van der Klink, 2017) maar ook in de aanvraag van 25 april 2017 om Wageningen aan te wijzen als een van de experimenteer gemeenten in het kader van het Tijdelijk Besluit Experimenten Participatiewet werd gesproken van het Vertrouwensexperiment Wageningen. Wageningen behoorde tot de vier initiatiefgemeenten die samen met Utrecht, Tilburg en Groningen in 2015 een andere aanpak voorstelden vanwege twijfels over de effectiviteit van de Participatiewet die per 1 januari 2015 van kracht was geworden. Er bestonden vooral twijfels over de vraag of de strenge regels en verplichtingen (re-integratie –en sollicitatieplicht plus tegenprestatie in combinatie met kortingen en sancties) wel het beoogde effect hebben. De maatschappelijke participatie van de kwetsbaarste groepen in de Participatiewet leek eerder te worden belemmerd door de strenge regels en verplichtingen dan bevorderd. De gemeenten wilde ruimte voor een andere aanpak waarin het mogelijk zou zijn om de vrijlating van verdiensten te verruimen, er minder (repressief) toezicht en controle zou zijn en meer maatwerk en vertrouwen in de begeleiding. De gemeente wilde daarom onderzoeken of:

- a. een andere benadering van klanten (met een grotere vrijheid zelf keuzen te maken, met intensieve begeleiding meer maatwerk) leidt tot hogere uitstroom naar werk, hogere maatschappelijke participatie en verhoging van welbevinden;
- b. een verruiming van de mogelijkheden tot bijverdienen voor mensen in de bijstand leidt tot hogere uitstroom naar deeltijd of voltijd werk, vermindering van bijstandsafhankelijkheid, hogere maatschappelijke participatie en verhogen van welbevinden.

In het onderzoeksvoorstel en in de aanvraag is in sterke mate aangesloten bij de opzet van de experimenten in de andere AMvB gemeenten, in het bijzonder die van de gemeente Tilburg. Het evaluatieonderzoek van het experiment wordt ook door dezelfde groep van onderzoekers aan de Universiteit van Tilburg uitgevoerd (zie Muffels & van der Klink, 2017). In de gekozen opzet wordt conform de opzet in de andere experimenteergemeenten, Tilburg, Groningen, Deventer, Nijmegen en Utrecht) een ‘random control trial’ (RCT) experiment uitgevoerd waarbij deelnemers na aanmelding aselekt worden toegewezen aan de ‘treatments’ of interventies. In tegenstelling tot het vertrouwensexperiment in Tilburg zijn geen combinatie interventies maar enkelvoudige interventies ingericht in Wageningen. De aanvraag spreekt over de drie volgende interventies:

Interventie 1. Zelf in actie (in de AMvB: de ontheffingsgroep)

De treatment ‘zelf in actie’ (met ontheffing van de formele arbeids- en re-integratieverplichtingen) duidt op de treatment waarbij de ‘eigen regie’ van de cliënt richtinggevend is en eigen initiatief van de cliënt voorop wordt gesteld.

Interventie 2. Extra begeleiding (in de AMvB: de intensiveringsgroep)

De tweede interventie van intensivering van de begeleiding (in de AMvB: de intensiveringsgroep) bestaande uit vraaggerichte begeleiding op maat is erop gericht mensen te ondersteunen en kansen te bieden waarbij de persoonlijke situatie en wensen van de cliënt centraal staan.

Interventie 3. Zelf verdienen loont (in de AMvB: de vrijlatingsgroep)

De derde interventie duiden we aan als 'zelf verdienen loont'. Het betreft het verruimen van het behoud van inkomsten uit deeltijd of voltijd werk (in de AMvB: de vrijlatingsgroep). Overigens wordt in deze groep de reguliere dienstverlening aangeboden en gelden dus ook de algemene arbeids- en re-integratie verplichtingen.

Interventie 4. De onderzoeksgroep of controlegroep

De onderzoeksgroep of controlegroep krijgt de standaard treatment ofwel reguliere dienstverlening aangeboden die alle andere niet-deelnemers aan het experiment ook krijgen.

Referentiegroep

Tot slot wordt een vergelijkingsgroep onderscheiden die bestaat uit alle bijstandsgerechtigden die wel voldoen aan de gehanteerde toelatingscriteria voor deelname maar die niet deelnemen aan het experiment en het onderzoek.

Wetgeving: de invoering van de Participatiewet per 1 januari 2015

Sinds 1 januari 2015 vallen bijstandsuitkeringen onder de regelingen van de Participatiewet (PW), die de oude Wet werk en bijstand (Wwb) heeft vervangen. Doel van de Participatiewet is om zoveel mogelijk mensen naar betaald werk te helpen of, als dat niet kan, op een andere manier te laten participeren. De wet regelt de uitgifte van bijstandsuitkeringen en geeft een kader voor de regels en verplichtingen die bijstandsgerechtigden moeten naleven. Hoewel de wet de nationale kaders vastlegt, ligt de verantwoordelijkheid voor de uitvoering van de wet bij de gemeenten. Het doel van de wet was om meer mensen sneller te laten uitstromen naar betaald werk (zie SCP rapport, 2019). Na invoering van de wet op 1 januari 2015 is bij meerdere gemeenten de behoefte ontstaan om de doeltreffendheid van de nieuwe wet te onderzoeken, met andere woorden of de Participatiewet de gestelde doelen bereikt; werkt het goed en zo niet, wat en hoe kan het dan beter werken. Doeltreffendheid richtte zich daarbij zowel op de toeleiding naar betaald werk als ook op de bevordering van de participatie in de samenleving. Bij het tweede doel kan gedacht worden aan vrijwilligerswerk, psycho-sociale of andere hulpverlening zodat mensen beter om kunnen gaan met hun persoonlijke en/of sociale problemen en daardoor beter kunnen participeren. In algemene zin vroegen de gemeenten zich af of de P-wet voldoende is toegerust om ondersteuning te kunnen bieden voor verbetering van de sociale participatie door middel van het bevorderen van de zelfredzaamheid, gezondheid en welbevinden van de bijstandsgerechtigde welke hiervoor bepalende factoren zijn. Bovendien vonden een aantal gemeenten de regelingen en verplichtingen in de Participatiewet niet altijd goed aansluiten bij de huidige praktijk. Het stelsel van regels en verplichtingen in de Participatiewet werd door gemeenten als complex en bureaucratisch ervaren, met weinig ruimte voor maatwerk en vormen van begeleiding waarbij de bijstandsgerechtigde zélf meer regie

krijgt. Uiteindelijk zijn 6 gemeenten aan de slag gegaan om te experimenteren met een andere manier van begeleiding op basis van een AMvB, zoals opgenomen in artikel 83 van de Participatiewet. Daarnaast zijn er vijf gemeenten die buiten artikel 83 om experimenteren; dat zijn Apeldoorn en Epe, Oss, Geldrop-Mierlo en Amsterdam.

Probleemstelling en doel Wageningse experiment

In de aanvraag is de centrale probleemstelling van het Wageningse Vertrouwensexperiment als volgt verwoord: “In hoeverre kunnen de doelstellingen van bevordering van (arbeids)participatie beter bereikt worden door een andere bejegening en een andere handelingsvrijheid enerzijds (eigen regie), gecombineerd met positieve financiële prikkels ofwel het belonen van eigen initiatief door middel van een hogere inkomstenvrijlating, alsmede door een vraaggerichte, intensieve begeleiding op maat”. Vervolgens wordt in verband met de gewenste uitkomsten gesteld dat:

“Het experiment is voor ons een succes als het experiment een positieve bijdrage levert aan de (arbeids)participatie en aan de zelfredzaamheid, welbevinden en (mentale) gezondheid van de cliënt en een bijdrage levert aan een duurzamer resultaat op de uitstroom en/of maatschappelijke participatie dan de standaard benadering. We willen weten wat (beter) werkt, dus zelfs als er geen positieve uitkomsten zijn is het toch een interessant experiment, waarbij elke uitkomst, zowel positief als negatief waardevol kan zijn. We zijn dus benieuwd naar de uitkomsten van het experiment en wat het effect is op het vinden van werk maar zeker ook wat het betekent voor de motivatie, gezondheid en welbevinden van de cliënt” (einde citaat).

In de eerste voortgangsrapportage van de gemeente Wageningen van april 2018 (Muffels e.a. 2018) wordt het doel van het experiment “Vertrouwen Werkt” als volgt verwoord:

“Het doel is om te onderzoeken wat het effect is van een andere benadering van mensen in de bijstand, gebaseerd op vertrouwen en een positieve bejegening. Met het vertrouwensexperiment wil de gemeente onderzoeken of naast het reguliere bijstandsbeleid een aanpak die is geënt op intrinsieke motivatie en een aanpak die is geënt op een ruimere mogelijkheid tot bijverdienen leiden tot een toename van uitstroom en bijverdiensten en/of een vorm van participatie, en/of een toename van welbevinden en/of een toename van de financiële zelfredzaamheid van mensen” .

Opzet van het onderzoek: product -en procesevaluatie?

Het onderzoek bestaat uit twee onderdelen: de product –en procesevaluatie (zie Muffels & van der Klink 2017). De *productevaluatie* is erop gericht na te gaan wat wel en niet werkt en waarom voor de deelnemende bijstandsgerechtigden; de nieuwe aanpak van eigen regie en ontheffing van verplichtingen, extra maatwerkbegeleiding en extra vrijlating van verdiensten of de reguliere aanpak met een strenge toetsing van verplichtingen en de toepassing van maatregelen en kortingen. Dit gebeurt aan de hand van de ‘primaire’ uitkomstmaat uitstroom naar werk en de ‘secundaire’ uitkomstmaten gezondheid, welbevinden, zelfvertrouwen, sociale participatie en zelfredzaamheid. Er worden drie vragenlijsten bij de deelnemers online afgenomen met behulp van “Qualtrics” een software programma.

De *procesevaluatie* is erop gericht na te gaan hoe de projectleider(s) en klantmanagers als professional het proces van implementatie van het experiment in de gemeente percipiëren en waarderen, en welke meerwaarde de klantmanagers ervaren van de andere bejegening van de deelnemer voor de uitoefening van hun taken in de nieuwe aanpak. Daartoe worden regelmatig focusgroep gesprekken gehouden met de projectleider en de klantmanagers. Het eerste gesprek vond plaats na 3 maanden nadat het experiment was gestart en het eerste gesprek met de cliënt was gehouden. Daarna zijn steeds na ruim 6 maanden nog twee focusgroep gesprekken gehouden. Aan het einde van het experiment, na 1 oktober 2019, zijn door de projectleider exit gesprekken gevoerd met de consultants. De verslagen daarvan zijn ter beschikking gesteld aan de onderzoekers. Onderdeel van de procesevaluatie was ook het online invullen van een korte vragenlijst door de klantmanager na het gesprek met de deelnemer over de afspraken en bevindingen van het coaching gesprek en de begeleiding in de afgelopen periode. In totaal zijn door de klantmanagers of consultants drie van dergelijke vragenlijsten ingevuld. Tot slot is nog een korte vragenlijst afgenomen over alle door de klantmanager begeleide deelnemers tezamen.

Vrijwillige deelname

In de gekozen opzet wordt conform de opzet in de andere experimenteergemeenten, Groningen, Utrecht, Wageningen, Deventer en Nijmegen) een 'random control trial' (RCT) experiment uitgevoerd waarbij deelnemers aselect worden toegewezen aan de 'treatments' of interventies. Dit is gebeurd nadat deelnemers zich hebben aangemeld voor deelname aan het experiment. Deelname aan het experiment is vrijwillig. Na toewijzing aan een groep kunnen deelnemers alsnog afzien van deelname.

Inhoud voorlopige eindverslag

Dit voorlopig eindverslag is als volgt opgezet. Eerst wordt de opzet van het experiment en het onderzoek nader toegelicht. Dit houdt in dat in hoofdstuk 2 de beleidsuitgangspunten en filosofie van het experiment, het theoretisch kader, zowel wetenschappelijk als beleidstheoretisch, de werving van de deelnemers, de keuze voor de treatmentgroepen, de wijze van steekproeftrekking, de gestelde onderzoeksvragen en de manier van dataverzameling worden toegelicht. Vervolgens worden in hoofdstuk 3 de resultaten besproken van de beschrijvende analyses van de product -en procesevaluatie op basis van de vragenlijsten. Hoofdstuk 4 bespreekt in dit voorlopige verslag summier de belangrijkste conclusies van het onderzoek.

Gebruik van CBS microdata en planning aanvullende rapportage

Dit eindverslag is een voorlopig eindverslag. Het rapporteert over de effecten van het experiment op basis van de informatie die op lokaal niveau bij de gemeenten is verzameld (vragenlijsten deelnemers en procesevaluatie vragenlijsten klantmanagers) dan wel beschikbaar is middels de administratie van de gemeente Wageningen (BUS ofwel bijstandsuitkeringsstatistiek). In overleg met de projectleider van de gemeente Wageningen is afgesproken dat een update van de informatie volgt nadat de effecten op volledige en gedeeltelijke uitstroom naar betaald werk ook zijn geanalyseerd door de onderzoekers en het CPB op de microdata van het CBS. In deze aanvullende rapportage zal worden ingegaan op de vergelijking van de resultaten voor Wageningen van de lokale BUS-analyse en de CPB analyse wat betreft de primaire uitkomstmaat uitstroom naar werk. Deze aanvullende rapportage zal in mei 2020 beschikbaar komen. De analyse op de CBS data gebeurt in gezamenlijk overleg met de onderzoekers verenigd in het landelijk overleg experimenten Participatiewet (LOEP). De zes AMvB-gemeenten maken gebruik van de

microdatabestanden van het CBS om de effecten op de zogenaamde primaire uitkomstmaat te analyseren. Deze analyse gebeurt mede in overleg met het Centraal Planbureau dat door het Ministerie van Sociale Zaken en Werkgelegenheid is verzocht om de evaluatie van deze uitkomstmaat voor haar rekening te nemen. De rapportage daarvan staat gepland begin mei 2020. Door gebruik te maken van de door het CBS bewerkte microdata en een iets andere analysemethode kunnen de uitkomsten van de beide analyses verschillen. De onderzoekers zullen deze verschillen in de aanvullende rapportage in kaart brengen en van toelichting voorzien opdat eenduidige beleidsconclusies mogelijk zijn. In het technische deel van de rapportage die later volgt worden de analyses op de microdata verwerkt en wordt een meer gedetailleerde en technische toelichting gegeven op de manier waarop het onderzoek is uitgevoerd. Dit gedeelte is bedoeld voor lezers die meer achtergrondinformatie willen hebben van het onderzoek. Dit deel zal onderdeel uitmaken van het definitieve verslag begin mei 2020.

1.2 Theoretisch kader

Tegen de achtergrond van de vragen die de gemeenten zich stelden over de doeltreffendheid van de nieuwe Participatiewet die per 1 januari 2015 werd ingevoerd en de wens in de gemeente Wageningen en andere gemeenten om te onderzoeken of een andere aanpak 'beter werkt' zijn door de onderzoekers naast de reguliere dienstverlening drie specifieke treatments uitgewerkt (eigen regie, intensieve begeleiding en vrijlating) welke zijn gebaseerd op verschillende theoretische inzichten uit de literatuur (zie Muffels & van der Klink, 2017). Voor de derde treatment van 'extra vrijlating van verdienen' bestond er tot op dat moment geen wettelijke basis, daarvoor moest het experimenteer artikel 83 van de Participatiewet ruimte bieden. De theoretische begrippen die aan de basis van de ontwikkelde treatments staan zijn: wederkerigheid, vertrouwen, intrinsieke motivatie, keuzevrijheid, belonen (ook niet-financieel door positieve bejegening), zelfredzaamheid en individueel maatwerk (afgestemd op de wensen en mogelijkheden van de bijstandsgerechtigde). Deze begrippen zijn afgeleid uit de gedragseconomie (belonen, vertrouwen, intrinsieke motivatie), psychologische cognitie -en motivatietheorieën (wederkerigheid, zelfredzaamheid en intrinsieke motivatie) en de 'capability' theorie van Amartya Sen (keuzevrijheid en maatwerk) (zie ook Groot, Muffels en Verlaat, 2019). Deze zullen hieronder nader worden toegelicht. Eerst gaan we in op de achterliggende beleidstheorie die aan de basis heeft gestaan van de vragen die de gemeente zichzelf stelde bij aanvang van het experiment.

Beleidstheorie

In artikel 83 van de Participatiewet wordt als 'primaire uitkomstmaat' van het experiment de uitstroom naar voltijd betaald werk genoemd. Dat betekent dat wanneer mensen minimaal 27 uur gaan werken tegen het minimumloon zij volledig onafhankelijk worden van de uitkering en dus uitstromen uit de bijstand die immers voor een alleenstaande 70% van het minimumloon bedraagt. Het primaire doel van de Participatiewet is zoals opgemerkt om meer mensen sneller naar voltijds betaald werk toe te leiden en daardoor participatie in de samenleving te bevorderen. Lukt dit niet dan zijn tussenstapjes op weg naar dit doel (activering) evenwel ook mogelijk (deeltijdwerk, vrijwilligerswerk, cursussen, training en opleiding, beschut werk, proefplaatsingen, aanpassingen werkplek en loonkostensubsidie voor (gedeeltelijk) arbeidsongeschikten. Loonkostensubsidie en beschut werk zijn nieuw in de P-wet. In de praktijk bleek dit doel zelfs in de herstelperiode van de arbeidsmarkt vanaf 2015 dus na de economische crisis moeilijk

haalbaar. De arbeidsmarkt herstelde zich langzaam in de eerste jaren na de ernstigste naoorlogse economische recessie die duurde van 2008 tot einde 2014. Vanaf 2017 ging het beter en veranderde de arbeidsmarkt hetgeen in de jaren 2018 en 2019 leidde tot een erg krappe arbeidsmarkt. Het aantal tijdelijke banen bleef hoog en zelfs stijgen maar tegelijkertijd nam het aantal vaste banen toe. Dit gebeurde in de tijd dat het experiment van start ging. Het experiment vond dus plaats in een gunstig economisch tij waardoor met name jongere bijstandsgerechtigden een plekje konden vinden op de arbeidsmarkt. Voor velen in de bijstand (mensen met slechte gezondheid, ouderen, geen startkwalificatie) was dat ook in deze periode niet mogelijk. De stap naar betaald werk is voor een groot deel van de bijstandsgerechtigden dus te groot; voor velen zijn tussenstapjes nodig voordat een terugkeer in betaald werk mogelijk is, voor anderen in het zogenaamde “granieten bestand” is uitstroom naar betaald werk een onhaalbare kaart zelfs op de wat langere termijn. Zij hebben speciale ondersteuning nodig met een specifieke maatgerichte aanpak in verband met hun persoonlijke en maatschappelijke problemen en problemen met hun fysieke of mentale gezondheid.

In de eerste tussenrapportage in april 2018 is met behulp van latente klassenanalyse¹ een analyse gemaakt van de deelnemers aan het experiment op basis van informatie in de eerste vragenlijst (nulmeting). Het zoek- en sollicitatiegedrag van de deelnemer is geanalyseerd plus de belemmeringen die zij ervaren onder andere vanwege hun gezondheidsproblemen en een laag niveau van arbeidsgeschiktheid (inclusief hun kwalificaties). Later is deze analyse herhaald op alle mensen die de nulmeting hebben ingevuld en waarvoor de informatie beschikbaar was (n=279). Het beeld dat hieruit komt is hetzelfde. De beste oplossing gezien de gebruikte ‘model fit’ indices was een indeling in drie groepen (zie ook Muffels & Gielens, 2019). De eerste groep noemen we de ‘actief, arbeidsgeschikte’ groep die bijna 43% van de populatie uitmaakt (42.6%); zij achten zich arbeidsgeschikt, zoekt nog actief naar betaald werk en ziet nog zeker mogelijkheden voor het vinden daarvan. De rest van de populatie, dat is meer dan de helft (57%) is passief op zoek en kan vanwege mentale of fysieke gezondheidsproblemen geen werk vinden. Deze groep valt uiteen in twee groepen; enerzijds een ‘passief, arbeidsongeschikte’ groep met fysieke gezondheidsproblemen die niet (meer) op zoek is omdat gezondheidsproblemen het vinden van werk in de weg staat. Dat is het geval voor 33% van de deelnemers (n=93). Anderzijds, is er een groep die zichzelf arbeidsongeschikt acht vanwege mentale gezondheidsklachten, niet meer actief zoekt en berustend of ontmoedigd is vanwege het zoeken naar werk en het niet slagen daarvan (n=67 of 24%). De groep die zichzelf arbeidsongeschikt acht vanwege mentale gezondheidsproblemen, ontmoedigd en passief is, en daarom niet meer zoekt naar werk omvat dus een kwart van de deelnemers in Wageningen. Deze groep vraagt om een specifieke aanpak en een passende beleidstheorie of filosofie om deze adequaat te begeleiden naar participatiemogelijkheden inclusief (aangepast) werk. De filosofie achter het vertrouwensexperiment alsook het design van de treatments die

¹Latente klasse analyse probeert aan de hand van een aantal waargenomen variabelen (bevroegd in de nulmeting) een achterliggende) niet waargenomen variabele te vinden (zie <https://nl.wikipedia.org/wiki/Latenteklassemiddel>). Minimaal bestaat deze variabele uit twee categorieën verwijzend naar twee latente klassen. Een latente klasse wordt in de analyse weergegeven door voorwaardelijke kansen die voor iedere mogelijke score op de waargenomen variabelen wordt berekend. Deze kans geeft aan hoe waarschijnlijk een deelnemer die tot een van de klassen behoort een bepaalde score heeft op de waargenomen variabelen. De drie klassen oplossing voor Wageningen is technisch gezien uitstekend: er is een sterk onderscheid tussen de klassen en het aantal classificatiefouten is zeer gering (1,9%). De analyses zijn uitgevoerd met het programma Latent Gold 5.1.

eruit voortkwam was niet specifiek ingericht op deze problematische groep. Wel werd in het onderzoeksvorstel gesteld dat ook deze groep baat kan hebben bij de voorgestelde treatment van intensieve begeleiding omdat deze meer vraaggericht is en gebaseerd is op maatwerk. In het onderzoeksvorstel (Muffels & Klink, 2017) is daarom verwezen naar experimentele studies van IPS (“individual placement support”)-projecten en ES (“employment support”)-programma’s die met name voor deze groep soelaas kunnen bieden (Bond et al. 2010; Marshall et al. 2014). Dergelijke IPS/ES-programma’s richten zich evenwel niet alleen op de aanbodzijde maar ook op de vraagzijde van de arbeidsmarkt hetgeen beter past bij een werkgeversaankpak waarin bijvoorbeeld ook mogelijkheden voor aanpassing van het werk worden meegenomen. Het vertrouwensexperiment beperkte zich evenwel tot de aanbodzijde.

Vanuit het besef van een problematische arbeidsmarkt voor deze kwetsbare groepen hebben de gemeenten bij aanvang van het experiment de wens geuit om naast de primaire uitkomstmaat van volledige en gedeeltelijke uitstroom naar betaald werk ook een aantal secundaire uitkomstmaten te gebruiken om te bepalen of het experiment al dan niet werkt voor de deelnemers. Juist voor mensen waarvoor de stap naar betaald werk te groot is kunnen deze secundaire uitkomsten van groot belang zijn om later ook andere stappen te kunnen zetten. De uitkomstmaten die de gemeente Wageningen heeft gekozen zijn: vertrouwen, keuzevrijheid, bevordering van zelfredzaamheid, welbevinden en gezondheid (fysiek en mentaal) en sociale participatie. Daarnaast werd belang gehecht aan een andere bejegening van de burger uitgaande van vertrouwen, individueel maatwerk en aandacht in plaats van ‘loslaten’, ‘weinig of geen contact’ en wantrouwen welke begrippen in de ogen van de gemeente beter passen bij de reguliere aanpak. Veel gemeenten die vanaf de invoering van de nieuwe P-wet werden geconfronteerd met de volledige verantwoordelijkheid voor de financiering van de bijstandsuitkeringen middels de BUIG-middelen zochten een weg om de wet doeltreffender uit te voeren. Een stringente toetsing op verplichtingen en voorwaarden was een van de manieren om dat te doen. Een aantal gemeenten zijn die ‘strenge’ weg op gegaan, anderen zochten juist naar mogelijkheden om de effectiviteit van de re-integratie te verbeteren door bijvoorbeeld meer vraaggerichte begeleiding en maatwerk te bieden. In het kader van het Vertrouwensexperiment is ook in Wageningen gezocht naar een andere aanpak waarin eigen regie en minder stringente toepassing van de verplichtingen een intensievere en meer op maat gesneden begeleiding en ondersteuning wordt aangeboden vooral ook aan mensen in het zogenaamde “zittende’ bestand dat ook bekend staat onder de geuzennaam “granieten” bestand.

Wetenschappelijk kader

Het experiment is theoretisch ontleend aan een viertal inzichten, die ontleend zijn aan een brede onderzoeksliteratuur in de gedragseconomie, de empirische sociologie en de cognitie -en motivatiepsychologie (zie Groot, Muffels en Verlaat, 2018):

- Het eerste inzicht gaat over recente onderzoeksresultaten over de invloed van armoede op de ‘mindset’ ofwel geestesgesteldheid van mensen. Onderzoek in dit relatief nieuwe onderzoeksgebied laat zien dat (financiële) schaarste en stress door armoede de cognitieve hulpbronnen van mensen verminderen (Manietal, 2013; Mullainathan en Shafir, 2013). Men heeft alle energie, geestkracht en inventiviteit nodig om de eindjes voor henzelf, eventuele partner en

kinderen aan elkaar te knopen door het ontwikkelen van zogenaamde huishoudbezuinigingsstrategieën om in de eigen behoeften en die van het huishouden te kunnen voorzien. Indien financiële schaarste en het voldoen aan bijstandsverplichtingen een groot deel van de cognitieve hulpbronnen van mensen innemen, is er weinig ruimte voor belangrijke en cognitief uitdagende taken, zoals het bijscholen voor een andere baan, het onderhouden van het sociaal netwerk of het actief zoeken naar betaald werk (Groot, Muffels en Verlaat, 2019).

- Het tweede inzicht is afkomstig uit de gedragseconomie en gaat over de invloed van instituties op de arbeidsmarkt en in de sociale zekerheid alsmede de impliciete waarden die zij vertegenwoordigen op het gedrag van mensen. De impliciete waarden die ten grondslag liggen aan deze instituties en die tot uiting komen in uitkeringsvoorwaarden, regels en verplichtingen beïnvloeden het gedrag van de gebruikers in de vorm van bijvoorbeeld het zoekgedrag op de arbeidsmarkt, maar ook de keuze tussen werken en vrije tijd in de vorm van bijvoorbeeld het aantal uren dat men wil werken. Daarnaast beïnvloeden zij andere economische beslissingen in de sfeer van consumptie en bestedingen, huisvesting, onderwijs en gezondheid niet alleen van henzelf maar ook van anderen in het huishouden waaronder ook eventuele kinderen. Onderliggende waarden zijn bijvoorbeeld reciprociteit (wederkerigheid, 'tit for tat') en vertrouwen. Reciprociteit betekent dat individuen een goede behandeling of het ontvangen van vertrouwen (een investering in sociale relaties), belonen door bijvoorbeeld zich extra in te spannen (positieve reciprociteit), terwijl ze het omgekeerde doen als ze slecht worden behandeld of op basis van wantrouwen worden bejegend (negatieve reciprociteit) (Fehr en Schmidt, 2003). Negatieve prikkels, zoals opgelegde financiële sancties en kortingen zijn niet noodzakelijkerwijze de beste manier om mensen te prikkelen mee te werken en zich aan de regels te houden. Economisch en sociologisch onderzoek laat zien dat er weliswaar bescheiden korte termijn effecten optreden (men accepteert sneller werk ook al is de 'fit' minder) maar dat de effecten op de langere termijn onzeker zijn bijvoorbeeld doordat een slechte 'fit' tussen mens en werk op de iets langere termijn welbevinden, gezondheid en de loopbaan negatief beïnvloeden (zie SCP, 2019; Blonk, 2018). Bevindingen uit de experimentele economie laten ook zien dat mensen in ruil voor het vertrouwen dat zij ontvangen extra gemotiveerd zijn en meer hun best doen voor hun taak en daarmee dus ook degene die hen vertrouwt beloont (Groot, Muffels en Verlaat, 2019). Op deze manier leidt vertrouwen dus tot gevoelens van positieve reciprociteit en daardoor ook tot blijvende inzet en toegenomen productiviteit (Bohnet et al., 2001).
- Het derde inzicht is afkomstig uit de psychologische motivatie theorie en leert ons dat extrinsieke prikkels de intrinsieke motivatie kunnen wegnemen (Frey en Jegen, 2001). De zelfbeschikkingstheorie (Deci en Ryan, 1985) stelt dat intrinsiek gemotiveerde mensen een activiteit ondernemen omdat ze het plezierig en interessant vinden, waardoor ze meer effectiviteit en volharding laten zien in hun gedrag en een verbeterd welbevinden (Ryan et al., 1997). Deze theorie stelt ook dat het geven van vertrouwen aan mensen een gevoel van zelfsturing oplevert, die weer effect heeft op het werkzoekgedrag en duurzame werkgelegenheid. Voorgaand onderzoek wijst ook uit dat de intrinsieke motivatie kan worden versterkt door een activiteit als keuze aan te bieden in plaats van als controlemiddel (Groot, Muffels en Verlaat, 2019).

- Het laatste inzicht gaat over ‘vrije keuze’ en komt voort uit Sen’s ‘capability theorie’. Binnen deze theorie zijn ‘capabilities’ de keuze opties die mensen worden aangeboden om de dingen te doen die zij belangrijk vinden in hun eigen leven. Op deze manier krijgen mensen kansen aangeboden waardoor ze persoonlijke autonomie en zelfvertrouwen ontwikkelen en daarmee hun welbevinden kunnen vergroten (Sen, 2004; 2009). Op basis van de inzichten van Sen wordt in recente theorievorming rond duurzame inzetbaarheid het waarde-aspect van werk benadrukt alsook het belang van eigen regie (Van der Klink et al., 2016). De focus op keuzemogelijkheden en eigen regie vergroot dus ook de intrinsieke motivatie om te gaan werken.

Deze vier inzichten vormen de uitgangspunten voor de manier waarop de eerder genoemde alternatieve interventie-benaderingen (of treatments) zijn vormgegeven binnen het vertrouwensexperiment. Tevens zijn ze inspiratiebron geweest voor het opstellen van vragen in de vragenlijsten aan de deelnemers en de coaches. De inzichten zijn niet alleen relevant voor de productevaluatie waarin het effect wordt onderzocht op arbeidsinschakeling en op welbevinden, gezondheid en sociale participatie maar ook voor de procesevaluatie omdat de procesevaluatie onderzoekt in hoeverre sturing op zelfredzaamheid, intrinsieke motivatie en keuzevrijheid ook beter werkt in de uitvoering. We hebben aan het begin de volgende hypothesen geformuleerd.

Hypothesen

- Op grond van gedragseconomische inzichten verwachten we dat positieve prikkels in de vorm van bijvoorbeeld een andere, positievere en meer ondersteunende bejegening met minder verplichtingen en gebaseerd op vertrouwen leidt tot een grotere motivatie en tot meer coöperatief gedrag bij de cliënten hetgeen zich uit in een grotere bereidheid tot deelname alsook inzet om de re-integratie tot een succes te maken. We verwachten positieve effecten op motivatie en daarmee op participatie en arbeidsinschakeling in deze groepen in vergelijking met de controlegroep.
- Omgekeerd verwachten we op grond van het wetenschappelijk empirische onderzoek dat een systeem van strenge handhaving, monitoring en sanctionering zoals bij de controle –en referentiegroep positieve effecten kan hebben op inschakeling in werk maar dat de vraag is of deze effecten korte termijn effecten zijn en dat duurzame uitstroom mogelijk minder zal zijn bij deze groep dan bij de treatment groepen
- Een streng systeem van controle en handhaving kan negatieve effecten hebben op de motivatie(motivatietheorieën) en de mentale belasting en stress die het systeem van controle en straffen met zich meebrengt met negatieve gevolgen voor welbevinden en de (mentale) gezondheid.
- In overeenstemming met de uitkomsten van het Fit4Work programma en de uitkomsten van review studies naar de effecten van IPS (individual placement support) en ES (‘employment support’) interventies verwachten we sterke effecten van intensieve begeleiding op maat op de arbeidsinschakeling van vooral de meest kwetsbare groepen in de bijstandpopulatie (het zittende bestand).

- De focus op eigen regie en zelfredzaamheid geeft mensen handelings- en keuzevrijheid terwijl intensieve begeleiding en maatwerk hen extra keuzemogelijkheden en kansen biedt. Hierdoor neemt hun zelfvertrouwen en welbevinden toe waardoor het zoekgedrag toeneemt alsook de intrinsieke motivatie om te gaan werken hetgeen uiteindelijk ook de arbeidsinschakeling bevordert. Dit verwijst naar Sen's 'capability' theorie en Ryan en Deci's zelfdeterminatie theorie.

Procesevaluatie

De procesevaluatie beoogt inzicht te krijgen in hoe het proces van experimenteren verloopt in de spreekkamer van de professional waar de ontmoeting met de burger plaatsvindt. Hoe vertaal je de treatments naar de uitvoeringspraktijk; wat betekent eigen regie en hoe intensief is intensief in de treatment intensieve begeleiding? Hoe worden de treatments uitgevoerd, op welke manier wordt een vertrouwensband met de deelnemer opgebouwd, welke condities belemmeren een goede uitvoering van de treatments etc.? De beschikbare wetenschappelijke literatuur over 'wat en waarom werkt' in het re-integratie beleid is beperkt. In de oratie van Blonk (2018) wordt gesteld dat de meeste studies gericht zijn op beleid en beleidsevaluatie meer dan op het begrijpen van de individuele gedragsverandering en het bieden van handvatten aan de professional. Evaluatiestudies van re-integratie beleid laten zien dat de effecten van dat beleid (bemiddeling, werkcreatie, loonkostensubsidies) niet erg groot zijn, en veelal korte termijn effecten laten zien (zie de overzichtsstudies van Card, Kluge en Weber, 2010; 2015). De effecten van het activerend arbeidsmarktbeleid bedragen gemiddeld over alle studies voor de middellange termijn (12 tot 24 maanden) 11,5%. Voor experimenteel onderzoek bedraagt het gemiddeld percentage 5,6%. Dit betreft het gemiddelde van begeleiding bij het zoeken naar werk, werk creatie in publieke sector, controle/sancties, subsidies voor private werkgelegenheid en training. Dit betreft dus zowel mensen in de werkloosheidsverzekering als in de bijstand. De studies tonen wel aan dat training en ontwikkeling dus investeren in menselijk kapitaal effectiever is dan de andere instrumenten voor de wat langere termijn. Een recente studie van een soortgelijk 2-jarig experiment met intensieve begeleiding en maatwerk in Lausanne in Zwitserland tussen 2015-2017 laat zien dat het effect op uitstroom naar werk gemiddeld zo'n 9% bedraagt over de gehele periode (Bigotta e.a. 2018). Internationaal onderzoek toont ook aan dat Intensieve begeleiding van mensen met een psychische beperking werkt en effectief om deze mensen naar werk toe te leiden. De effecten hiervan zijn groter dan de hierboven genoemde effecten en variëren gemiddeld tussen 10 tot 25% (gemiddeld 15%). De meest sprekende voorbeelden van succesvolle begeleiding naar werk van mensen met (mentale en fysieke) gezondheidsbeperkingen komen dan ook uit de psychologische literatuur rond "Individual Placement Support" en "Employment Support" (zie Bond et al. 2010; Marshall et al. 2014). Daarin wordt specifiek en gericht individueel maatwerk geboden (zowel aan de aanbod als de vraag -of werkgeverskant) welke leidt tot de genoemde additionele uitstroompercentages.

Blonk (2018) stelt in zijn oratie daarom het volgende: "Je moet dus investeren in mensen en dat vooral volhouden. Dat is het beste dat je kunt doen. Maar we weten ook dat niet elke training helpt en ook niet voor iedereen in gelijke mate. Om daar meer grip op te krijgen is inzicht nodig in de werkzame bestanddelen van en de condities rond een training. Niet alleen wat werkt, maar waarom, zodat de professional onderbouwde keuzes kan maken wat wordt ingezet. Die kennis is nog maar beperkt beschikbaar. Er is wat dat betreft nog steeds een hoog 'black box' gehalte. Om die kennis te ontwikkelen zijn meer studies nodig

waarin naar gedragsaspecten wordt gekeken. Aspecten zoals kwaliteit van zoekgedrag, naar de intensiteit daarvan en de focus, naar vertrouwen in eigen in dat zoekgedrag en naar vaardigheden die zelfregie ondersteunen en hoe op deze aspecten te interveniëren. Wat dat betreft zijn we nog maar net begonnen.” (citaat oratie Blonk, 2018). Vervolgens verwijst Blonk naar studies van Liu, Wang en Huang uit 2014 waarin wordt onderzocht wat werkt bij een aantal specifieke sollicitatietrainingen die hebben geleid tot een driemaal grotere kans om werk te vinden voor de deelnemers. Het is volgens Blonk de combinatie van motivatie en gedrag die bepalend zijn voor het succes van deze trainingen. Deze overwegingen hebben aan de basis gestaan van de procesevaluatie en de invulling van de ‘skills’ training die aan klantmanagers is gegeven om deelnemers te leren zelfredzaam te zijn of te worden. De basis van de training wordt gevormd door de theorie van ‘planned behavior’ ofwel het integratieve gedragsmodel van Fishbein & Ajzèn (2010). In het model (Fig. 1.1) staat zoekgedrag naar werk centraal. Het gaat om intenties ofwel motivatie en vaardigheden. De motivatie is afhankelijk van het belang van het te bereiken doel, de ervaren druk van de omgeving en hoe je daarmee omgaat en de verwachting dat je het doel kunt halen. Bij vaardigheden gaat het om vaardigheden die je ten dienste staan bij het eigen regie nemen en houden over je leven, over zelfredzaamheid, volhouden en niet bij de pakken te gaan neerzitten ondanks je belemmeringen. Deze vaardigheden zijn te leren. De laatste component in het model zijn de mentale en fysieke beperkingen maar ook kenmerken zoals leeftijd, nationaliteit of culturele achtergrond. Het omgaan met deze beperkingen en kenmerken is onderdeel van zelfredzaamheid en eigen regie. De training is enkele keren in Wageningen gegeven. In de drie focusgroep gesprekken met de klantmanagers is de betekenis en rol van de training in het uitvoeringsproces besproken. In de verslaglegging van deze focusgroep gesprekken wordt hier nader op ingegaan.



Fig. 1.1. Het integratieve gedragsmodel van Fishbein & Ajzèn (2010).

Bron: Oratie Roland Blonk (2018):

Vanuit wetenschappelijk oogpunt is nu vooral van belang om te toetsen of als de uitvoering niet tot het gewenste resultaat leidt dit het gevolg is een 'theory failure', de uitgangspunten zijn niet goed geformuleerd of een 'uitvoeringsfailure', de treatments zijn niet conform de onderliggende uitgangspunten uitgevoerd. De vragenlijsten bij de coaches en de focusgroep gesprekken hebben tot doel om hier meer inzicht in te krijgen. Daarbij analyseren we in samenhang de invloed van contextuele factoren (politiek, gemeentelijke uitvoering), het gedrag van de coaches en deelnemers, de bedoelde en onbedoelde effecten van de treatments, en de uitkomsten en effecten.

1.3 Beleids -en onderzoeksvragen: hypothesen

Het doel van het onderzoek is om te achterhalen wat het effect is van de drie alternatieve interventies of aanpakken op de gekozen uitkomstmaten in vergelijking met de standaard aanpak. Het idee is dat door een andere meer positieve bejegening (onder andere inhoudende uitgaan van vertrouwen en geen wantrouwen richting deelnemer, belonen eerder dan straffen, vraaggericht en "wat heb jij nodig om de dingen te doen die je wilt doen?") en een andere manier van het 'gesprek' aangaan, meer vanuit intrinsieke motivatie en vertrouwen dan vanuit extrinsieke prikkels en wantrouwen, het gedrag positief wordt beïnvloed in de richting van meer eigen initiatief en zelfredzaamheid. Uiteindelijk wordt daarmee ook het hoofddoel van de P-wet om meer mensen sneller toe te leiden naar betaalde arbeid beter bereikt, althans dat is de verwachting.

Wetenschappelijk gezien zijn de experimenten interessant omdat er nog weinig bekend is over de effectiviteit van een alternatief beleid van motiveren en belonen (positieve prikkels) in plaats van een beleid vooral gericht op financiële prikkels inclusief het straffen en sanctioneren dat erbij hoort (negatieve prikkels). De doeltreffendheid van het beleid is voor dit onderzoek breed gedefinieerd. Dat houdt in dat er wordt gekeken naar effecten op vijf domeinen: (1) arbeidsparticipatie, zowel in voltijds werk als in deeltijds werk (2) sociale participatie en vertrouwen, (3) gezondheid en welbevinden en (5) de financiële situatie van de deelnemers (financiële stress en armoede). De doeltreffendheid wordt vastgesteld door de drie alternatieve interventies in beginsel te vergelijken met de reguliere aanpak in de onderzoek -of controlegroep. In het onderzoeksvoorstel en de aanvraag bij het Ministerie om te worden erkend als artikel 83 gemeente zijn destijds de volgende beleids-en onderzoeksvragen gesteld:

Beleids -en onderzoeksvragen

- In het experiment worden de verplichtingen, die gekoppeld zijn aan de inspanningen die van cliënten worden verwacht om te participeren en re-integreren, bijvoorbeeld de sollicitatie -en re-integratieplicht versoepeld (met name ook in de eigen regiegroep). De vraag voor beleid is dan wat is het effect van deze ontheffingen op de gekozen uitkomstmaten?
- Een tweede beleidsvraag betreft het effect op de arbeidsinschakeling wanneer mensen op een positieve vraaggerichte manier worden ondersteund en intensiever en op maat worden begeleid (met name intensieve begeleidingsgroep).
- Een derde vraag die de gemeente stelde was in hoeverre met vrijlating van verdiensten conform artikel 83 van de Participatiewet mensen meer voltijds dan wel meer in deeltijdwerk aan de slag zullen gaan en/of het aantal uren uitbreiden.

- Een vierde beleidsvraag voor de gemeente was of mensen die deelnemen aan het experiment meer ruimte ontwikkelen om eigen initiatief te ontplooiën, meer in brede zin gaan participeren, en wat het effect hiervan is op de arbeidsinschakeling (eigen regie, intensieve begeleiding).
- Een belangrijke vraag voor de gemeente was tevens wat de neveneffecten zijn van de ruimere handelingsvrijheid op maatschappelijke participatie, maar ook op zelfredzaamheid, welbevinden en gezondheid (eigen regie, intensieve begeleiding).
- In het experiment wordt onderzocht wat de uitkomsten over de tijd gezien zijn van meer handelingsvrijheid en een intensievere begeleiding op maat. Dat vraagt een andere manier van ondersteuning en andere kwalificaties van de klantmanagers. De klantmanagers krijgen ook meer tijd (lagere caseload) om de deelnemer te begeleiden. De gemeente wil graag weten hoe dit proces verloopt. Worden de treatments uitgevoerd zoals beoogd en in welke mate zijn ze onderscheidend? Hoe verloopt de interactie met de klant? Wat werkt in de uitvoering, in de spreekkamers beter en wat werkt niet ook vanuit het perspectief van de klant? Welke condities, aspecten maken het verschil en wat betekent dit voor de implementatie van de uitkomsten (procesevaluatie).

Analysemodel productevaluatie

Het integratieve gedragsmodel fungeerde als theoretische basis voor de wijze van uitvoering van de treatments. Daarbij staan sturing op intrinsieke motivatie en gedrag centraal. In de productevaluatie worden vervolgens de effecten onderzocht van deze treatments op de gekozen uitkomstmaten zoals uitstroom naar werk, welbevinden, gezondheid, sociale participatie en zelfredzaamheid ervan uitgaande dat de treatments conform de theoretische uitgangspunten zijn uitgevoerd. Daarbij wordt rekening gehouden met factoren die deze effecten beïnvloeden zoals de sociaaleconomische en culturele context, de positie op de arbeidsmarkt, de beleving en keuzemogelijkheden en het zoekgedrag. Dit wordt in Figuur 1.2 weergegeven.

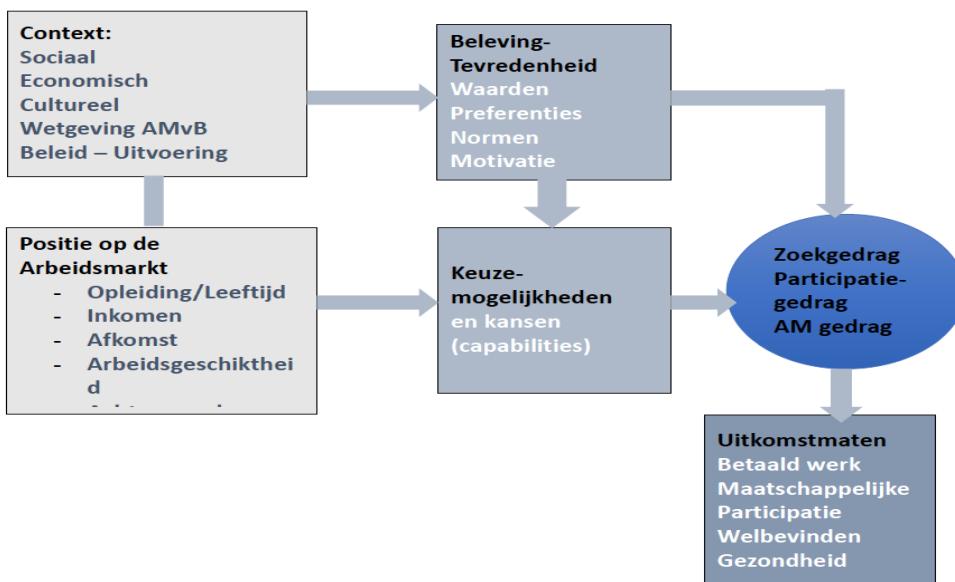


Fig. 1.2. Analysemodel onderzoek effecten Vertrouwensexperimenten PW

2 Opzet van het onderzoek

2.1 Experimentgroepen

In Wageningen zijn zoals eerder opgemerkt vier treatment of experimentgroepen onderscheiden. Hieronder volgt een nadere uitwerking en invulling van de gekozen treatments op basis van de beschrijvingen die aan het begin van het experiment door de gemeente en de projectleider en in overleg met de klantmanagers zijn opgesteld. Dit is beschreven in een werkdocument bij de start van het experiment dat is opgesteld door de projectleider samen met het team van klantmanagers (Zuidema, 2017). Hierin wordt de methodiek van begeleiding op basis van een individueel plan van aanpak met een specifiek omschreven doel of doelen en bijbehorende inzet van instrumenten beschreven. Dit document geldt als leidraad voor de uitvoering van het experiment door de klantmanagers. Daarin wordt ook de bestaande dienstverlening ('care as usual') omschreven. In het werkdocument wordt hierover opgemerkt dat de beschrijving van de bestaande dienstverlening is uitgewerkt in een notitie 'Werk en participatie in Wageningen – de huidige uitvoering' en de hierbij horende 'bijlage bij Werk en Participatie in Wageningen'. In het werkdocument wordt bij de beschrijving van de onderzoekgroep de doelgerichte methodiek als volgt toegelicht: "In de context van het vertrouwensexperiment zal voor alle mensen in de bijstand de basis zijn het kijken naar het op de persoonlijke mogelijkheden afgestemde doel: wat is passend gelet op de mogelijkheden van betrokkene richting werk en participatie?". De feitelijke omschrijvingen gebruikt in het experiment zijn uitgebreider dan hieronder weergegeven. Wekelijks vindt er overleg plaats tussen de projectleider en de klantmanagers over de werkwijze, voortgang en casuïstiek.

Experimentgroep 1: Zelf in actie (in de AMvB: de ontheffingsgroep)

De treatment 'zelf in actie' (met ontheffing van de formele arbeids- en re-integratieverplichtingen) duidt op de treatment waarbij de 'eigen regie' van de cliënt richtinggevend is en eigen initiatief van de cliënt voorop wordt gesteld. Dit treatment is erop gericht de cliënt op een andere manier te bejegenen en minder controle uit te voeren, de regeldruk te verminderen, mensen te waarderen in hun proactief gedrag en in beginsel geen maatregel op te leggen. Deelnemers aan dit treatment kunnen – net als anderen in de bijstand – een beroep blijven doen op ondersteuning van de gemeente en het inzetten van voorzieningen. Deze ondersteuning en het inzetten van voorzieningen is volledig op basis van het gekozen uitgangspunt van regie vanuit betrokkene zelf.

Experimentgroep 2: Extra begeleiding (in de AMvB: de intensiveringsgroep)

De tweede interventie van intensivering van de begeleiding (in de AMvB: de intensiveringsgroep) bestaande uit vraaggerichte begeleiding op maat is erop gericht mensen te ondersteunen en kansen te bieden waarbij de persoonlijke situatie en wensen van de cliënt centraal staan. Het totale pakket aan dienstverlening in het sociale domein staat hiervoor ter beschikking. Hierbij wordt een integrale aanpak beoogd. Het uitgangspunt van vraaggerichte maatwerkbegeleiding is richtinggevend voor deze treatment. De deelnemers hebben vaker contact en krijgen speciale aandacht. Hiervoor zijn in beginsel tot 6 contact momenten per jaar mogelijk.

Experimentgroep 3: Zelf verdienen loont (in de AMvB: de vrijlatingsgroep)

De derde interventie duiden we aan als 'zelf verdienen loont'. Het betreft het verruimen van het behoud van inkomsten uit deeltijd of voltijd werk (in de AMvB: de vrijlatingsgroep). Ook gelden hier de algemene arbeids- en re-integratie verplichtingen. Voor deze interventie wordt voorzien in een 'goede' ondersteuning van mensen bij het verrekenen van inkomsten uit deeltijdwerk. Er wordt de standaard of reguliere dienstverlening aangeboden en er gelden dezelfde verplichtingen als in de controlegroep. De vrijlatingsbepalingen zijn conform artikel 83 van de Participatiewet op basis waarvan het experiment is ingericht. Daarbij kunnen deelnemers voor de duur van het experiment 50% van hun verdiensten houden (die worden dus niet gekort op de uitkering) tot een maximum van €199,- per maand.

Experimentgroep 4: De onderzoeksgroep of controlegroep

De onderzoeksgroep of controlegroep krijgt de standaard ofwel reguliere dienstverlening aangeboden die alle andere niet-deelnemers aan het experiment ook krijgen. Voor hen verandert er niets, ook de vrijlating van verdiensten zoals die gold vóór deelname aan het experiment blijft gehandhaafd. De vrijlatingsregeling in de Participatiewet houdt in dat 25% van de verdiensten in de eerste 6 maanden na instroom in de bijstand niet worden ingehouden op de uitkering. Het maximum dat kan worden vrijgelaten bedraagt €199,- per maand. Bijstandsgerechtigden in de controlegroep zijn deelnemers aan het experiment en het onderzoek. Dat geldt niet voor de referentiegroep.

Referentiegroep(niet-deelnemers)

Daarnaast is er een vergelijkingsgroep van alle niet-deelnemers die ook de reguliere dienstverlening krijgen maar niet meedoen aan het onderzoek, de zogenaamde referentiegroep. In Wageningen hebben zij wel een uitnodiging gekregen maar hebben ze zich niet aangemeld. Deze groep is niet gerandomiseerd. De referentiegroep van niet-deelnemers biedt de mogelijkheid om te onderzoeken in hoeverre de aanmeldingen selectief zijn geweest maar kan ook dienen als extra check op de uitkomsten voor de 'primaire' uitkomstmaat voltijdse uitstroom naar werk. Door selectie bij de aanmeldingen of afmeldingen kunnen de treatment en controlegroepen qua samenstelling afwijken van de oorspronkelijke doelpopulatie van bijstandsgerechtigden of de groep van niet deelnemers. In de aanvraag is geopteerd om alle niet-deelnemers (exclusief diegenen die zich hebben afgemeld) maar die wel tot de doelpopulatie behoren, dus aan de toelatingscriteria voor deelname aan het experiment voldoen, tot de referentiegroep te rekenen. Bij een aantal van bruto 410 deelnemers zijn dat 387 bijstandsgerechtigden die op 1 oktober 2019 een lopende uitkering hadden en 171 bijstandsgerechtigden die tussen 1 oktober 2017 en 1 oktober 2018 instroomden in de bijstand (bij elkaar 968 bijstandsgerechtigden).

Experiment of 'Hawthorne' effecten

Hiervoor is de doelgerichte methodiek die in Wageningen is gebruikt toegelicht. Deze is zoals beschreven in het werkdocument ook van toepassing op de reguliere dienstverlening in de onderzoek -of controlegroep. Dat betekent dat bij de start reeds de reguliere dienstverlening in de onderzoeksgroep werd afgestemd op deze methodiek die voor alle treatments het uitgangspunt vormde in het experiment. De reguliere dienstverlening kan daarnaast veranderen door verandering in de uitvoering in de loop van de

tijd. Dit heeft te maken met autonome beleidsontwikkelingen maar wellicht ook met uitstralingseffecten van het experiment op de begeleiding van de bijstandsgerechtigden in de reguliere dienstverlening omdat de klantmanagers van de controle – of onderzoeksgroep onderdeel uitmaken van het team dat het experiment uitvoert. Doordat men als team opereert en wekelijks overlegt over de werkwijze in de verschillende treatments en de casuïstiek kunnen er zogenaamde cross-over effecten zijn opgetreden tussen de treatments. Dit geldt temeer aangezien de schaal van het experiment in Wageningen groot is en het experiment expliciet is ondersteund door de organisatie. Veel klantmanagers hebben ook meer dan een treatment voor hun rekening genomen. Dat geldt ook voor de begeleiding van de onderzoeksgroep waarbij vier klantmanagers 73% van de deelnemers onder hun hoede hadden maar tegelijkertijd ook andere groepen begeleiden. Dergelijke cross-over effecten kunnen zijn versterkt door deelname aan het experiment en het onderzoek (invulling van procesevaluatie vragenlijsten). De treatments kunnen elkaar dan beïnvloeden en zijn niet meer hetzelfde als bij aanvang of zoals bij de niet-deelnemers. In de RCT-experimenten literatuur spreekt men dan over gedragseffecten in dit geval bij de klantmanagers die de zuivere vergelijking tussen de vier experimentgroepen bemoeilijkt. De deelname aan het experiment door de deelnemers en de speciale behandeling die ook zij daarbij krijgen zoals een uitnodiging bij de start van het experiment aan bijeenkomsten met projectleider en klantmanager, het contact met klantmanagers over het invullen van de vragenlijsten (de deelnemers worden enkele keren eraan herinnerd de vragenlijst in te vullen en hebben dan bijvoorbeeld telefonisch contact met de klantmanager) kunnen leiden tot gedragseffecten bij de deelnemers waardoor zij zich anders gedragen dan bijstandsgerechtigden die niet aan het experiment deelnemen. Dit wordt in de literatuur het ‘Hawthorne’ effect genoemd².

Dit heeft uiteraard in dat geval invloed op het onderzoek naar de effecten van het experiment. Doordat een referentiegroep is onderscheiden van bijstandsgerechtigden in de Participatiewet die niet heeft deelgenomen aan het experiment kan worden nagegaan of bij aanmelding selectie is ontstaan waardoor de deelnemers niet meer representatief zijn voor de doelpopulatie. Deze referentiegroep biedt echter ook de mogelijkheid om bijvoorbeeld de uitstroom naar werk uitkomsten van de vier treatment groepen te vergelijken met deze vergelijkingsgroep en daarbij te corrigeren voor verschillen in compositie. Immers, de bijstandsgerechtigden in deze groep kunnen administratief worden gevolgd doordat de gemeente de reden van beëindiging (op basis van een beëindigingsonderzoek) registreert.

Beleidswijzigingen

De onderzoekers zullen uiteraard in het onderzoek met deze mogelijke verandering in de reguliere begeleiding sinds de start van het experiment per 1 oktober 2017 rekening proberen te houden. Van belang is dat ook de overige ‘treatments’ zoveel als mogelijk is in stand worden gehouden en niet veranderen maar ook dat is niet vanzelfsprekend aangezien beleid voortdurend verandert en het experiment invloed heeft op de begeleiding door de klantmanagers in de treatmentgroepen. In de praktijk blijkt dit een lastige opgave omdat er experiment -en leereffecten optreden bij de klantmanager waardoor de treatment zich kan wijzigen. Daarnaast worden de treatments uitgevoerd door klantmanagers die voorheen de reguliere

² Een variant hierop is het “John Henry” effect waarbij er competitie (“rivalry”) ontstaat tussen de treatments en de controlegroep waarbij klantmanagers in de controlegroep zich extra inzetten voor bijvoorbeeld uitstroom naar werk om niet onder te doen voor de collega’s of wanneer deelnemers zich extra gemotiveerd tonen voor deelname aan het experiment, ook in de controlegroep, om de begeleiding succesvol te laten verlopen.

dienstverlening uitvoerden. De kans dat na verloop van tijd men toch terugvalt op de ‘oude’ werkwijze is aanwezig zeker bij “moeilijke” cases waar weinig vooruitgang of beweging te bespeuren valt. Dit kan tot gevolg hebben dat treatments in de praktijk minder verschillen dan beoogd en de effecten daardoor kleiner zijn. De vergelijking met de referentiegroep (wat uitstroom naar werk betreft) kan om deze redenen belangrijke additionele informatie opleveren, vooropgesteld dat de treatmentgroepen en de referentiegroepen wat samenstelling betreft vergelijkbaar zijn. Deze vergelijking met de referentiegroep is overigens alleen mogelijk bij de uitkomstmaat uitstroom naar werk, maar niet bij de overige uitkomstmaten van welbevinden, gezondheid, zelfredzaamheid, keuzemogelijkheden en sociale participatie. Daarvoor ontbreekt een referentiegroep omdat over deze uitkomstmaten geen administratieve data beschikbaar zijn.

2.2 Procedure va random selectie en toewijzing plus werving

Voor het onderzoek is er gebruik gemaakt van een RCT design, dat wil zeggen een gerandomiseerd experiment met een controlegroep³, waarbij deelnemers aselekt worden toegewezen aan de vier ‘treatments’ of interventiegroepen. De doelgroep omvatte alle bijstandsgerechtigden in Wageningen met uitzondering van enkele specifieke groepen die zijn uitgesloten van deelname aan het experiment omdat ze niet voldeden aan de uitsluitingscriteria. Uitgesloten zijn mensen die geen recht (meer) hebben op een uitkering Participatiewet, mensen die in een inrichting verblijven, dak -en thuislozen, mensen die gedurende het experiment met pensioen gaan (>65 jaar), mensen die de Nederlandse taal niet machtig zijn, mensen die volledig arbeidsongeschikt zijn en mensen die andere regelingen zitten (IOAW, IOAZ, BBZ) bij aanvang van het experiment.

Uitnodiging, werving en afmelding

In Wageningen is besloten om alle 724 bijstandsgerechtigden die per 31 september 2017 tot de doelpopulatie behoorden (91% van de totale bijstandspopulatie) uit te nodigen voor deelname aan het experiment en voor deelname aan een informatiebijeenkomst. Voor hen die niet wensten mee te doen was een ‘nee’ kaartje bijgevoegd. De anderen werden uitgenodigd voor een gesprek met de klantmanager. Nadat men zich heeft aangemeld werd men aselekt (randomisatie) toegewezen aan een bepaalde groep. Vervolgens werd men uitgenodigd voor een gesprek en werd middels ondertekening van de deelnemersovereenkomst de deelname bevestigd. Zoals in de tussenrapportages is vermeld geschiedde aanmelding op vrijwillige basis. Na indeling in een groep mochten deelnemers alsnog weigeren deel te nemen aan het experiment. Ze kunnen dan niet worden geplaatst in een andere groep.

Afmeldingen

Het aantal deelnemers dat op eigen verzoek is gestopt is opvallend klein (17 deelnemers of 4%). We hadden verwacht dat na toewijzing in de standaard of onderzoeksgroep het aantal afmeldingen daar het grootste en aanzienlijk zou kunnen zijn. Dat is niet gebeurd. De persoonlijke benadering van de Wageningse aanpak en van de cliënten door de klantmanagers heeft waarschijnlijk tot dit gunstige resultaat geleid. Tabel 1 geeft de aantallen aanmeldingen, uitstroom en afmeldingen weer per 1 oktober 2019. Per 1 oktober 2019 zijn er

³ Engels: Randomised Controlled Trial (RCT).

410 aanmeldingen geregistreerd waarvan 18 afmeldingen (17 op eigen verzoek plus een door de gemeente). Het netto aantal deelnemers komt daarmee op 392. De meeste afmeldingen (11) dateren van bij de start. Daarna zijn er nog 6 afmeldingen op eigen verzoek bijgekomen. Dit betreft vooral mensen die vanwege persoonlijke, fysieke -en mentale gezondheid redenen alsnog afhaken van het onderzoek mede vanwege de belasting van het onderzoek (invullen vragenlijsten) of in verband met de persoonlijke situatie. Het aantal afmeldingen in Wageningen vanwege non-compliance met het experiment/onderzoek is daardoor klein, zo'n 4%.

Tabel 2.1: Aantallen bereikte deelnemers in Wageningen incl. afmeldingen en uitstroom, cijfers per 1-10-2019

VEX-groep (deelnemers en afmeldingen)	Wageningen	%
Deelnemer in begeleiding (exclusief herinstroom)	229	56%
Afgemeld	18	4%
Uitstroom naar werk	80	20%
Uitstroom vanwege verhuizing	25	6%
Uitstroom vanwege studie	12	3%
Herziening uitkering plus overige uitstroom	46	11%
Bruto aantal deelnemers	410	100%
Afgemeld om uiteenlopende redenen	18	4%
Netto aantal deelnemers	392	96%

Bron: Gegevens Vertrouwensexperiment Wageningen en eigen berekeningen onderzoekers, Tilburg University, 1 oktober 2019

2.3 Onderzoekpopulatie: deelnemers en afmeldingen

In Tabel 2.2 wordt de verdeling van de afmeldingen over de treatments weergegeven. Tevens geeft de tabel de bereikte aantallen weer voor elk van de vier groepen inclusief en exclusief de 18 afmeldingen en het percentage deelnemers dat voltijds is uitgestroomd naar werk per 1 oktober 2019 (op basis van uitstroomgegevens ultimo 2019) plus de overige uitstroom vanwege verhuizing, studie en dergelijke.

Vergelijking tussen treatments, controlegroep en referentiegroep

Het percentage voltijdse uitstroom naar werk in de drie treatment groepen ligt met 20,4% een stuk hoger dan de 11%-15% die we in het eerste voortgangsrapport berekenden voor de periode 2016-2018 voor respectievelijk lopende en nieuwe uitkeringen (Muffels e.a. 2018). Deels heeft de stijging evenwel te maken met de betere arbeidsmarktsituatie maar deels mogelijk ook met de interventies aangezien het verschil met de gerealiseerde uitstroom in de controlegroep per 1 oktober 2019 ongeveer 4% bedraagt en met de referentiegroep van niet-deelnemers uit het zittende bestand, zelfs bijna 11%. Het verschil tussen de beide treatment groepen eigen regie en extra begeleiding met de onderzoeksgroep is net niet significant op 90% betrouwbaarheidsniveau ($p < 0.12$). Het verschil van de vrijlatingsgroep met de controlegroep is niet significant. Echter, we vinden wel significante verschillen tussen de deelnemers uit het zittende bestand in de beide treatmentgroepen eigen regie en extra begeleiding en de referentiegroep op respectievelijk 99%

($p < 0.01$) en 95% betrouwbaarheidsniveau ($p < 0.05$). De referentiegroep is evenwel niet gerandomiseerd zoals de controlegroep waardoor we niet weten in hoeverre dit verschil causaal geïnterpreteerd kan worden en nader onderzoek nodig is in hoeverre deze 4-10% stand houdt bijvoorbeeld na correctie op compositieverschillen tussen de groepen. De overige uitstroom bedraagt voor de deelnemers 38% zodat in totaal bijna drie van de vijf deelnemers (58%) na 2 jaar is uitgestroomd uit de bijstand.

Tabel 2.2: Deelnemers inclusief afmeldingen naar treatmentgroep en feitelijke uitstroom naar werk op einddatum per 1 oktober 2019

Treatment groepen	Bruto deeln. N=	Afmeldingen N=	Netto deeln. ²⁾ N=	Uitstr. werk ³⁾ N=	Uitstr. werk in % ³⁾	Uitstr. Overig N=	Uitstr. Overig In %
Eigen regie (zelf in actie)	106	8	98	23	23,5%	40	37,0%
Extra begeleiding (op maat)	98	4	94	22	23,4%	41	41,8%
Vrijlating (verdiene loont)	113	4	109	18	16,5%	41	36,3%
Totaal treatments	317	16	301	65	21,6%	122	38,5%
Onderzoeksgroep (controle)	93	2	91	15	16,5%	34	36,6%
Totaal deelnemers	410	18	392	80	20,4%	156	38,0%
Referentiegroep A ¹⁾	385	-	385	42	10,9%	110	28,6%
Referentiegroep B ¹⁾	171	-	171	53	31,0%	94	55,0%

Noot: ¹⁾ Referentiegroep A: zittend bestand per 1 okt. 2017 en B: nieuwe instroom tussen 1 okt. 2017 en 1 okt. 2018;

²⁾ Het netto aantal deelnemers (C) betreft alle deelnemers (A=bruto) minus de afmeldingen (B).

³⁾ Uitstroom naar voltijds werk betreft beëindiging van de uitkering door het vinden van een voltijdse baan als werknemer, als zelfstandige of middels een loonkostensubsidie, c.q. loonwaardebepaling. De uitstroom in % is berekend op basis van het netto aantal deelnemers exclusief afmeldingen zoals weergegeven in de derde kolom.

Bron: Vertrouwensexperiment Wageningen, Tilburg University, 2020

Reguliere dienstverlening in onderzoeksgroep

Het is dus vooral nog niet geheel duidelijk waarom voltijdse uitstroom van deelnemers naar betaald werk in de onderzoeksgroep significant hoger is (15,9% versus 10,9%) dan in referentiegroep A ($p < 0.10$). Deelnemers hebben zich bewust aangemeld voor het experiment en zich dus niet afgemeld nadat hen werd medegedeeld dat ze in de controlegroep zijn ingedeeld. Het betreft daardoor mogelijk een extra gemotiveerde groep van bijstandsgerechtigden, aangezien er voor hen niet veel te halen is in deze groep in de vorm van extra vrijlating of extra begeleiding (extra inzet van instrumenten). Uit de focusgroep gesprekken met de klantmanagers maken wij op dat de deelnemers het nut en de noodzaak om mee te doen snel inzagen waardoor het mogelijk een selectieve groep is die extra gemotiveerd is om een bijdrage te leveren aan het verbeteren van de werking van de Participatiewet. Wellicht dat vanwege die motivatie zij meer kans hebben om uit te stromen.

Wat de klantmanagers van de controlegroep betreft behoorden zij tot het team dat het experiment uitvoerde en dat onderling wekelijks overlegde over de werkwijze en methodiek van begeleiding en de bejegening in geval van moeilijke casuïstiek. Er is door de projectleiding in samenspraak met de klantmanagers een werkdocument opgesteld waarin afspraken zijn opgenomen over de werkwijze en methodiek. Op deze manier is er een soort intervisie in de groep ontwikkeld waarvan de klantmanagers van

de onderzoeksgroep onderdeel uitmaakten. In Wageningen is weliswaar een training volgens de “skills” methodiek aangeboden maar deze is slechts twee keer ook daadwerkelijk gehouden en daarna gecontinueerd in de vorm van intervisie binnen het team. Niet uitgesloten kan worden dat deze werkwijze van invloed is geweest op de aanpak in de controlegroep maar daarvoor is nader onderzoek nodig. In het eindrapport van de gemeente over het experiment wordt op deze materie tevens ingegaan (zie pag. 10). Indien de reguliere treatment niet veranderd zou zijn is er geen reden om een verschil tussen onderzoeks- of controlegroep en referentiegroep te verwachten. De referentiegroep is evenwel als vergelijkingsgroep dan geschikter mits de deelnemers een goede afspiegeling vormen van de doelpopulatie. Is dat niet het geval dan heeft er selectie plaatsgevonden bij de aan- of afmeldingen voor het experiment. Indien de deelnemersgroep geen goede afspiegeling is van de populatie zal de vergelijking met de referentiegroep gecorrigeerd moeten worden voor compositieverschillen. Daartoe is onder andere het gebruik van lineaire regressie -en matchingmethoden geschikt.

Representativiteit

Doordat de deelnemers aselekt zijn toegewezen aan de vier interventiegroepen verwachten we geen grote verschillen tussen de deelnemers en de doelpopulatie. Mensen moeten zich evenwel aanmelden waardoor er zelfselectie zou kunnen ontstaan. In Tabel 2.5 worden de deelnemers vergeleken met de referentiegroep. Zelfselectie kan ontstaan doordat mensen met bepaalde kenmerken zich af -of aanmelden en andere niet. Het zou kunnen zijn dat lager opgeleiden of ouderen boven 55 jaar bijvoorbeeld doordat ze meer moeite hebben met invulling van online vragenlijsten zich minder snel aanmelden en dus ondervertegenwoordigd zijn. Voorts verwachten we dat mensen met weinig kansen op de arbeidsmarkt en vaak langere verblijfsduren in de bijstand, minder genegen zijn zich aan te melden. Om de representativiteit te onderzoeken is gebruik gemaakt van de BUS gegevens op individueel niveau bij de start van het experiment. Voor het zittende bestand zijn dat gegevens per 30 september 2017. Voor de nieuwe instromers zijn dat de gegevens ultimo van de maand die zo dicht mogelijk ligt bij de start van het experiment veelal de maand ervoor. De informatie beschikbaar in de BUS betreft gegevens zoals geslacht, leeftijd, duur uitkering, aantal keer uitkering bij start, nationaliteit, leefvorm en huisvestingssituatie. Gegevens over opleiding waren slechts bekend voor deelnemers die de vragenlijst hebben ingevuld maar niet voor de anderen. De referentiegroep omvat 556 mensen waartoe mensen behoren met een lopende uitkering per 1 oktober 2017 (385) of mensen die tussen 1 oktober en 1 juli 2018 zijn ingestroomd in de bijstand maar hebben besloten niet deel te nemen aan het experiment (171). De drie treatment groepen betreffen alle 410 mensen die zich hebben aangemeld voor het experiment dus inclusief de 18 afmeldingen.

Nadere analyse (zie Tabel 2.5) laat zien dat er weinig verschillen zijn tussen de drie treatment groepen en de controlegroep. De controlegroep is iets ouder (43 versus 41), woont iets meer in gezinsverband en heeft iets langere duren (3 jaar versus 2.9 jaar). Door zelfselectie in het experiment (wie melden zich aan en wie melden zich af) zijn er evenwel verschillen ontstaan tussen de experimentgroepen en de referentiegroep, alle mensen die niet deelnemen aan het experiment en zich niet hebben afgemeld. Verschillen blijken te zijn ontstaan in verband met vooral nationaliteit, de duur van de uitkering, met leeftijd (iets meer ouderen en iets minder jongeren tot 27 jaar) en geslacht, mannen zijn licht oververtegenwoordigd. Vooral naar duur is echter sprake van een oververtegenwoordiging van duren van 1 tot 3 jaar en een ondervertegenwoordiging van lange duren (>3 jaar). De bijstandsgerechtigden met duren tot 1 jaar zijn evenredig vertegenwoordigd in de treatment groepen Dit betekent dat een vergelijking van de

uitkomstmaten en zeker ook de uitstroomkansen naar werk tussen de vier groepen die sterk afhankelijk is van de duur van de uitkering, rekening moet houden met deze onder -en oververtegenwoordiging naar vooral duur van de uitkering maar ook naar nationaliteit voorafgaand aan het experiment. Opvallend in Wageningen is dat er een oververtegenwoordiging is van buitenlanders in plaats van een verwachte ondervertegenwoordiging.

Tabel 2.5: Representativiteit onderzoekspopulatie (inclusief 18 afmeldingen) naar achtergrondkenmerken bij start van experiment, Wageningen, 1 oktober 2017-2018 (in %).

Achtergrond-kenmerken bij start	Categorie	Deelnemers	Eigen regie	Extra begeleiding	Vrijlating	Controlegroep	Refgroep A+B
Geslacht	Man	49.3	46.2	52.0	51.3	47.3	45.7
	Vrouw	50.7	53.8	48.0	48.7	52.7	54.3
Leeftijd	gemiddeld	41	39	42	41	43	41
	<27 jaar	13.9	20.8	12.2	13.3	8.6	21.6
	27-34 jaar	20.2	25.5	24.4	14.2	17.2	16.2
	35-44 jaar	23.9	17.0	16.3	32.7	29.0	14.6
	45-54 jaar	23.9	24.5	24.5	23.9	22.6	24.6
	55 jaar (tot 65)	18.1	12.3	22.5	15.9	22.6	23.0
Migratieachtergrond	Buitenlands	30.0	32.1	29.6	27.4	31.2	18.5
	Nederlands	70.0	67.9	70.4	72.6	68.8	81.5
Leefvorm	Allst. ouder	14.4	16.0	12.2	15.0	14.0	10.8
	Alleenstaande	58.8	57.6	62.2	59.3	55.9	65.5
	Samenwonend	26.8	26.4	25.5	25.7	30.1	23.7
Gemiddelde duur	In jaren	2.9	2.8	3.4	2.5	3.0	4.0
Bijstandsduur in klassen	Tot 1 jaar	50.2	58.5	48.0	51.3	41.9	50.0
	1 tot 3 jaar	21.7	17.9	21.4	23.9	23.7	13.0
	3 jaar of meer	28.1	23.5	30.6	24.8	34.4	37.1
Huisvesting	eigenaar	1.2	0	2.0	0.9	2.2	1.4
	huurder	92.9	93.4	88.8	94.7	94.6	89.9
	overig	5.9	6.7	9.1	4.4	3.3	8.7
Aantal	N=	410	106	98	113	93	556

Bron: BUS-gegevens Wageningen, 1 oktober 2019

Noot: Achtergrondkenmerken van deelnemers in de onderzoeksgroep (kolom 4) en in de twee experimentgroepen (kolom 2-3). T1: Eigen regie-Zelf aan het stuur; T2: Intensieve begeleiding-Op maat; Controlegroep=Onderzoeksgroep; Referentiegroep (niet-deelnemers).

Om voor deze selectie-effecten te corrigeren zijn logistische regressie en matchingtechnieken toegepast. In de logistische regressie modellen voor de uitstroomkans naar betaald werk corrigeren we voor verschillen

in duur of nationaliteit door de nationaliteit en duur bij de start van het experiment op te nemen in de regressievergelijking. In de matching procedure gaan we op zoek naar mensen in de bijstandpopulatie van niet-deelnemers die zoveel mogelijk lijken op mensen die aan het onderzoek deelnemen en dus ook qua duur uitkering, leeftijd, geslacht, nationaliteit en dergelijke vergelijkbaar zijn.

2.4 Product- en procesevaluatie

Het voorliggende onderzoek bestaat zoals eerder aangegeven twee onderdelen: de product- en procesevaluatie (zie Muffels & van der Klink 2017). De *productevaluatie* is erop gericht na te gaan wat beter werkt voor de deelnemer; de nieuwe vraaggerichte aanpak met meer handelingsvrijheid en extra vrijlating van verdiensten of de standaardaanpak met een strenge toetsing van verplichtingen en de toepassing van maatregelen en kortingen. Over een periode van 2 jaar zijn deelnemers aan het experiment gevolgd. In het onderzoek wordt niet alleen gekeken naar of mensen eerder betaald werk vinden maar ook wat het met deelnemers doet, met hun vertrouwen, gezondheid, hun welbevinden, hun zelfvertrouwen en hun zelfredzaamheid. Daartoe vulden deelnemers aan het begin van het experiment, na 1 jaar en na 2 jaar een online vragenlijst in. De eerste noemen we de nulmeting en de laatste de eindmeting. De uitkomsten van de analyses op deze vragenlijsten zullen een antwoord geven op eerder gestelde onderzoeksvragen. Wanneer deelnemers uitstromen naar werk of bijvoorbeeld verhuizen, blijven ze deelnemer en worden ze gevraagd de vragenlijsten in te vullen. De bereidheid om dat te doen zal evenwel naar verwachting geringer zijn bij de uitstromers waardoor de non-response naar verwachting zal toenemen naarmate meer mensen uitstromen. In het onderzoek worden degenen die zich hebben afgemeld na de start van het experiment ook na afmelding administratief gevolgd. Dat gebeurt via de gemeentelijke administratie (BUS) alsook door koppeling van de gegevens van deze oorspronkelijke deelnemers aan de bestanden van het CBS. Dit is voor een valide en zuivere meting van de uitkomsten van het RCT experiment noodzakelijk.

De *procesevaluatie* is erop gericht na te gaan hoe de projectleider(s) en consulenten/ (werk)coaches als professional het proces van implementatie van het experiment in de gemeente percipiëren en waarderen, en welke meer- of minderwaarde de consulenten/coaches ervaren van de andere bejegening van de cliënt voor de uitoefening van hun taken in de nieuwe aanpak. Daartoe zijn er regelmatig focusgroep gesprekken gehouden met de projectleider en de coaches. In totaal zijn er 3 focusgroep gesprekken gehouden. Het eerste gesprek vond plaats op 15 maart 2017 nadat het experiment enkele maanden van start was gegaan en het eerste gesprek met de cliënt is gehouden. Het tweede gesprek heeft op 9 oktober 2018 plaatsgevonden, het derde gesprek volgde op 21 maart 2019. Ten tweede vulde de coach aan het begin van het experiment, na 1 jaar en na 2 jaar een online vragenlijst in over de afspraken en bevindingen van de coaching gesprekken. Tezamen met het eerste focusgroep gesprek vormt dit de nulmeting van het procesevaluatiedeel. Ten slotte hebben de consulenten aan het einde van het experiment (november 2019) een afsluitende vragenlijst ingevuld over hun eigen ervaringen gedurende het experiment. Door het uitvoeren van deze procesevaluatie kan inzicht worden verkregen in de effectiviteit van de interventie, maar ook in de mate waarin ineffectiviteit is terug te voeren tot een 'theory failure' (beleidstheorie) of een 'implementation failure' (uitvoering). In de volgende paragrafen zal de dataverzameling van zowel de product- als procesevaluatie worden toegelicht.

2.5 Dataverzameling product- en procesevaluatie

Productevaluatie: vragenlijsten

De vragenlijsten bevatten vragen over onderwerpen zoals gezondheid en welbevinden, zelfredzaamheid, zoek- en sollicitatiegedrag, ervaren keuzemogelijkheden (capabilities), sociale participatie, sociaal vertrouwen, vertrouwen in overheidsinstanties en financiële situatie en inkomens -en deprivatiearmoede. Wanneer deelnemers uitstromen, bijvoorbeeld naar werk, blijven ze deelnemer en worden ze gevraagd de vragenlijsten in te vullen. In de praktijk blijkt dat ze zich dan sneller afmelden. Het gevolg hiervan is dat de respons afneemt naarmate het experiment vordert. Het doel van de eerste vragenlijst (de zogenaamde nulmeting) was om de situatie van de cliënt vast te stellen aan het begin van het experiment. Het invullen ervan gebeurde tijdens of direct na het eerste gesprek met de deelnemer aan het experiment. De tweede en derde vragenlijst (ook wel eerste en tweede vervolgmeting) zijn 1 jaar en 2 jaar na de start van de begeleiding ingevuld. De deelname aan het experiment is zoals eerder opgemerkt vrijwillig, men mag te allen tijde het experiment weer verlaten maar er wordt wel van een deelnemer verwacht dat deze meedoet aan het onderzoek en de vragenlijsten zorgvuldig invult. Daarover zijn afspraken gemaakt met de deelnemer, hoewel het in de praktijk soms moeilijk is gebleken om een deelnemer aan deze afspraak te houden zeker na de nulmeting.

Invulling deelnemer vragenlijst (productevaluatie)

Een persoonlijke benadering bleek uitermate belangrijk om deelnemers te stimuleren de online vragenlijst voor de nulmeting in te vullen. Zoals hiervoor toegelicht vraagt dat een persoonlijke aanpak. Het vraagt overredingskracht en uitleg van de klantmanager om de cliënt te overtuigen de vragenlijst volledig in te vullen. Vanaf de start van het experiment op 1 oktober 2017 jl. hebben de deelnemers aan het experiment de online vragenlijst ingevuld. De bereidheid bij de deelnemers om de tweede en vooral ook de derde vragenlijst in te vullen bleek sterk te zijn afgenomen na de nulmeting ondanks een intensieve belcampagne gericht op invulling door de deelnemer opgezet door de gemeente Wageningen. Veel vragen gaan over werk zoeken en vinden hetgeen voor de mensen die na 2 jaar nog geen werk hebben gevonden ondanks de andere bejegening en begeleiding geen haalbare kaart is waardoor de motivatie om de tweede en derde vragenlijst in te vullen veel geringer bleek. Dat heeft uiteindelijk geresulteerd in een respons van 60% voor de beide laatste deelnemersvragenlijsten tegenover 77% voor meting 1. Uiteindelijk is een respons bereikt van 70% hetgeen bij deze als 'moeilijk bereikbare' onderzoeksgroep bekend staande groep als 'goed' mag worden bestempeld zeker indien in beschouwing wordt genomen dat 40% van de deelnemers na 2 jaar om vele redenen is uitgestroomd uit de bijstand vanwege verhuizing, herzien recht op uitkering, overgang naar andere uitkeringen maar ook door het vinden van werk. In totaal heeft 20% in de periode van 2 jaar werk gevonden (inclusief uitstroom naar zelfstandigheid) en heeft bijna 20% van de deelnemers de uitkering beëindigd om andere redenen. De 765 door hen "online" ingevulde vragenlijsten⁴ zijn door de onderzoekers

⁴ De meeste deelnemers hebben de vragenlijst online ingevuld. De deelnemers kregen een door de Universiteit gemaakte persoonlijke weblink opgestuurd waarmee zij de vragenlijst in de thuissituatie online konden invullen. Alleen wanneer de deelnemer de voorkeur gaf aan invulling tijdens of na een gesprek met de klantmanager bij de gemeente werd die mogelijkheid gefaciliteerd. Sommigen hebben een schriftelijke vragenlijst ingevuld die vervolgens door de onderzoekers zelf online is ingevuld.

nader geanalyseerd. Voordat we de resultaten verder bespreken wordt hieronder inzicht geven in de bereikte aantallen ingevulde vragenlijsten per treatment groep.

Procesevaluatie vragenlijsten

Ten behoeve van de procesevaluatie zijn er ook periodiek vragenlijsten ingevuld door de klantmanagers die de deelnemers in begeleiding hebben. Voor elke deelnemer die de klantmanager in begeleiding heeft, vult deze een vragenlijst in. Net als bij de deelnemersvragenlijsten zijn deze klantmanager vragenlijsten drie keer afgenomen: aan het begin, na 1 jaar en na 2 jaar begeleiding (maar wel in volgorde van aanmelding van een deelnemer). In deze vragenlijsten zijn vragen gesteld over het proces van begeleiding, zoals de doelen die zijn ingezet in overleg met de deelnemer, de activiteiten die zijn ondernomen en de trajecten die zijn aangevraagd, de motivatie van de deelnemer voor de begeleiding door de consulent, een inschatting van de klantmanager over de kansen op de arbeidsmarkt van de deelnemer en de leereffecten van het experiment voor zowel de klantmanager als de deelnemer. Wanneer een deelnemer uitstroomt, bijvoorbeeld vanwege verhuizing, overgang naar een andere uitkering of het vinden van werk, dan is bij ontbreken van contact het lastig voor de klantmanager om de vragenlijst in te vullen.

Deelname onderzoek: ingevulde vragenlijsten, privacy en 'informed consent'

De onderzoekers hebben twee vragenlijsten uitgezet, een bij de deelnemer als onderdeel van de metingen ten behoeve van de productevaluatie en een bij de klantmanager als onderdeel van de procesevaluatie. Deelnemers die uitstromen bijvoorbeeld naar werk blijven deelnemer en worden geacht de vragenlijsten in te vullen. In het onderzoek worden degenen die zijn uitgestroomd of zich hebben afgemeld en/of geen vragenlijsten meer invullen ook na afmelding administratief gevolgd. Dat gebeurt via de gemeentelijke administratie (BUS) alsook door koppeling van de gegevens van deze oorspronkelijke deelnemers aan de bestanden van het CBS. Dit is voor een valide en zuivere meting van de uitkomsten van het RCT experiment noodzakelijk. De wijze van gegevensverzameling is getoetst aan de AVG privacy regels en aan de ethische procedures voor 'informed consent' welke gelden voor wetenschappelijk onderzoek. De persoonsgegevens (NAW) zijn door de gemeente geanonimiseerd met behulp van een anoniem identificatienummer waarbij alleen de gemeente en het CBS toegang tot de persoonsgegevens hebben. Daartoe is een bewerkingsovereenkomst met de gemeente gesloten.

In Tabel 2.4 worden de aantallen respondenten weergegeven van de drie vragenlijsten ingevuld door de deelnemers en de klantmanagers. In de eerste voortgangsrapportage was ruim 60% van de nulmeting vragenlijsten ingevuld, uiteindelijk is dat opgelopen tot 77%. Daarna neemt de bereidheid tot invulling geleidelijk af en is het responsepercentage in de derde meting gedaald naar 56%. De klantmanagers vullen alleen vragenlijsten in voor deelnemers die nog in begeleiding zijn dus niet meer wanneer mensen zijn uitgestroomd vanwege werk of andere redenen. In Wageningen hebben de klantmanagers voor nagenoeg alle deelnemers de vragenlijst ingevuld: 93% bij de eerste procesevaluatie meting tot 91% bij de laatste meting. Deze response percentages zijn voor onderzoek onder bijstandsgerechtigden relatief hoog te noemen zowel wat de metingen bij de deelnemers als de klantmanagers betreft.

Tabel 2.4: Ingevulde vragenlijsten product -en procesevaluatie naar treatment groepen in Wageningen

Groepen	Deelnemers			Klantmanagers		
	CV- meting 0	CV- meting 2	CV- meting 3	PCEV meting 0	PCEV meting 2	PCEV meting 3
	Eigen regie	90	56	57	96	63
Extra begeleiding	68	53	46	77	61	53
Vrijlating + regulier	88	72	64	93	86	75
Onderzoeksgroep (controle)	59	58	53	76	65	59
Totaal	306	239	220	342	275	243
%	77%	61%	56%	93%	87%	91%
N=	N=399	N=392	N=392	N=368	N=317	N=266

Bron: Vragenlijsten deelnemers en klantmanagers Wageningen, Tilburg University, gegevens ultimo 2019

Administratieve BUS data en koppeling aan CBS microdatabestanden

Zoals opgemerkt zijn begin februari 2020 de gegevens over de deelnemers geupload naar het CBS. In deze rapportage maken we vooral gebruik van de gemeentelijke BUS bestanden om de uitstroomkansen naar werk te bepalen maar ook om bij de berekening van deze kansen te kunnen corrigeren voor eventuele compositieverschillen tussen de treatment groepen. We zullen echter tevens de uitkomsten van de eigen analyses en van het CPB kort presenteren en deze vergelijken met de analyse op basis van de BUS bestanden. De BUS-bestanden bevatten informatie over demografische kenmerken van de bijstandspopulatie, de reden van beëindiging van de uitkering (geregistreerd op basis van een beëindigingsonderzoek), de arbeidsgeschiedenis, de aanwezigheid van schulden en de loonwaarde voor bijstandsgerechtigden met beperkingen die met loonkostensubsidie aan de slag zijn gegaan. Daarnaast zijn de data geupload naar het Centraal Bureau voor de Statistiek (CBS) om deze te verrijken met gegevens over werk (loopbaan) en inkomen, uitkeringen en andere achtergrondkenmerken.⁵ Hierdoor is het mogelijk om de loopbanen van de deelnemers te reconstrueren en ook gedeeltelijke uitstroom naar flexibele banen, (kleine) deeltijdbanen en vaste banen te onderzoeken. Door het gebruik van CBS-data zijn de gegevens tussen de experimenten met de Participatiewet bovendien beter vergelijkbaar.

⁵ In alle gebruikte databestanden werden persoonsgegevens bij voorbaat door middel van een sleutel vervangen door een betekenisloos administratienummer (pseudonimiseren). Middels dit nummer kunnen databestanden uit verschillende bronnen en over de tijd heen aan elkaar worden gekoppeld. Omdat geen persoon toegang heeft tot zowel de databestanden als de sleutel, is informatie niet naar deelnemers herleidbaar.

3 Resultaten productevaluatie

3.1 Productevaluatie en uitkomstmaten

In de productevaluatie staat niet alleen de uitkomstmaat voltijdse uitstroom naar betaald werk centraal maar ook de verbetering van de gezondheid, het welbevinden, het zelfvertrouwen, de keuzemogelijkheden, de sociale participatie en de zelfredzaamheid van de deelnemer. Eerst worden de resultaten besproken van de analyse op de uitkomstmaat volledige uitstroom naar betaald werk en daarna de andere. Onder volledige uitstroom verstaan we dat bijstandsgerechtigden vanwege uitstroom naar betaald werk volledig onafhankelijk worden van de uitkering. Bij een loon ter hoogte van het minimumloon betekent dit dat men minimaal 27 uur gaat werken (70% van 38-40 uur) aangezien de uitkering voor een alleenstaande 70% bedraagt van het volwassen minimumloon.

Deeltijdse uitstroom en uitstroom naar tijdelijk werk

Er zijn in de BUS data uitsluitend gegevens beschikbaar over volledige uitstroom uit de uitkering maar niet over deeltjduitstroom of over uitstroom naar tijdelijk werk. De gemeenten beschikken over uitkeringsgegevens van mensen in de bijstand op grond waarvan een goede inschatting gemaakt zou kunnen worden van deeltijdse uitstroom en voor hoeveel uren. Deze informatie is eveneens van de gemeente Wageningen ontvangen maar verder niet door ons geanalyseerd. In het verslag van de gemeente zijn deze gegevens aan een beschrijvende analyse onderworpen. In plaats daarvan zullen we in de definitieve rapportage in mei 2020 tevens de analyse van dezelfde type gegevens (zoals geregistreerd in de polis-administratie welke onderdeel vormt van de CBS microdatabestanden) opnemen.

De tweede en derde cliëntvragenlijst bevatten echter wel enige informatie van de deelnemer over zijn of haar situatie op het moment van invulling van de vragenlijst (zie Tabel 3.1). Deze informatie geeft evenwel geen betrouwbaar beeld van de uitstroom naar deeltijd -of tijdelijk werk omdat deze slechts door een deel van de deelnemers is ingevuld (60% response). Van de deelnemers is 1 en 2 jaar na de start bekend of ze een voltijdse, deeltijdse of tijdelijke baan hebben of dat ze andere activiteiten hebben ondernomen zoals vrijwilligerswerk of een opleiding of studie. Tevens is bekend of ze sinds de vorige keer een baan hebben aangenomen en welke type baan of contract dat is (zie Tabel 3.2).

Uit de gegevens in tabel 3.1 blijkt dat 36% van de respondenten zegt te werken als werknemer of als zelfstandige, 4% zegt te werken in een proefplaatsing of een stage en 7% in de sociale werkvoorziening. Ongeveer 8% zegt volledig aan het werk te zijn als werknemer of als zelfstandige. De meerderheid werkt evenwel in deeltijd (26%) of gedeeltelijk als zelfstandige (1.6%). Dat betekent dat ongeveer 42% van de deelnemers naar eigen zeggen aan het werk is in de tweejarige experimenteer periode. Daarnaast zegt 7% van de deelnemers vrijwilligerswerk te verrichten. De gunstige arbeidsmarktsituatie is hier zeker debet aan maar uiteraard ook de begeleiding door de klantmanagers. Daarnaast is gevraagd of men sinds de vorige keer (vorige meting) een baan heeft gevonden en wat voor type baan dat is. De respons in de tweede en derde cliëntvragenlijst bedroeg evenwel slechts 60% waardoor geen volledig beeld mogelijk is (zie Tabel 3.2).

Tabel 3.1: Type sociale participatie activiteiten, meting 2-3, Wageningen

Type activiteit	CV-Meting 2+3	
	N=	%
Voltijdse baan	27	6%
Deeltijdse baan, uitzendwerk, enkele uren werk	116	26%
Zelfstandige, volledig	7	1,4%
Gedeeltelijk als zelfstandige	8	1,6%
Werkstage, proefplaatsing, overig werk (voorbereiding werk)	19	4%
Opleiding, cursus of studie	37	8%
Gedeeltelijke/volledige ao-uitkering	26	6%
Sociale werkvoorziening	32	7%
Vrijwilligerswerk/mantelzorg	33	7%
Bijstand	136	30%
Verhuisd	6	1%
Ziek, overig	4	1%
Totaal	451	100%
Totaal N deelnemers meting 2+3=459		

Bron: Vragenlijsten eerste en tweede vervolgmeting deelnemers Wageningen, TiU, 2019

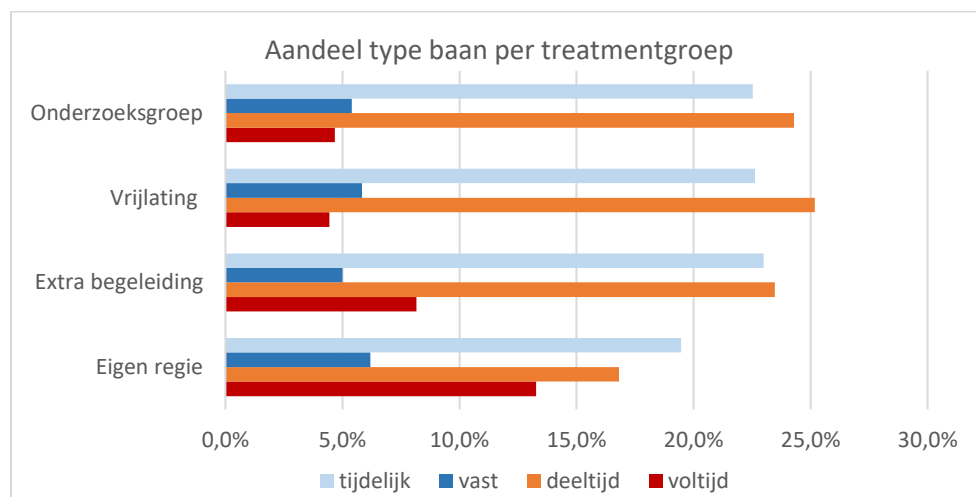
Tabel 3.2 laat zien dat een op de drie deelnemers zegt werk te hebben gevonden sinds de vorige meting (29,%). Daarvan heeft een op de vijf deelnemers die werk heeft gevonden een vast contract gekregen (19%). Ongeveer drie van de vier deelnemers die werk heeft gevonden komt in een tijdelijke baan terecht (75%). Een kleine 7% gaat als zelfstandige of freelancer aan de slag. Van de deelnemers die werk vinden gaat 8% werken in een uitzendbaan, dat is 2,5% van alle deelnemers. Veel tijdelijke banen zijn deeltijdse banen dus we vermoeden dat van de deelnemers die een tijdelijke baan vonden dat voor velen een deeltijdbaan was.

Tabel 3.2: Type contract van deelnemers die werk vonden sinds de vorige keer, meting 2-3, Wageningen

Type contract	Meting 2+3		
	N=	%	In % deelnemers
Type contract indien werk			
Vast contract (onbepaalde tijd)	26	19,1%	5,7%
Tijdelijke werk (contract voor bepaalde tijd)	85	62,5%	18,5%
Oproepkracht	5	3,7%	1,1%
Uitzendkracht	11	8,1%	2,4%
Zelfstandige/vrij beroep/freelancer	9	6,6%	2,0%
Aantal deelnemers meting 2+3=459	136	100%	29,6%

Bron: Vragenlijst eerste en tweede vervolgmeting deelnemers Wageningen, TiU, 2019

Op basis van deze gegevens is het aandeel berekend van de verschillende type banen (voltijd, deeltijd, vast, tijdelijk) in de vier Wageningse treatments (Tabel 3.3). Dan blijkt dat de onderzoeksgroep en de vrijlatingsgroep de meeste deeltijdbanen hebben en niet veel voor elkaar onderdoen. De vrijlatingsgroep en de onderzoeksgroep hebben echter ook de minste voltijdbanen. Het aandeel deeltijdbanen is het laagst in de eigen regiegroep (verschil met vrijlatingsgroep is significant) maar eigen regie heeft ook de meeste voltijdbanen in vergelijking met de controlegroep (verschil is significant).



Bron: Vragenlijst eerste en tweede vervolgmeting deelnemers Wageningen, TiU, 2019

Fig. 3.1: Aandeel type banen per treatmentgroep in het Vertrouwensexperiment Wageningen, 2017-2019

3.2 Resultaten: voltijdse uitstroom naar werk

Op basis van de BUS-gegevens is een analyse verricht op voltijdse (minimaal 27 uur) uitstroom naar werk en daarmee uitstroom uit de bijstand. Dit is gebeurd op basis van de registratie in de BUS van de reden van beëindiging van de uitkering welke wordt geregistreerd op basis van een gedegen beëindigingsonderzoek waarin wordt gecontroleerd wat de inkomsten zijn uit betaald werk. In hoeverre deze registratie een betrouwbare bron is om volledige uitstroom naar werk te achterhalen, is niet onderzocht. We nemen voorsnog aan dat de administratie van de gemeente betrouwbare informatie verschaft over uitstroom uit de uitkering.

Peildatum en mutaties in de BUS-registratie

De peildatum van uitstroom is 1 oktober 2019. Alle beëindigingen van de bijstandsuitkering volgens de Participatiewet vanwege uitstroom tot 1 oktober 2019 zijn in de analyses verwerkt. Dat betekent dat de observatie van de uitstroom naar werk een periode van twee jaar omvat. Gemeenten passen mutaties toe op de administratie tot 3 maanden na de eerste registratie. Dit heeft te maken met het herzien van het recht op uitkering, overgang naar andere typen bijstandsregelingen (bijvoorbeeld van alleenstaande naar samenwoning) en verandering in onder andere de persoonlijke inkomens -en vermogenssituatie van de bijstandsgerechtigde. Daarom zijn de uitstroomcijfers tot en met 31 december 2019 gebruikt om definitieve

cijfers per 1 oktober 2019 beschikbaar te hebben. De onderzoekers maken dus gebruik van de uitstroomcijfers in de BUS-data per 30 september 2019 maar op basis van de registratie op 31 december 2019⁶. Daarnaast is echter ook gekeken naar de ontwikkeling van de uitstroom in de drie maanden na het experiment dus van 1 oktober 2019 tot en met 31 december 2019. De deelnemers behouden ook na afloop in de regel dezelfde klantmanager waardoor naar verwachting de begeleiding door de klantmanager indien gedrag -of leereffecten zijn opgetreden niet ineens zal veranderen. Ook kan sprake zijn van na-ijlingseffecten, het gedrag van de deelnemer is veranderd maar het heeft tijd nodig voordat dit zich vertaalt in een 'uitstroom naar werk'- effect.

Volledige uitstroom naar werk

In Fig. 3.1 presenteren we de cijfers over de bereikte uitstroompercentages naar werk voor elk van de treatmentgroepen, de onderzoeksgroep en de referentiegroep per 1 oktober 2019 en per 31 december 2019.

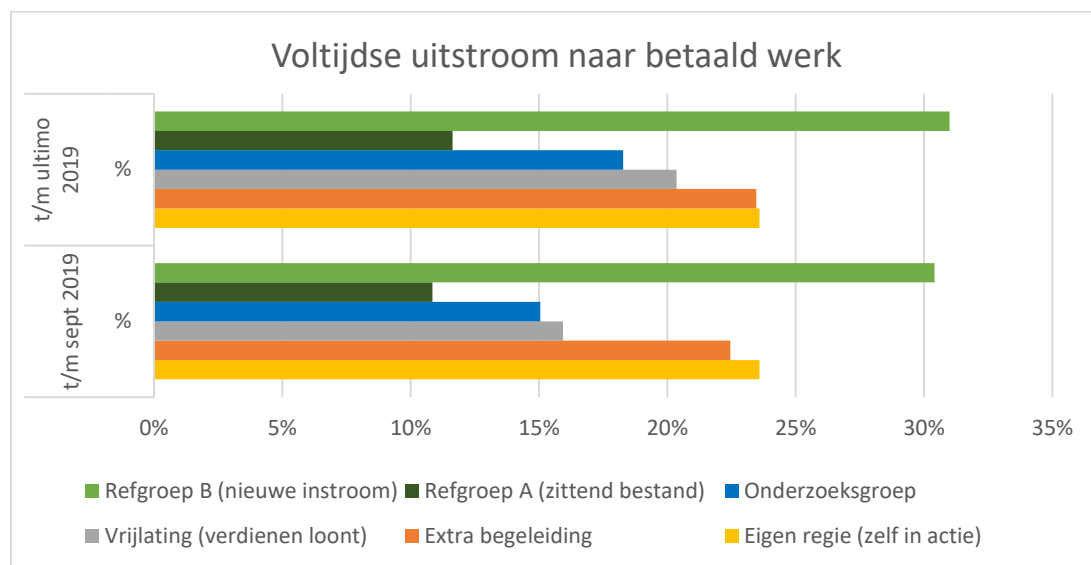


Fig. 3.1: Voltijdse uitstroom als percentage van deelnemersgroep aan het begin per treatment-groep en referentiegroep

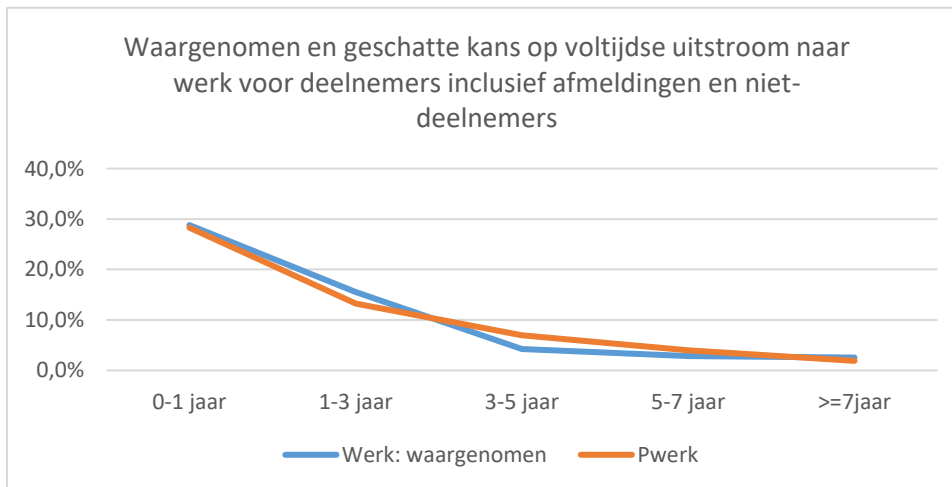
Hierbij is nog niet gecorrigeerd voor compositieverschillen, ofwel verschillen in samenstelling tussen de treatment groepen. De uitstroomcijfers van referentiegroep B (nieuwe instromers) springen eruit maar als we deze vergelijken met de nieuwe instromers in de treatment groepen dan zijn deze lager zoals Tabel 3.2 laat zien. Ten eerste is opvallend dat het percentage uitstroom naar werk in het zittend bestand bij de niet-deelnemers in de referentiegroep (Refgroep A) zo laag is over deze 2 jaar (11%) ofwel 5,5% voltijdse uitstroom per jaar. Dit geeft weer dan bij langere dueren het erg lastig is uit de bijstand te geraken door

⁶ De BUS cijfers die in het experiment zelf zijn gebruikt registreren een uitstroom van 84 voor de deelnemers exclusief afmeldingen terwijl wij op basis van de BUS-data per 1 oktober 2019 een uitstroom naar werk registreren van 79 voor alle deelnemers exclusief afmeldingen en 80 inclusief afmeldingen. Het verschil ontstaat doordat de uitstroomcijfers gebruikt in het experiment zijn gecorrigeerd voor de wijze van registratie van uitstroom naar de BBZ die niet als beëindigingsgrond is geregistreerd. Het betreft 5 mensen die uitstromen naar de BBZ.

voltijds betaald werk te vinden. In deze cijfers is uitstroom naar deeltijdwerk niet opgenomen. Daarover wordt in het eindverslag van de projectleider verder ingegaan.

Daarnaast zien we dat eigen regie en extra begeleiding relatief de meeste voltijdse uitstroom naar werk realiseren ten opzichte van onderzoek -of controlegroep. Ondanks het verschil in uitstroom van 7,5% en 6,3% tussen respectievelijk eigen regie en extra begeleiding met de controlegroep is alleen het verschil tussen de eigen regiegroep en de controlegroep significant. Tussen extra begeleiding en de controlegroep is het verschil net niet significant op 90% niveau van betrouwbaarheid ($p < 0.14$). Tussen 1 oktober 2019 en 31 december 2019 is de uitstroom naar voltijds werk vooral toegenomen bij de vrijlatingsgroep en de controlegroep en in iets mindere mate bij de groep extra begeleiding.

Eerder zagen we al dat duren van 1-3 jaar en langer zijn oververtegenwoordigd. Dit vertekent de uitkomsten aangezien de kans op uitstroom sterk samenhangt met de duur van de uitkering. Daardoor zullen de uitstroompercentages hoger zijn dan wanneer er geen vertekening naar duur was opgetreden. Uit de regressieanalyse die we later presenteren zal blijken dat deze vertekening niet erg groot is. In Fig. 3.2 geven we de waargenomen en de geschatte uitstroomkans weer voor de doelpopulatie. De doelpopulatie omvat alle mensen die voldoen aan de toelatingscriteria, op 1 oktober 2017 een lopende uitkering hadden of tussen 1 oktober 2017 en 1 juli 2018 in de PW uitkering instroomden ($N=968$). Hiertoe behoren dus ook de 410 deelnemers aan het experiment inclusief de 18 afmeldingen. De geschatte uitstroomkans geeft de uitstroomkans weer gecorrigeerd voor compositieverschillen tussen de vier treatment groepen (zie paragraaf 3.4. van dit verslag).



Figuur 3.2. Uitstroomkans naar werk naar duur uitkering 2017-2019, Wageningen, N=968 'Intention to treat' (ITT) analyse

De afmeldingen zijn meegenomen in de analyse omdat zij zich hebben aangemeld als deelnemer ook al hebben zij de treatment niet of slechts zeer gedeeltelijk ondergaan. Dit wordt in de experiment literatuur een ‘intention-to-treat’ analyse (ITT-analyse) genoemd. Iedereen die zich heeft aangemeld en tot de doelpopulatie behoort en daarna aselect is toegewezen (gerandomiseerd) aan een treatment groep wordt in principe in deze analyse gevolgd. In Fig. 3.2 zien we zowel bij de waargenomen als de geschatte uitstroomkans een sterke en tot aan een verblijf van 7 jaar of langer in de bijstand sterke daling van de uitstroomkans naar werk bij volledige uitstroom uit de uitkering. De gemiddelde duur in de uitkering bedraagt ruim 3,5 jaar in Wageningen voor deze deelpopulatie. De geschatte kansen wijken niet veel af van de waargenomen kansen. In Tabel 3.3 worden de ongecorrigeerde uitstroompercentages per treatmentgroep gegeven inclusief de 18 afmeldingen De resultaten gepresenteerd in Tabel 3.3 laten kleine verschillen zien voor de vergelijking met de onderzoeksgroep maar onverwachte en grotere verschillen in vergelijking met de referentiegroep. Zoals gezegd zijn deze laatste cijfers vertekend naar duur en nationaliteit en moet daarvoor modelmatig gecorrigeerd worden.

Zittend bestand versus nieuwe instroom

In tabel 3.3 maken we verder onderscheid tussen deelnemers die op 1 oktober 2017 tot het zittende bestand behoren en deelnemers die na 1 oktober 2017 tot 1 oktober 2018 zijn ingestroomd in het experiment, de zogenaamde nieuwe instromers. Aangezien in Wageningen vrij snel na de start het streefaantal van 300 deelnemers werd bereikt en zelfs meer dan 400 deelnemers zich al op 1 oktober 2017 hadden aangemeld is het aantal nieuwe instromers erg beperkt gebleven. In totaal zijn er 35 deelnemers ingestroomd na 1 oktober 2017 maar 17 deelnemers hadden al een lopende uitkering op 1 oktober 2017. Slechts 19 deelnemers zijn dus echte nieuwe instromers. In het vervolg van dit verslag zullen we geen onderscheid meer maken tussen beide groepen.

Tabel 3.3: Uitstroom naar werk voor deelnemers inclusief afmeldingen, naar zittend bestand en nieuwe instroom, treatmentgroep, op einddatum per 1 oktober 2019 en ultimo 2019 (ITT-analyse)

Strata	Zittend bestand		Nieuwe instroom		Totaal	
	01-Oct-19	31-Dec-19	01-Oct-19	31-Dec-19	01-Oct-19	31-Dec-19
wvexgroep5	VT-werk	VT-werk	VT-werk	VT-werk	VT-werk	VT-werk
Eigen regie (zelf in actie)	23,3%	23,3%	33,3%	33,3%	23,6%	23,6%
Extra begeleiding	21,7%	21,7%	33,3%	50,0%	22,5%	23,5%
Vrijlating (verdiene loont)	13,9%	18,5%	60,0%	60,0%	15,9%	20,4%
Onderzoek -controlegroep	14,8%	18,2%	20,0%	20,0%	15,1%	18,3%
Totaal %	18,4%	20,5%	36,8%	42,1%	19,5%	21,5%
N=	N=391	N=391	N=19	N=19	N=410	N=410
Refgroep A	10,9%	11,4%	n.v.t	n.v.t.	10,9%	11,4%
Refgroep B	n.v.t	n.v.t	31,0%	31,0%	31,0%	31,0%
Totaal niet-deelnemers	N=385	N=385	N=171	N=171	N=556	N=556

Bron: Vertrouwensexperiment Wageningen, Tilburg University, gegevens februari 2020.

Het aantal van 19 is te klein om verschillen in uitstroom zinvol te kunnen duiden maar het is duidelijk dat de nieuwe instromers in alle treatment groepen aanmerkelijk meer kans hebben voltijds aan het werk te gaan (33-60%) dan in de controlegroep (22% kans) of de referentiegroepen (31% kans).

Wat het zittende bestand betreft is het verschil tussen de treatments en de controlegroep alleen significant voor de eigen regiegroep maar niet voor de extra begeleidingsgroep of de vrijlatingsgroep. Op peildatum 1 oktober 2019 realiseren alle groepen inclusief de controlegroep een hogere uitstroom dan het geval is in de referentiegroep die evenals de controlegroep de reguliere treatment heeft gekregen. De standaard treatment in de controlegroep is kennelijk niet goed vergelijkbaar geweest met de standaard treatment van de mensen in de referentiegroep. Dit zou kunnen wijzen op het eerder genoemde "Hawthorne" effect dat de standaard treatment in de controlegroep onder invloed van het experiment en het onderzoek is veranderd. Zoals gezegd zijn de uitstroomcijfers nochtans overschat door de ondervetegenwoordiging van langere duren. Hieronder bespreken we de methode en de uitkomsten daarvan die is gebruikt om daarvoor te corrigeren.

Modelmatige schatting van de uitstroomkans

Om te corrigeren voor compositieverschillen tussen de groepen ook wat duur betreft hebben de onderzoekers bij de analyse van voltijdse uitstroom naar betaald werk gebruik gemaakt van een analysemodel dat zoveel mogelijk vergelijkbaar is gemaakt met het model dat in het landelijk overleg van onderzoekers is overeengekomen. Dat LOEP-model zal later worden gebruikt voor de analyses op de CBS-data. Het LOEP-model is afgestemd op de beschikbare informatie in de CBS-data, niet alleen de BUS-data waarvan het CBS ook gebruik maakt, maar ook andere registratiedata die het CBS gebruikt. In het model dat wij gebruiken voor Wageningen zijn zoveel mogelijk dezelfde variabelen gebruikt als in het LOEP-model. Bij bepaalde variabelen bleek dat moeilijk te realiseren zoals bij opleiding aangezien de gemeenten hoogst behaalde opleiding bijvoorbeeld op een eigen manier registreren welke niet volledig vergelijkbaar is met die van het CBS op basis van andere bronnen. Voor Wageningen ontbraken betrouwbare gegevens over opleiding in de BUS-data. In Tabel 3.4 worden de in het model opgenomen variabelen weergegeven.

De afhankelijke variabele is voltijdse uitstroom naar betaald werk tot 1 oktober 2019. De onafhankelijke variabelen zijn gemeten in beginsel een maand voor de start van iedere deelnemer aan het experiment tenzij voor die maand geen informatie in de BUS-data beschikbaar was. Dan is de dichtstbijzijnde maand genomen waarvoor wel informatie beschikbaar was. De gemiddelde duur bedraagt bij aanvang van het experiment meer dan 3 jaar en 6 maanden. Voor de niet-deelnemers is de informatie genomen van de maand september 2017.

Tabel 3.4: Aantal, gemiddelde en standaardafwijking gebruikte variabelen in het analysemodel volledige uitstroom naar werk, Wageningen, 2017-2019

Naam model variabele	Omschrijving	N=	Gemiddelde	sd
Uitstroom werk	Reden van uitstroom	966	0.18	0.38
wvexgroep6	Onderzoeksgroep (treatments, onderzoeksgroep, referentiegroep A/B)	966	4.10	1.62
duuruitk_t0	Duur uitkering PW bij start experiment ¹⁾	966	4.35	5.67
nkeeruitk_t0	Aantal keer uitkering voor start experiment	966	1.55	1.60
lft_t0	Leeftijd bij start	966	43	13.7
geslacht_t0	Geslacht (vrouw=1; man=0)	966	0.53	0.50
buitnat_t0	Nationaliteit (Nederlands of buitenlands)	966	0.17	0.38
leefv_t0	Leefvorm (alleenstaande, alleenstaande ouder, samenwonend)	966	2.00	0.57
huisv_t0	Huisvesting (eigenaar, huurder, inwonend, overig)	966	1.37	0.79
nizb	Nieuwe instroom of zittend bestand ¹⁾	966	0.83	0.38

Noot: sd=standaarddeviatie. ¹⁾ Nagenoeg alle deelnemers in de treatmentgroepen komen uit het zittende bestand (n=391). In de referentiegroep is dat 70% (n=385).

Bron: BUS-data gemeente Wageningen, 2017-2019

Resultaten

Tabel 3.5 geeft de geschatte uitstroomkansen en de verschillen of effecten weer. We presenteren hier de resultaten voor het zittende bestand (n=391) maar het model is ook geschat op alle deelnemers inclusief de nieuwe instromers (n=19). De resultaten zijn nagenoeg hetzelfde. Het model voor zittend bestand verklaart 36% van de variantie (pseudo R²) en het model inclusief nieuwe instromers 32%. Het model corrigeert zoals gezegd voor duurverschillen tussen de groepen maar ook voor andere compositie verschillen die te maken hebben met geslacht, leeftijd, nationaliteit, duur uitkering, aantal keer in bijstand voorafgaand aan start, leefvorm, huisvesting situatie bij de start. We hebben het model geschat op het laatste record van de deelnemers en niet-deelnemers. Dat zijn 1718 records waarvan 968 tot de doelpopulatie behoorden en 776 tot het zittende bestand per 1 oktober 2017. Dat betekent dat het laatste record van de mensen waarvan de uitkering voor 1 oktober 2019 is beëindigd betrekking heeft op de maand waarin de uitkering is beëindigd. Het bestand is dus een gestapeld of gepoold bestand van deelnemers en niet-deelnemers over de maanden vanaf 1 oktober 2017. Om die reden is tevens een tijdvariabele toegevoegd die de maand weergeeft. Op deze manier corrigeren we tevens voor de verandering in arbeidsmarktsituatie over de tijd voor zover deze verschillend uitwerkt op de treatments en controlegroep of referentiegroep. We hebben twee modellen geschat, een model met de controlegroep als vergelijkingsgroep en een model met de referentiegroep als vergelijkingsgroep.

Tabel 3.5. Resultaten logistische regressie: geschatte uitstroomkansen zittend bestand naar treatment groep, inclusief en exclusief afmeldingen, plus treatment-effecten in procentpunten

Totaal	Inclusief afmeldingen (ITT)		Exclusief afmeldingen		Effect t.o.v. controle groep	Effect t.o.v. referentie groep
	N=	P-werk	N=	P-werk	P-werk	P-werk
Eigen regie	103	23.0%	95	23.3%	7.6%	12.9%**
Extra begeleiding	92	21.5%	88	22.2%	6.1%	11.4% ¹⁾
Vrijlating	108	14.7%	104	15.2%	-0.7%	4.6%
Controlegroep	88	15.4%	86	15.6%	0.0%	5.3%
Referentie groep A	378	10.1%	378	10.1%	-5.3%	0.0%
Treatments	303	19.6%	287	20.0%	4.2%*	9.5%***
Deelnemers	391	18.6%	373	19.0%	nvt	8.5%***
N=953						
Pseudo R ²	36,5%					
Wald chi2(22)	134,4					
Prob > chi2	0.000					
<i>Contrast-analyse</i>						
ER-EB					ns	ns
ER-VRL of EB-VRL					ns	ns
ER/EB					*	***
ER/EB/VRL					ns	***

Noot: ***p<0,01; **p<0,05; *p<0,10; ¹⁾significant op 85% niveau: p<0.15. Logit regressiemodel met controlevariabelen leeftijd, nationaliteit, geslacht, log duur uitkering bij start, leefvorm bij start, huisvesting bij start, maand. ER=eigen regie; EB=extra begeleiding; VRL=Vrijlating; ER-EB=contrast tussen eigen regie en extra begeleiding; ER-VRL of EB-VRL=contrast tussen eigen regie of extra begeleiding en vrijlating; ER/EB/VRL=ER, EB en VRL samen; ER/EB=ER en EB samen.

Bron: Bewerking BUS data Wageningen, Tilburg University, 2019.

De uitstroomkansen zijn in beide modellen hetzelfde maar de effecten zijn verschillend. Het effect van de treatments wordt in de tabel weergegeven door het verschil in de gemiddelde uitstroomkansen te berekenen ten opzichte van respectievelijk de controlegroep en de referentiegroep. Het model met de controlegroep als vergelijkingsgroep laat geen significante of betekenisvolle verschillen zien tussen de afzonderlijke treatment groepen en de controlegroep op deze uitkomstmaat. Gezien het verschil in percentage van 6-7% is dit opvallend. Dit kan duiden op een power probleem vanwege de kleine aantallen. Het effect van beide treatments tezamen (ER/EB) met een verschil van 6,7% vergeleken met de controlegroep is daarentegen wel significant op basis van een ttest maar ook op basis van hetzelfde

logistische regressiemodel maar met beide groepen tezamen genomen in de treatment variabele. Datzelfde geldt echter niet voor het verschil tussen alle drie treatments tezamen en de controlegroep (4,2%)⁷.

Het model met de referentiegroep als vergelijkingsgroep laat significante verschillen zien tussen eigen regie en de referentiegroep. Het verschil van extra begeleiding met de referentiegroep is in het model voor het zittend bestand net niet significant op 90% betrouwbaarheidsniveau ($p < 0.15$) maar wel significant in het model met referentiegroep A en B samen als vergelijkingsgroep ($p < 0.10$). Het verschil in percentage tussen eigen regie en extra begeleiding en de referentiegroep is met 11-13% ook hoger dan in vergelijking met de controlegroep. De referentiegroep is echter niet gerandomiseerd waardoor we niet weten of het verschil causaal geïdentificeerd kan worden. Door selectie bij aanmelding is de deelnemersgroep minder representatief voor de doelpopulatie maar door opname van de variabelen duur bijstand bij start experiment, geslacht, leeftijd en dergelijke in het model wordt daarvoor zo goed mogelijk gecorrigeerd.

Contrasten

Voor beide regressiemodellen is een contrast analyse uitgevoerd op de treatment variabele om te bezien of er significante of betekenisvolle verschillen tussen de vier treatments. In het model met de controlegroep als ijkpunt vinden we geen significante treatmenteffecten tussen elk van de treatments en de controlegroep maar wel als we de beide treatments eigen regie en extra begeleiding samen analyseren. In het model met de referentiegroep als vergelijkingsgroep heeft eigen regie een significant effect en is ook de combinatie van de beide treatments significant maar is er geen significant verschil tussen elk van de treatments afzonderlijk.

Marginale effecten

Tevens hebben we de marginale effecten berekend (dy/dx =toename in uitstroomkans door deelname aan de treatment in vergelijking met de reguliere dienstverlening onder constant houden van de andere variabelen). Dit hebben we alleen gedaan voor de vergelijking met de referentiegroep omdat de effecten in het regressiemodel vergeleken met de controlegroep niet significant en dus statistisch nul zijn. In Tabel 3.7 worden ook de standaardfouten en bijbehorende betrouwbaarheidsintervallen weergegeven. Door de grootte van de standaardfout zijn de betrouwbaarheidsintervallen breed. De tabel bevestigt de eerdere conclusies dat eigen regie een positief significant effect heeft op uitstroom naar werk in vergelijking met de referentiegroep. Het effect van extra begeleiding is positief en net niet significant op 90% niveau maar wel op 87% niveau ($p < 0.13$). Een marginaal effect van 5,6% en 4,8% op de uitstroomkans ten opzichte van de referentiegroep (gemiddelde uitstroomkans 10,9%) is groot te noemen.

⁷ In de ITT-analyse met de controlegroep (dus inclusief afmeldingen) zijn de gemiddelde treatment effecten berekend voor de treated maar niet de LATE effecten ('local average treatment effecten'). De LATE effecten zijn wel berekend op de CBS data door LOEP en het CPB maar alleen indien de ITT effecten significant zijn. Voor Wageningen zijn de door LOEP en het CPB berekende ITT-treatment effecten weliswaar positief, maar niet significant. Zij bevestigen het beeld hier geschetst op de BUS-data.

Tabel 3.7. Marginale effecten op uitstroom naar voltijds werk voor mensen in zittend bestand plus betrouwbaarheidsintervallen voor vergelijking met referentiegroep,

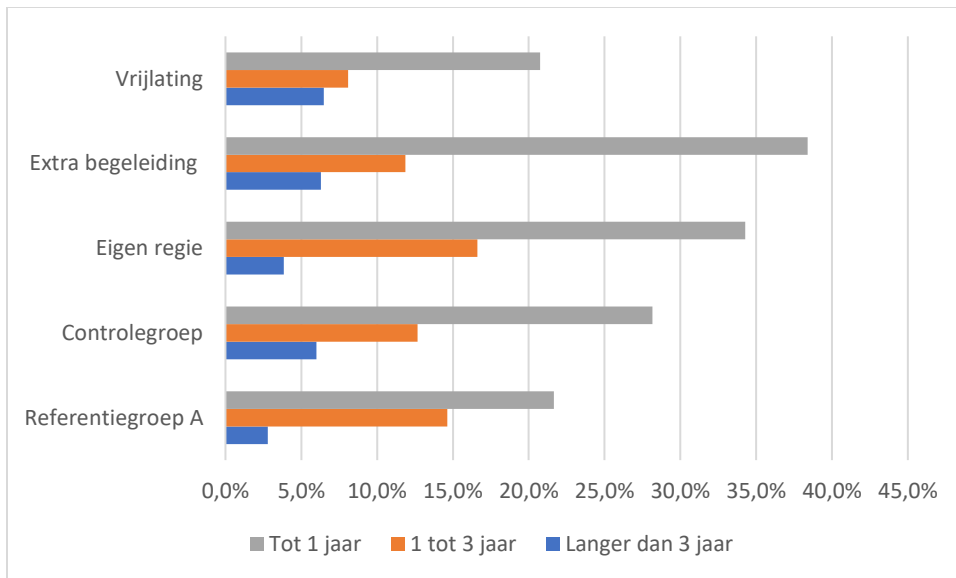
Referentiegroep A/B	Marg. Effecten (dy/dx)	Std. Fout	P>z	Betrouwbaarheidsinterval (95%)	
Treatments					
Eigen regie (zelf in actie)	0.056	0.029	0.055*	-0.001	0.114
Extra begeleiding (maatwerk)	0.048	0.031	0.127 ¹⁾	-0.014	0.110
Vrijlating (verdiene loont)	0.001	0.029	0.967	-0.056	0.058
Onderzoekgroep (controlegroep)	0.051	0.036	0.158	-0.020	0.122

***p<0,01; **p<0,05; *p<0,10

Bron: Bewerking BUS data, Tilburg University, 2019.

3.3 De invloed van duur nader onderzocht: matching modellen

De vraag stelt zich nochtans of het lineaire kansmodel (logit) in staat is om op correcte wijze te corrigeren voor de verschillen in duren. Het effect van duur in het model met de referentiegroep als vergelijkingsgroep was sterk negatief en significant op 95% niveau. Indien we de geschatte kansen vergelijken naar duur en treatmentgroep (Fig. 3.3) dan blijven de verschillen tussen de beide treatment groepen eigen regie en extra begeleiding en de referentiegroep overeind, inclusief de verschillen tussen de controlegroep en de referentiegroep. Dat geldt voor de korte duren (tot 1 jaar) en de langere duren (>3 jaar) maar niet voor de duren van 1-3 jaar, althans bij extra begeleiding. Voegen we in het model een interactieterm toe tussen treatment en duurklasse (drie of vijf klassen), dan worden de hoofdeffecten van de treatments sterker (interacties zijn niet significant) in plaats van zwakker. Voor beide treatments samen is het berekende marginale effect (dy/dx) ten opzichte van de referentiegroep (de extra kans op werk voor niet-deelnemers indien zij de treatment zouden krijgen) gemiddeld 4,7%. Dat betekent dat de duurvariabele een deel van het treatment effect wegneemt en dat de causaliteit deels verloopt via de treatment; de treatment vergroot de uitstroomkans van mensen met een langere duur bij aanvang van het experiment.



Bron: Bewerking BUS data, Tilburg University, 2019.

Fig. 3.3: Geschatte uitstroomkansen naar voltijds werk naar duur en treatment groep, Wageningen, zittend bestand, 1 okt2017- 1okt2019

Rol van klantmanager

Naast de duur spelen andere factoren een rol die het verschil zouden kunnen verklaren. Een van die factoren is wellicht de klantmanager zelf. Het eerdere logit model met de controlegroep als vergelijkingsgroep en klantmanager nummer als verklarende variabele verklaart de verschillen in uitstroomkansen naar werk tussen de groepen echter slechts 7% beter als hetzelfde model zonder klantmanager nummer (33% in plaats van 26%). Dus de rol van de klantmanager is niet de enige verklaring en er zijn nog andere factoren die de verschillen verklaren.

Matching: correctie voor compositieverschillen

Als duur enerzijds corrigeert voor verschillen tussen de treatments groepen en de referentiegroep maar anderzijds deels ook het effect van treatment verklaart dan is niet duidelijk wat het zuivere effect van treatment is. Een tweede manier om daar achter te komen is het gebruik van 'matching' om te corrigeren voor verschillen tussen de groepen. Wat dan overblijft is wellicht het 'zuivere' treatment effect. Wellicht, omdat niet uitgesloten kan worden dat niet waargenomen verschillen tussen de groepen onderliggend het effect mee beïnvloeden. Een van die variabelen is opleidingsniveau waarvoor in de BUS data geen informatie voorhanden is maar het kunnen ook gezondheidsverschillen zijn of verschillen in motivatie en zoekgedrag. Voor deze niet-waargenomen verschillen zou kunnen worden gecorrigeerd met behulp van 'random' en 'fixed effects' panel regressie modellen. Vooralsnog is dit (nog) niet gebeurd en ook niet voorzien in het onderzoek maar zal eventueel later alsnog middels de koppeling aan de microdata van het CBS ter hand kunnen worden genomen.

Indien de populatie van deelnemers door (zelf)selectie bij de aanmelding niet representatief is op bepaalde aspecten, in dit geval vooral duur, maar ook enigszins op nationaliteit, geslacht en leeftijd, dan kan dus deels met 'matching technieken' hiervoor gecorrigeerd worden. Met matching wordt gezocht naar zoveel mogelijk vergelijkbare mensen aan de deelnemers in de groep van niet-deelnemers die de treatment niet

hebben gehad. Op deze manier kunnen we de resultaten bij de deelnemers vergelijken met een gematchte groep van mensen die qua duur en andere kenmerken zoveel mogelijk op de deelnemers lijken. Hiermee corrigeer je voor verschillen in waargenomen kenmerken zoals duur, leeftijd, geslacht, leefvorm, huisvestingssituatie en nationaliteit. De analyse is opnieuw alleen uitgevoerd op de deelnemers in het zittende bestand (n=391) dus met een lopende uitkering per 1 oktober 2017. We hebben ‘een-op-een’ matching toegepast, dus voor iedere deelnemer is één andere persoon gezocht in de referentiegroep die op de deelnemer lijkt. Indien meer personen vergelijkbaar zijn wordt degene gekozen die het meest lijkt op de deelnemer. Dat wordt de ‘dichtstbijzijnde buur’ (‘nearest neighbour’) genoemd⁸. Voor de matching zijn zoveel mogelijk dezelfde variabelen gebruikt als in het regressiemodel. In beide methoden leverde een indeling in vijf duurklassen betere ‘matching’ resultaten op. De beide methoden laten toe om het effect van de interventie bij de treated nauwkeuriger vast te stellen (ATT-average treatment effect van de treated) maar ook het gemiddelde effect van de treatment voor de gehele populatie van deelnemers en niet-deelnemers. Voor de niet-deelnemers wordt dan op basis van hun kenmerken berekend wat het effect zou zijn geweest als zij ook de treatment hadden ondergaan (ATE-‘average treatment effect’). Tot slot is het mogelijk om het effect te bepalen voor degenen die geen interventie hebben gehad ervan uitgaande dat zij die wel zouden hebben gehad (ATU-‘average treatment effect van de untreated’). In onderzoek wordt veelal de ATT en ATE gebruikt om het effect van een interventie te meten. De ATT en ATE geven dan het verschil in geschatte uitstroomkans weer tussen de beide treatmentgroepen en de vergelijkingsgroep. Omdat we geen verschil vonden van de twee treatmentgroepen met de controlegroep wordt hier alleen ingegaan op de analyses gepresenteerd met de referentiegroep als vergelijkingsgroep⁹.

Correctie voor duurverschillen

Voordat we in Tabel 3.7 de resultaten presenteren laten we zien dat het model inderdaad in staat is geweest om te corrigeren voor verschillen in duur. In Fig. 3.4 (links nn-matching; rechts ps-matching) staan in beide plots links de feitelijke uitstroomkansen naar werk en rechts de gecorrigeerde of geschatte uitstroomkansen voor de referentiegroep en de beide treatments dus na matching.

⁸ De matching (zoeken en bepalen van vergelijkbare mensen) kan op basis van een wegingsfunctie van de verklarende variabelen (bijvoorbeeld middels nearest neighbour matching) of op basis van zogenaamde ‘propensity scores’ (de kans op treatment bepaald met behulp van een logit of probit model met verklarende variabelen) van de covariaten. Beide methoden zijn toegepast met vergelijkbare resultaten.

⁹ De modellen zijn ook geschat met de onderzoek -of controlegroep als vergelijkingsgroep maar de resultaten bevestigen de uitkomsten van de logistische regressies; er werden geen significante verschillen gevonden tussen beide treatmentgroepen en de controlegroep.

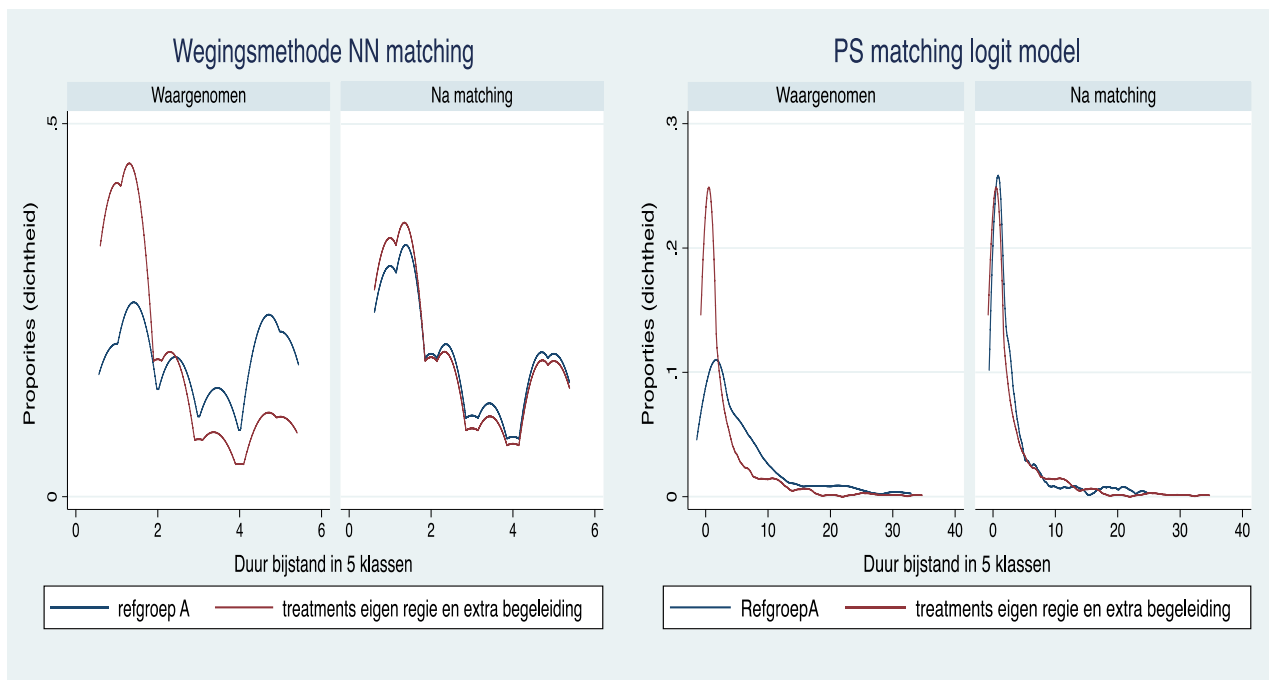


Fig. 3.4 Verschillen tussen controlegroep en treatments naar duur uitkering van zittend bestand en matchingmethode (op basis van schatting 'average treatment effect' (ATE) over deelnemers en niet-deelnemers)

Tabel 3.7 geeft de effecten weer voor de twee treatment groepen in vergelijking met de referentiegroep en de onderzoeksgroep. Aangezien we eerder vonden dat er geen verschillen zijn tussen de groepen presenteren we de effecten voor beide groepen (eigen regie en extra begeleiding) samen. Bevestigd wordt in deze cijfers dat in vergelijking met de controlegroep de effecten van de beide treatments op de uitstroomkans niet significant of betekenisvol zijn en daardoor statistisch nul. In vergelijking met de referentiegroep zijn ze echter -ook na correctie op basis van matching- wel betekenisvol maar gehalveerd (ATE=5-6%). De resultaten van de analyse worden in Tabel 3.7 weergegeven.

Ten eerste wordt bevestigd dat de uitstroom naar werk effecten niet significant zijn voor de vergelijking met de onderzoeksgroep of controlegroep. In vergelijking met de referentiegroep vinden we wel significante effecten. Vergelijken we de twee treatments met de referentiegroep dan bedraagt de ATT 6,4%, en de ATE 4,8% in de wegingsmethode. In de 'propensity score' matching methode bedraagt de ATT 8,7% en de ATE 5,6%¹⁰. Omdat de ATE het gemiddelde effect berekent van zowel de treated als de untreated geeft deze een betere schatting van het in beginsel onbekende treatment effect voor de gehele populatie (we weten immers niet wat het treatment effect zou zijn geweest bij degenen die geen deelnemer zijn indien zij de treatment zouden hebben gehad).

¹⁰ De modelschattingen op basis van de wegingsmethode zijn minder gevoelig voor modelspecificatie dan de schattingen op basis van de propensity score methode waardoor we de eerstgenoemde methode betrouwbaarder achten.

Tabel 3.7. Resultaten matching modellen uitstroomkans werk (voltijds), deelnemers zittend bestand

	Wegingsmodel (NN-matching) ¹⁾		PS-matching ¹⁾		N=
	ATT	ATE	ATT	ATE	
Vershil t.o.v. referentiegroep A					
Treatments (eigen regie, extra begeleiding)	6,4%*	4,9%*	8,7%**	5,6%**	N=681
Vershil t.o.v. onderzoekgroep					
Treatments (eigen regie, extra begeleiding)	3,6%	1,8%	5,2%	3,7%	N=391
Model variabelen					
Duur uitkering t0 in 5 klassen (<1 jr.; 1-3jr;3-5jr;5-7jr;>7jr)	x	x	x	x	
Leeftijd op t0	x	x	x	x	
Leeftijd t0 kwadraat	x	x	x	x	
Geslacht (man, vrouw)	x	x	x	x	
Buitenlandse nationaliteit (Nederlands, buitenlands)	x	x	x	x	
Leefvorm (alleenstaand, samenwonend, overig)	x	x	x	x	
Huisvesting (eigenaar, huurder, inwonend)	x	x			
Bias (duur uitkering, nationaliteit, leeftijd)	x	x			

Noten: nn-matching='nearest neighbour matching'; ps-matching=propensity score matching.

***p<0,01; **p<0,05; *p<0,10, getest op basis van een nearest neighbour matching model en een propensity score matching logit model (N=681); ¹⁾ In het model voor de 'nearest neighbourhood' matching methode is een correctie opgenomen voor bias (selectie) op duur uitkering, nationaliteit, leeftijd en leeftijd kwadraat. Beide modellen zijn geschat met Stata 15.1.

Bron: Bewerking BUS data Wageningen, Tilburg University, 2019.

Dat betekent dat als we uitgaan van de ATE er een 4,8% tot 5,6% hogere kans bij de treatmentgroepen is om volledig uit te stromen in vergelijking met de referentiegroep. Dat is lager dan we vonden in de ruwe data of bij de logit regressie resultaten (11-13%) maar niettemin een substantieel effect. Door via matching te corrigeren voor verschillen in samenstelling tussen de treatments en de vergelijkingsgroep van niet-deelnemers is zoals verwacht het uitstroom naar werk effect ongeveer gehalveerd, maar is het wel significant en nog steeds substantieel.

3.4 Conclusies uitstroom naar werk

De waargenomen uitstroomkansen naar werk (voltijds) zijn voor het zittende bestand (N=391 deelnemers) 4-5% hoger voor de treatment groepen eigen regie en extra begeleiding in vergelijking met de controlegroep of onderzoekgroep. Voor de vrijlatingsgroep is er geen of zelfs een klein negatief verschil met de onderzoek -of controlegroep. Nadere analyse van de uitstroom tot aan 31 december 2019 wijst evenwel uit dat met name in de vrijlatingsgroep maar ook in de controlegroep na 1 oktober 2019 (het einde

van het experiment) relatief veel mensen zijn uitgestroomd. De resultaten voor de vrijlatingsgroep zijn over deze langere periode dan gunstiger. In deze cijfers is uitstroom naar deeltijdwerk nog niet meegenomen. Uit de analyse van de BUS cijfers over de inkomsten verkregen uit werk blijkt dat met name in deze vrijlatingsgroep de uitstroom naar deeltijdwerk relatief groot is (zie eindverslag projectleider). De geschatte en voor een aantal achtergrondvariabelen (leeftijd, geslacht, duur uitkering bij start, nationaliteit, aantal keer bijstand bij start) gecorrigeerde kansen liggen dicht bij de waargenomen kansen. Dit geeft aan dat de vertekening die mogelijk is opgetreden erg klein is. Kennelijk is de randomisatie goed gelukt. Nochtans vinden we geen significante verschillen in de ITT-analyse in vergelijking met de controlegroep (de analyse inclusief de afmeldingen). Dat betekent dat er geen betekenisvolle verschillen zijn tussen elk van de drie treatmentgroepen en de controlegroep en dat het effect van elk van de treatments ten opzichte van de controlegroep dus nul is. Dit is opvallend en kan te maken hebben met een power probleem; de aantallen per groep zijn dermate klein dat statistisch gezien significantie op 95% en zelfs 90% betrouwbaarheid niet gerealiseerd wordt waardoor effecten net niet significant zijn op 90% niveau maar wel op 85% of 89% niveau. Door groepen samen te nemen kan dat wellicht wel. Vergelijken we de gerealiseerde uitstroomkansen van eigen regie en extra begeleiding samen met de controlegroep dan is er wel een klein significant effect op basis van een ttest.

Vergelijking met referentiegroep A (zittend bestand)

Vergelijken we de controlegroep met de referentiegroep van niet-deelnemers in het zittende bestand (refgroep A) dan valt op dat de controlegroep hogere uitstroomcijfers laat zien dan de referentiegroep. Dat geldt voor de waargenomen maar ook de geschatte uitstroomcijfers op basis van het regressiemodel. Gemiddeld genomen zijn de ruwe of waargenomen uitstroomkansen voor de beide treatmentgroepen 22-23%, hetgeen 11-12% hoger is dan voor de groep van niet-deelnemers in de referentiegroep (11%). Dat is een fors verschil. De treatment groep eigen regie laat nu wel significante resultaten zien (op 90% betrouwbaarheidsniveau). Voor de groep extra begeleiding is het effect net niet significant op 90% niveau maar wel op 89% niveau. Het effect voor de vrijlatingsgroep is niet significant en bijna nul.

Matching

De aangemelde deelnemers zijn aselekt toegewezen aan de verschillende treatments maar door zelfselectie kan de groep van deelnemers qua samenstelling afwijken van de populatie waartoe ze behoorde. Dat blijkt ook het geval te zijn. Vooral naar duur uitkering en nationaliteit, en enigszins naar geslacht en leeftijd zijn er verschillen tussen de deelnemers en de populatie van niet-deelnemers ontstaan waarvoor gecorrigeerd moet worden. Om dat te doen is gebruik gemaakt van twee matching methoden een matching wegingsmethode (nn-matching) en een model matching methode (ps-matching). Met beide methoden bleek het mogelijk voldoende geschikte matches te vinden voor de deelnemers onder de niet-deelnemers (100%). Op basis daarvan berekenden we dat uitgaande van de ATE er een gemiddeld treatment effect van 4,8% (nn-matching) tot 5,6% (ps-matching) bestaat voor de beide treatment groepen tezamen op basis van de beide matching methoden. De treatment in de controlegroep is in het proces van experimenteren mogelijk wat meer gaan lijken op de beide andere treatments. De uitstroompercentages zijn in ieder geval hoger dan in de referentiegroep die ook de reguliere dienstverlening heeft gekregen. De wijze van ondersteuning door de klantmanager lijkt van belang maar minder dan verwacht. Indien we voor de 391 deelnemers in het zittende bestand waarvoor we betrouwbare informatie hebben (inclusief dus de

afmeldingen) een model schatten om de verschillen tussen de groepen te verklaren in uitstroomkansen dan blijkt dat 7% van de verklaarde verschillen (in totaal 32%) te zijn verklaard door de persoon van de consulent. Dat is niet veel waardoor er andere factoren moeten zijn die het verschil mede verklaren.

Uitstroom naar deeltijdwerk

Gegevens over deeltijdwerk en flexwerk zijn niet beschikbaar in de lokale BUS-data. Wel zijn in de tweede en derde meting bij de deelnemers vragen gesteld over deeltijdwerk en flexwerk. Dit geeft vanwege de response van 60% echter geen volledig beeld. Hieruit bleek dat de meeste deeltijdbanen werden gerealiseerd in de vrijlatingsgroep en de onderzoeksgroep en de minste in de eigen regiegroep (verschil is significant $p < 0,02$). De meeste voltijdbanen werden gerealiseerd in de eigen regiegroep (verschil met onderzoeksgroep is significant, $p < 0,06$) en de minste in de vrijlatingsgroep en onderzoeksgroep. In hoeverre er substitutie heeft plaatsgevonden tussen deeltijdwerk en voltijdwerk in de vrijlatingsgroep is evenwel niet duidelijk. Omdat in de vrijlatingsgroep afgezien van vrijlating de reguliere dienstverlening werd geboden lijkt vrijlating een positief effect te hebben gehad op uitstroom naar deeltijdwerk.

3.5 Secundaire uitkomstmaten

In de drie vragenlijsten voor de deelnemers (nulmeting, meting 2 en meting 3) zijn vragen gesteld die beogen het experiment te evalueren op de uitkomstmaten die niet direct gerelateerd zijn aan betaald werk zoals welbevinden, subjectieve gezondheid (fysieke en mentale gezondheid), sociale participatie, sociaal en institutioneel vertrouwen, zelfredzaamheid, keuzemogelijkheden, financiële situatie, materiële deprivatie en armoede. In overleg met de andere onderzoekers in LOEP zijn uit een mogelijke lijst van indicatoren waarover informatie in de vragenlijsten beschikbaar was, er negen tot 10 geselecteerd die als leidraad zullen dienen bij de rapportage van de AMvB experimenten. Dat zijn welbevinden, subjectieve gezondheid, sociale participatie, ervaren keuzemogelijkheden (capabilities), zoekintensiteit, zelfredzaamheid, vertrouwen in de ander en in instituties, financiële stress en armoede en/of deprivatie. Deze zullen ook in dit rapport worden gepresenteerd. In de beide tussenrapportages voor Wageningen is over een aantal indicatoren reeds verslag gedaan maar is nog geen informatie gegeven over de vergelijking naar treatmentgroep. Deze indicatoren passeren achtereenvolgens de revue hier.

Analysemethode

We hebben nu voor elk van deze indicatoren een regressiemodel (OLS) geschat waarmee we de kans op de uitkomst bepalen waarbij we controleren voor compositieverschillen mogelijk veroorzaakt door verschillen in duur uitkering bij de start, leeftijd, geslacht, aantal keer uitkering bij start, opleiding, nationaliteit, klantmanager en of men behoort tot het zittende bestand op 1 oktober 2017 ($n=391$) of tot de nieuwe instroom ($n=19$) die na 1 oktober 2017 is ingestroomd. Voor de meeste variabelen is gebruik gemaakt van de BUS informatie. Over opleiding was er geen BUS informatie beschikbaar en hebben we deze afgeleid uit de informatie over de hoogst behaalde opleiding in de nulmeting. Voor mensen waarvoor de opleiding niet bekend is, is een categorie onbekend toegevoegd. De modelmatig gecorrigeerde of geschatte cijfers zullen we hierna presenteren en dus niet de feitelijk waargenomen cijfers. Het effect van de correctie is overigens erg klein omdat door de randomisering er kleine verschillen zijn tussen de vier groepen. In de rapportage van de analyse, wordt gebruik gemaakt van een beschrijvende tabel waarin de gemiddelde gecorrigeerde

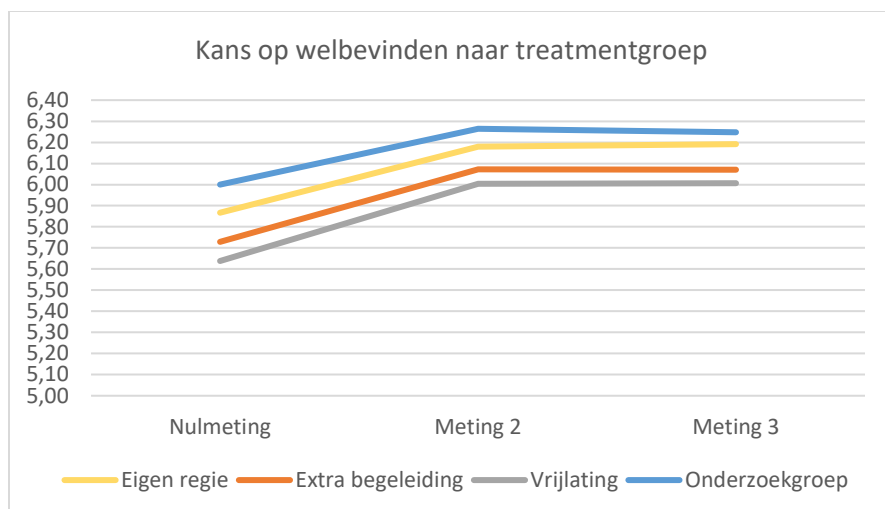
scores voor de treatment groepen op een uitkomstvariabele worden getoond per meting en in totaal. Deze gemiddelden zijn geschat op basis van de OLS regressieanalyse. Een nadere uitleg over hoe de tabellen gelezen moeten worden staat in paragraaf 4.1. Of een effect betekenisvol is of niet en met welke mate van betrouwbaarheid wordt aangegeven met sterretjes; drie sterretjes staat voor 99% betrouwbaarheid twee sterretjes voor 95% en een sterretje voor 90% betrouwbaarheid. Er zijn steeds twee modellen geschat een zonder interactie-effect en een met interactie-effecten tussen treatment en tijd (de drie vragenlijst metingen) (zie Bijlage A voor de regressieresultaten). De effecten in de randtotalen geven op basis van het model zonder interacties de effecten weer van verschillen over de tijd (rijtotalen onder) of van verschillen tussen elk van beide treatments en de onderzoeksgroep gemiddeld over de tijd (kolomtotalen rechts). We zijn met name hierin geïnteresseerd (zie ook Tabel A1.1 in Bijlage A). De sterretjes bij de scores in de andere cellen geven op basis van de schattingen in het interactiemodel weer in hoeverre er een betekenisvol effect is van de treatment ten opzichte van de onderzoeksgroep in de nulmeting bovenop het effect dus dat er was aan het begin van het experiment (zie Tabel A1.2 in Bijlage A). Aangezien aan het begin de treatment de meeste effecten naar verwachting nul zijn geven deze veelal de volledige treatment effecten weer. In het model zijn naast de treatment variabele controlevariabelen opgenomen maar tevens een variabele die de persoon van de klantmanager weergeeft.

3.5.1 Gezondheid en welbevinden

In de vragenlijsten zijn vragen gesteld die te maken hebben met de gezondheid -en welzijnsbeleving van de klant. Het welbevinden van de deelnemer wordt gepeild aan de hand van vragen over de tevredenheid met het leven (schaal van 0 tot 10), de mate waarin men het leven betekenisvol of zinvol acht (schaal van 0 tot 10) en de mate waarin men zich gelukkig voelt (op een schaal van 1 tot 5)¹¹. In Fig. 3.5 presenteren we de voor compositieverschillen gecorrigeerde of geschatte kansen (Pwelbevinden) op welbevinden (het gemiddelde van tevredenheid, geluk en betekenisvol leven).

In de nulmeting zijn de geschatte scores op welbevinden iets hoger voor de eigen regiegroep en de onderzoeksgroep om vervolgens voor alle groepen toe te nemen in de meting daarna. Na meting 2 vlakt die groei weer af. Het verschil tussen de onderzoeksgroep en de andere groepen bij de start duidt erop dat er bij de onderzoeksgroep enige zelfselectie op welbevinden heeft plaatsgevonden voor het experiment. Figuur 3.5 doet vermoeden dat er in de tijd een stijging plaatsvindt bij alle treatmentgroepen. De stijging wijst echter niet op een treatment effect aangezien alle groepen tot aan meting 2 inclusief de onderzoek - of controlegroep een soortgelijke stijging in welbevinden laten zien.

¹¹ Deze drie vragen zijn vervolgens gebruikt om een gemiddelde score te berekenen van welbevinden waarbij de schaal van geluk eerst omgecodeerd werd naar 0 tot 10.



Noot: Pwelbevinden=geschatte kans op welbevinden.

Bron: Vragenlijsten deelnemers productevaluatie vertrouwensexperiment Wageningen, Tilburg University, 2020.

Fig. 3.5: Geschatte kans op welbevinden (geschatte kans) in de drie metingen

De OLS regressieanalyses (zie Tabel 3.8) laten zien welke treatmenteffecten of effecten over de tijd betekenisvol zijn. Ten eerste vinden we een significant effect over de tijd in meting 2 en 3 welke de stijging in welbevinden weergeeft. We vinden een treatment effect bij eigen regie in meting 2 welke een betekenisvol en positief treatment verschil laat zien met de controlegroep (rij totaal).

Tabel 3.8: Gemiddelde scores welbevinden gecorrigeerd voor compositieverschillen tussen treatmentgroepen (0-10)^{1,2}, 2017-2019

	Nulmeting	Meting 2	Meting 3	Totaal	N=
Eigen regie	5.78	6.18***	6.19	6.05	197
Extra begeleiding	5.76	6.07	6.07	5.92	165
Vrijlating	5.78	6.00	6.01	5.85	221
Onderzoekgroep	5.89	6.26	6.25***	6.16	165
Totaal	5.80	6.12**	6.13**	5.99	748

Noot¹: ***p<0,01; **p<0,05; *p<0,10, getest op basis van OLS regressiemodel (N=748); Model met treatment, meting, consulentnummer en de controle variabelen.

Noot²: Pseudo R²=0.128; Bijdrage verklaarde variantie door consulent=26%

Bron: Vragenlijsten deelnemers productevaluatie vertrouwensexperiment Wageningen, Tilburg University, 2020.

Voor het overige vinden we geen treatment effecten behalve bij de controlegroep in meting 3 ten opzichte van de nulmeting. Dat is opvallend omdat op grond van de literatuurstudie voorafgaand aan het onderzoek voor eigen regie en extra begeleiding wel betekenisvolle effecten verwacht werden. Heeft dit te maken met de uitvoering van de treatment of met de treatment zelf? Het welbevinden in de controlegroep blijkt na de nulmeting betekenisvol te zijn toegenomen hetgeen opvallend is en de vraag doet rijzen of de treatment hetzelfde is gebleven of is veranderd gedurende de experiment periode. In het eerste geval zouden de

welbevinden scores min of meer hetzelfde zijn gebleven als aan het begin van het experiment tenzij andere factoren zoals de economische situatie of de arbeidsmarktsituatie die in deze 2 jaar zijn verbeterd ook geleid hebben tot een stijging van het niveau van welbevinden. Helaas ontbreekt een andere vergelijkingsgroep dan de controlegroep om dit te kunnen testen. Wel geven de cijfers over welbevinden en geluk in Nederland voor de onderkant van de inkomensverdeling geen reden om een stijging te verwachten. Het SCP constateert zelfs in een recente publicatie dat de subjectieve levenssatisfactie score van mensen met een score van 6 of lager in de afgelopen jaren (hetgeen bij bijstandsgerechtigden het geval is) is afgenomen in plaats van gestegen ondanks de gunstige economische situatie (zie Wennekers e.a. 2019.). Naar deze bevinding is verder onderzoek nodig.

Lage welbevinden scores bij bijstandsgerechtigden

Tot slot valt op dat het niveau van welbevinden relatief laag is onder de deelnemers met een score van rond de 6; de gemiddelde score in Nederland is net onder 8 voor tevredenheid en geluk. Ook de verschillen tussen de deelnemers in de scores zijn relatief klein. De onderzochte groep van bijstandsgerechtigden scoort ruim twee punten lager dan gemiddeld in Nederland (zie SCP) hetgeen als een groot verschil kan worden gekenschetst. Eerder is geconstateerd dat de respons op de vragenlijsten is afgenomen na de nulmeting van 85% naar 60%. Dit heeft vermoedelijk te maken met een geringere motivatie om de vragenlijsten nog in te vullen na uitstroom uit de bijstand. Dit kan het beeld enigszins hebben gekleurd.

Invloed klantmanager op welbevinden

We hebben zoals gezegd ook een variabele in het model opgenomen die de ondersteuning door de klantmanager indiceert. Voegen we deze variabele toe aan het model dan stijgt de verklaarde variantie met 3,3% van 9,5% naar 12,8%. Verschillen in de ondersteuning door de klantmanager hebben dus invloed op het welbevinden van de deelnemer en verklaren voor een deel ook verschillen tussen de deelnemers in de verschillende treatments. De klantmanager doet er dus toe. De invloed van de treatments op welbevinden veranderen evenwel niet indien we de klantmanager-variabele toevoegen aan het model. Dat betekent dat als we een treatment effect vinden dat dat een echt treatment effect is en geen effect van de persoon van de klantmanager.

Subjectieve gezondheid

Een tweede en derde indicator die we gebruikt hebben zijn een indicator voor subjectieve gezondheid en mentale gezondheid. De tweede indicator gebruikt de algemene vraag over subjectieve gezondheid (Wat vindt u van uw gezondheid?, met als antwoord categorieën: uitstekend, zeer goed, goed, matig, slecht) alsook de vijf vragen over mentale gezondheid. De derde indicator maakt gebruik van vijf vragen over de mentale gezondheid die vragen naar de mate waarin men zich kalm en rustig voelt, men zich somber voelt en neerslachtig, zich droevig voelt, zich erg nerveus voelt en zich gelukkig voelt. De negatief gestelde items zijn omgecodeerd zodat alle items in dezelfde (positieve) richting wijzen. Voor 'de mate van' is gebruik gemaakt van een Likert schaal (altijd, vaak, soms, bijna nooit, nooit). Hiervan is een 10-puntsschaal geconstrueerd van 0-10 voor 'goede' mentale gezondheid die vervolgens weer is omgezet naar een positieve tien-puntsschaal van mentale gezondheid. Fig. 3.6 geeft de ontwikkeling op beide indicatoren weer.

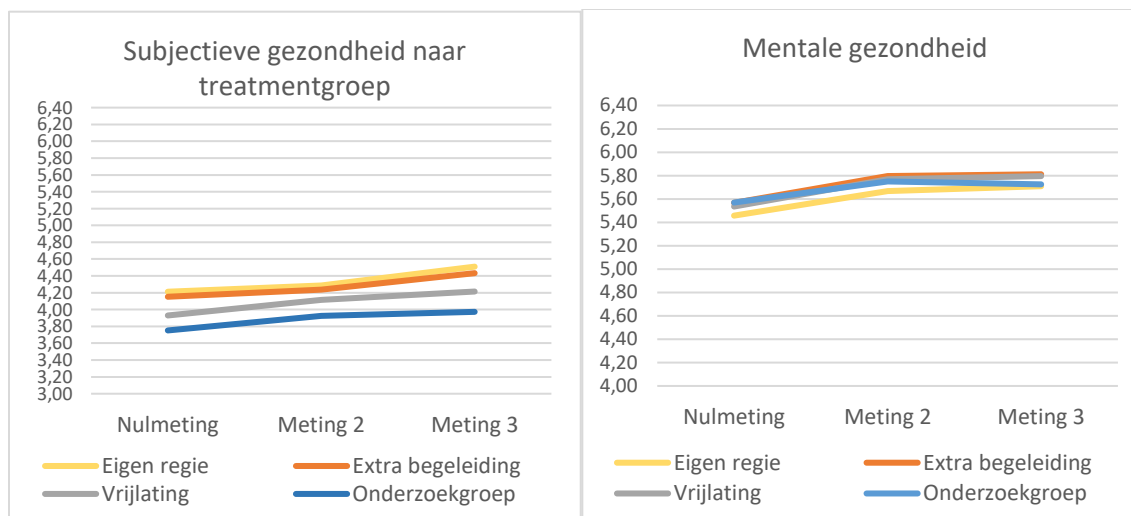


Fig. 3.6. Mate van subjectieve en mentale gezondheid naar treatment en meting, Wageningen 2017-2019

Tabel 3.9 geeft de effecten weer voor subjectieve en mentale gezondheid. Bij subjectieve gezondheid zien we voor extra begeleiding (maar niet bij eigen regie) een betekenisvol treatment effect. Dat houdt in dat het rapportcijfer dat deelnemers in extra begeleiding geven aan de eigen gezondheid gemiddeld is toegenomen met bijna een half punt ($dy/dx=.47$) op een schaal van 0 tot 10 ten opzichte van de controle of onderzoeksgroep.

Tabel 3.9. Gemiddelde scores subjectieve en mentale gezondheid (SGezh, MGezh) gecorrigeerd voor compositieverschillen tussen treatmentgroepen, 2017-2019

Treatment	Subjectieve gezondheid				Mentale gezondheid				N=
	Nul-meting	Meting 2	Meting 3	Tot.	Nul-meting	Meting 2	Meting 3	Totaal	
Eigen regie	4.21	4.29 ¹⁾	4.51	4.32	5.46	5.67 ^{**}	5.71	5.59	201
Extra begeleiding	4.15	4.24 ¹⁾	4.43 ^{**}	4.26 [*]	5.57	5.80	5.81 ^{**}	5.71	167
Vrijlating	3.93	4.12	4.21	4.07	5.54	5.77	5.80 ¹⁾	5.69	225
Onderzoeksgroep	3.75	3.92	3.97	3.88	5.57	5.75	5.73	5.68 ^{**}	170
Totaal	4.02	4.14	4.28 [*]	4.14	5.53	5.75	5.76 [*]	5.66	763

Noot: ^{***} $p < 0,01$; ^{**} $p < 0,05$; ^{*} $p < 0,10$, ¹⁾ $p < 0,15$; getest op basis van OLS regressiemodel (N=763); Model met treatment, meting, klantmanagernummer en de controle variabelen.

Noot²: $Pseudo R^2 = 0.21$ voor SGezh; 0.09 voor MGezh. ; Bijdrage verklaarde variantie door klantmanager=12%;28%.

Bron: Vragenlijsten deelnemers productevaluatie vertrouwensexperiment Wageningen, Tilburg University, 2020.

Tevens vinden we een significant effect op subjectieve gezondheid in meting 3 bij deze groep maar die geldt voor alle groepen (rij totaal). Voor mentale gezondheid zien we een significant treatment effect voor de groep eigen regie in meting 2 en voor extra begeleiding in meting 3 maar of dat een treatment of een ander effect is weten we niet omdat deze stijging zich ook voordoet bij de controlegroep. Is de treatment in de controlegroep verandert dan kan het toch een treatment effect zijn. Overigens stijgt de mentale gezondheid in alle vier groepen, inclusief de vrijlatingsgroep. De enigszins opvallende stijging van mentale

gezondheid bij de controlegroep in meting 3, dat is dus ten opzichte van de controlegroep in de nulmeting, is significant op 95% niveau. Dit zou erop kunnen wijzen dat de onderzoeksgroep treatment mogelijk toch is veranderd. De veronderstelling was dat beide treatments ten opzichte van de controlegroep een positieve bijdrage aan ervaren gezondheid zouden kunnen leveren. Deze hypothese wordt dus deels bevestigd maar mogelijk speelt zoals gezegd een rol dat de reguliere begeleiding in de controlegroep zelf mee is veranderd. Uit de literatuur bleek dat intensieve begeleiding en coaching middels “Individual Placement Support” en “Employment Support”, dat zijn maatgerichte programma’s voor mensen met fysieke of mentale gezondheidsproblemen en specifiek gericht op hun bijzondere problematiek, soelaas kan bieden om hen naar werk te begeleiden. Deze specifieke programma’s zijn in dit experiment niet toegepast maar kennelijk draagt extra begeleiding en maatwerk bij tot ervaren gezondheid.

3.5.2 Keuzemogelijkheden (capabilities)

Voor veel mensen is keuzevrijheid een belangrijk goed. De Nobelprijswinnaar Stiglitz heeft in zijn boek “De prijs van ongelijkheid” (2010) de zere vinger gelegd op het probleem van de gebrekkige ‘opportunities’ die mensen aan de onderkant van de inkomensverdeling in Amerika ervaren en hebben. Dat probleem geldt niet alleen voor de VS, ook in Nederland zijn er voor veel mensen weinig kansen en mogelijkheden om de eigen situatie te verbeteren en de dingen te doen die men voor het eigen leven belangrijk vindt. In de visie van Stiglitz is dat slecht voor het welbevinden van de mensen zelf maar ook voor de economie. Sen (1985) benadrukt dat mensen de kansen of mogelijkheden (keuzeopties) moeten hebben of krijgen om de dingen te doen in het leven die men belangrijk vindt. Het gaat zowel bij Sen als bij Stiglitz niet om gelijkheid van uitkomsten maar om gelijkheid van kansen en mogelijkheden. Om die te bieden is een ongelijke behandeling soms nodig om die gelijke startposities te kunnen realiseren terwijl in beleid veelal het gelijkheidsbeginsel leidend is. De toegang tot de arbeidsmarkt is voor mensen met beperkingen van welke aard ook beperkt en vraagt om extra ondersteuning. Het experiment is erop gericht om die kansen en mogelijkheden voor mensen aan de onderkant te bieden (maatwerk) of in ieder geval hen te leren die kansen zelf te scheppen en voor zover aanwezig te grijpen. Er is in het experiment meer vrijheid voor de klantmanager om aan de vraag van de deelnemer voor ondersteuning te voldoen of zelf met nieuwe oplossingen of instrumenten te komen hoewel het instrumentarium in Wageningen is beperkt tot de instrumenten beschikbaar in het sociale domein. Er zijn ook minder verplichtingen voor de deelnemers van beide treatments in het experiment of beter gezegd, er wordt op een andere manier met verplichtingen omgegaan waardoor de betrokkenen meer keuzevrijheid ervaren en hebben.

In de vragenlijst is aan het aspect keuzevrijheid aandacht besteed doordat vragen zijn gesteld over de dingen die men belangrijk vindt in het leven en die men op de een of andere manier wel of niet kan doen omdat men daar de mogelijkheden wel of niet voor heeft (keuzemogelijkheden). Het zijn 7 vragen gebaseerd op Sen’s capabilities benadering (1985) die beogen inzicht te krijgen in de keuzevrijheid ofwel in de mate waarin mensen de mogelijkheden hebben om de doelen die men stelt voor het eigen leven en die men belangrijk en waardevol vindt ook daadwerkelijk in het werk of anderszins te realiseren. De lijst zou een inzicht moeten geven in de keuzemogelijkheden en beperkingen die mensen ervaren in het leven in de bijstand. In de vragenlijst zijn zeven items voorgelegd aan de deelnemers. Ten eerste wordt gevraagd

of ze deze items belangrijk vinden en ten tweede of ze de mogelijkheden hebben om dit te kunnen doen in hun eigen situatie. De zeven items zijn:

- Meebeslissen over belangrijke dingen in werk en leven
- Goede contacten met anderen hebben
- Eigen doelen stellen
- Een voldoende inkomen hebben
- Iets waardevols bij te dragen aan het leven van anderen
- Om te kunnen leren en nieuwe dingen te doen (kennis en vaardigheden ontwikkelen en gebruiken)
- Om iets waardevols bij te dragen aan het leven van anderen. Gevraagd wordt op een vijfpunts Likert schaal naar de mate waarin men dit belangrijk vindt en in de eigen situatie ook kan doen: helemaal niet, niet, soms wel, soms niet, vaak of altijd.

Capabilities indicator (keuzemogelijkheden)

In overleg met de andere onderzoekers in LOEP is een soort capabilities indicator ontwikkeld die aangeeft in welke mate mensen in staat zijn, dus keuzemogelijkheden hebben, om dingen die men belangrijk vindt te realiseren. Daarbij zijn twee varianten onderscheiden. De eerste variant berekent een score op basis van uitsluitend de items die men vaak of altijd belangrijk vindt. Items die men niet of minder belangrijk vindt worden niet meegenomen en krijgen een gewicht 0. De tweede variant weegt de score op de mate dat men deze zeven dingen kan doen met de mate waarin ze belangrijk worden geacht. Daardoor worden alle items meegenomen in de score, ook items die men minder belangrijk vindt. Deze laatste krijgen dan wel een lager gewicht. De antwoord categorieën (helemaal niet tot altijd) worden omgezet in een score van 0 tot 10. Vervolgens wordt een somscore bepaald van alle items gewogen met de mate van belangrijkheid (score 1=helemaal niet tot 5=altijd). De laatste score wordt gedeeld door 5 om gewichten te krijgen variërend van 0,2 tot 1. Deze tweede variant wordt hier gebruikt. In Fig. 3.7 worden de resultaten weergegeven. Het niveau van keuzemogelijkheden is gemiddeld bij alle treatmentgroepen laag te noemen met een gemiddeld niveau van 4.0 op een schaal van 0 tot 10. Qua keuzemogelijkheden scoren de deelnemende bijstandsgerechtigden dus een ruime onvoldoende. Reeds in de nulmeting is het niveau van ervaren keuzemogelijkheden of capabilities groter in de eigen regiegroep en de controlegroep dan in de extra begeleidingsgroep en de vrijlatingsgroep.

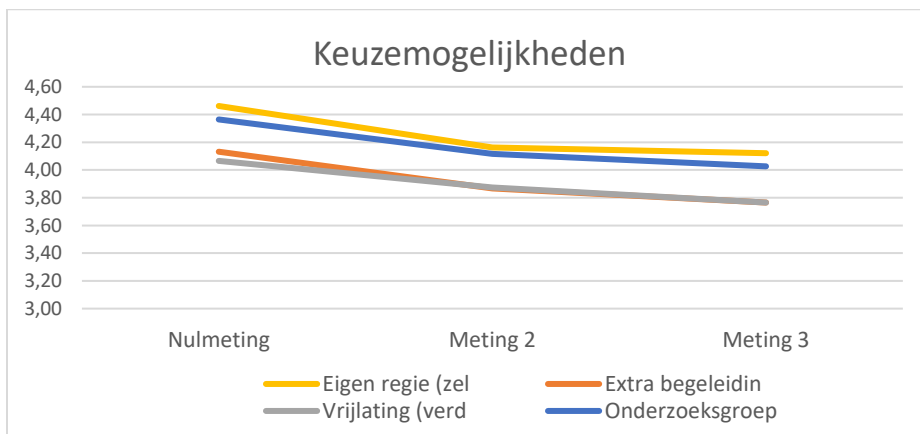


Fig. 3.5.7. Gecorrigeerde keuzemogelijkheden naar treatment en meting

De ervaren keuzemogelijkheden dalen in de tijd gezien voor alle groepen ongeveer in dezelfde mate. Na meting 1 is de daling het kleinst in de extra begeleidingsgroep en de eigen regiegroep. De afstand tussen de groepen verandert nauwelijks over de tijd. In tabel 3. 10 worden de geschatte effecten voor keuzemogelijkheden weergegeven. De verschillen tussen de treatments blijken bij toetsing niet significant te zijn. Over de tijd gezien zijn de verschillen tussen de groepen niet significant maar vinden we vanwege de dalende trend wel een significant negatief effect in de tijd in meting 3 (zie rij totaal). In het model zonder consulentnummer vinden we een negatief treatment effect in de vrijlatingsgroep maar met toevoeging van consulentnummer is dat effect niet meer significant. Kennelijk draagt de klantmanager bij aan verbetering van de ervaren keuzemogelijkheden.

Tabel 3.10. Gemiddelde scores keuzemogelijkheden (0-10) gecorrigeerd voor compositieverschillen tussen treatmentgroepen, 2017-2019

Treatments	Nulmeting	Meting 2	Meting 3	Totaal	N=
Eigen regie	4.46	4.16	4.12	4.28	201
Extra begeleiding	4.13	3.87	3.77	3.95	167
Vrijlating	4.07	3.87	3.77	3.92	225
Onderzoekgroep	4.36	4.12	4.03	4.17	170
Totaal	4.25	4.00	3.92*	4.08	763

Noot¹: ***p<0,01; **p<0,05; *p<0,10, getest op basis van OLS regressie analyse (N=763); Model met treatment, meting, klantmanagernummer en de controle variabelen.

Noot²: *Pseudo R*²=0.18; Bijdrage verklaarde variantie door klantmanager=22%.

Bron: Vragenlijsten deelnemers productevaluatie vertrouwensexperiment Wageningen, Tilburg University, 2020.

Dat de ervaren keuzemogelijkheden over alle groepen zo sterk dalen is opvallend en kan te maken hebben met een beter maar realistischer inzicht in de eigen arbeidsmarktkansen en mogelijkheden. De dalende trend kan ook veroorzaakt worden doordat mensen voltijds werk vinden en vervolgens minder genegen zijn de vragenlijst in te vullen terwijl zij wellicht de meeste keuzemogelijkheden hebben. Wanneer we de uitstroom naar werk als covariaat opnemen heeft deze echter geen significant positief effect op keuzemogelijkheden. Vooralsnog gaan we dus ervan uit dat de treatments ertoe leiden dat men meer inzicht krijgt in de beperkte mogelijkheden die men heeft.

3.5.3 Sociale participatie

De bevordering van sociale participatie is voor veel gemeenten een belangrijk doel van het experiment. Zelfs indien de uitstroom naar betaald werk niet haalbaar is voor de deelnemer vanwege zijn of haar problemen en beperkingen, kunnen het experiment en de treatments er wel toe bijdragen dat de sociale participatie van de betrokkenen wordt gestimuleerd. Uit de literatuur over geluk en welbevinden is bekend dat sociale participatie van belang is om verschillen in geluk tussen mensen en veranderingen in geluk van mensen te kunnen verklaren (zie Wenneker e.a., 2019; Headey et al. 2010)

In de vragenlijsten zijn daartoe vragen gesteld over de mate waarin men zich opgenomen voelt in de samenleving dus de perceptie van sociale integratie in de samenleving. Daarnaast zijn vragen gesteld over

de sociale participatie, dat is het aantal uren per week dat men besteed aan vrijwilligerswerk en mantelzorg, en er zijn vragen gesteld over het sociale netwerk door te vragen naar het aantal keer per maand dat men contact heeft met buurtgenoten, kennissen en vrienden en familie. Deze vragen zijn nu gebruikt om sociale participatie indicatoren te construeren. We zullen elk van deze drie indicatoren apart bespreken.

- *Perceptie sociale integratie*: De eerste gaat uit van de beleving van de bijstandsgerechtigde deelnemer over de mate waarin hij of zij zich welkom en opgenomen voelt in de samenleving;
- *Sociale participatie*: De tweede baseert zich op het aantal contacten per maand met anderen in het sociale netwerk (buren, kennissen-vrienden, familie). Deze indicator meet de frequentie van contacten in het sociale netwerk. De derde indicator betreft tot slot het aantal uren besteed aan vrijwilligerswerk en mantelzorg (maatstaf voor sociale participatie).

De gemeente acht de bevordering van de sociale participatie een belangrijk doel van het experiment. Voor een belangrijk deel van de cliënten is dit ook het hoogst bereikbare doel vanwege de grote afstand tot de arbeidsmarkt als gevolg van gezondheidsproblemen of een lage opleiding. Meer dan 55% heeft fysieke of mentale gezondheidsproblemen zoals uit de eerste tussenrapportage en de inleiding bleek (zie Muffels et al. 2018). In de procesevaluatie later in dit rapport zullen we de verandering op de participatieladder, een andere indicator voor sociale participatie, ook bekijken zoals deze door de consulenten gedurende de experimenteerperiode wordt geregistreerd in de administratie. Eerst wordt de beleving of perceptie van sociale integratie onderzocht, ofwel de mate waarin men zich opgenomen of welkom voelt in de samenleving en hoe deze zich ontwikkeld heeft tijdens het experiment (Fig. 3.8 links).

Opvallend is dat de controlegroep bij aanvang reeds de hoogste score kent welke 0,6 punt boven die van de vrijlatingsgroep ligt. Dat geldt in iets mindere mate ook voor de eigen regie groep. Dit kan duiden op enige zelfselectie in deze groepen. Vanaf de nulmeting daalt het niveau van ervaren sociale integratie licht voor alle groepen maar sterker voor de eigen regiegroep. De eigen regie groep herstelt zich echter enigszins van deze daling na meting 2 waarna het weer stijgt. De verschillen over de tijd en tussen de afzonderlijke treatmentgroepen en de onderzoekgroep zijn evenwel al bij de nulmeting niet significant en later ook niet. In Fig. 3.8 (rechts) worden de effecten voor het aantal contacten in het sociale netwerk weergegeven. De grafiek laat een opvallend beeld zien, aanvankelijk voor alle groepen een sterke daling en vervolgens na meting 1 een opvallende stijging die de aanvankelijke daling weer geheel teniet doet.

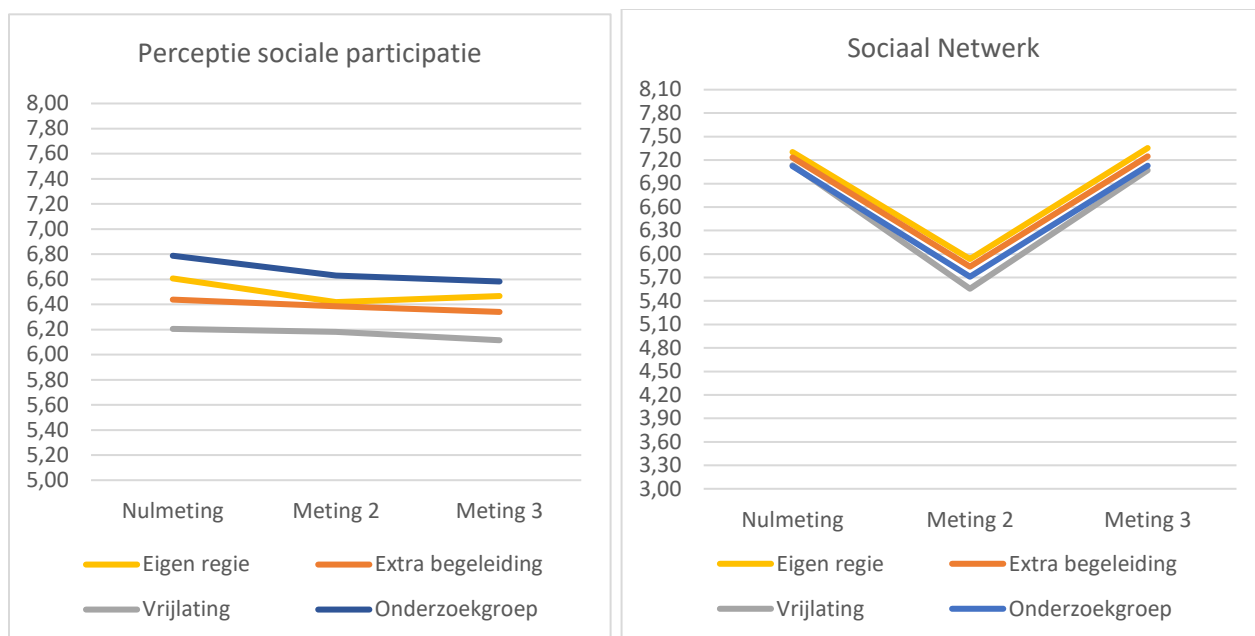


Fig. 3.8. Gecorrigeerde indicator perceptie sociale integratie (links) en participatie sociale netwerken (rechts) naar treatment en meting

In Tabel 3.11 worden de resultaten en treatment effecten voor beide uitkomstmaten per meting weergegeven. De daling tussen meting 1 en 2 in sociale contacten vertaalt zich in een significant negatief effect in meting 2 op het aantal contacten in het sociale netwerk in alle groepen. Vooralsnog is onduidelijk of dit een statistisch artefact is of een werkelijk treatment effect. We vermoeden het eerste.

Tabel 3.11. Gemiddelden geschatte kansen perceptie sociale integratie en sociale netwerken (0-10) gecorrigeerd voor compositieverschillen tussen treatmentgroepen, 2017-2019

Treatments	Sociale Participatie			Tot.	Sociale netwerk			Tot.	N=
	Nul-meting	Meting 2	Meting 3		Nul-meting	Meting 2	Meting 3		
Eigen regie	6.61	6.42	6.47	6.51	7.30	5.94***	7.35	6.94	201
Extra begeleiding	6.44	6.39	6.34	6.39	7.23	5.84*	7.25	6.79	167
Vrijlating	6.20	6.18	6.11	6.17	7.13	5.56**	7.07	6.61	225
Onderzoekgroep	6.79	6.63	6.58	6.67	7.12	5.71***	7.13	6.64	170
Totaal	6.49	6.39	6.36	6.42	7.20	5.75**	7.19	6.74	763

Noot¹: ***p<0,01; **p<0,05; *p<0,10, getest op basis van OLS regressie analyse (N=763); Model met treatment, cliëntvragenlijstmeting, klantmanagernummer en de controle variabelen.

Noot²: Pseudo R²=0.14 (links);0.14 (rechts). Bijdrage verklaarde variantie door klantmanager=42%;10%

Bron: Vragenlijsten deelnemers productevaluatie vertrouwensexperiment Wageningen, Tilburg University, 2020.

Vrijwilligerswerk en mantelzorg

Tot slot bespreken we de resultaten voor de geschatte uren besteed aan vrijwilligerswerk en mantelzorg.

In Fig. 3.9 worden de resultaten voor vrijwilligerswerk en mantelzorg weergegeven. Het aantal uren besteed aan vrijwilligerswerk en mantelzorg is gemiddeld ongeveer 9,5 uur in de nulmeting maar daalt over de tijd voor alle groepen min of meer in dezelfde mate naar gemiddeld 8 uur in meting 3. Opvallend is dat de onderzoek- of controlegroep bij de start al bijna 3 uur per week meer aan deze activiteiten besteed als de eigen regiegroep. Dat verschil blijft bestaan na de nulmeting. Dit duidt mogelijk op enige zelfselectie in deze groep. De vrijlatingsgroep en de eigen regiegroep liggen dicht bij elkaar bij de start en blijven dat doen nadien. De daling kan veroorzaakt worden door selectieve uitval van mensen die uitstromen naar deeltijdwerk of voltijdwerk en niet meer bereid zijn de vragenlijst in te vullen. Indien degenen die werk vinden ook degenen zijn die vrijwilligerswerk doen daalt daardoor vanwege deze uitval het gemiddeld aantal uren. Het opnemen van een interactie tussen treatment en voltijds of deeltijds uitstromen naar werk laat zien dat dit inderdaad een verklaring biedt voor de daling bij zowel de extra begeleidingsgroep als de vrijlatingsgroep aangezien de treatment effecten dan insignificant worden en overgenomen worden door het effect van werk vinden welke gemiddeld leidt tot een daling in het aantal uren vrijwilligerswerk en mantelzorg met 8 uur per week. De daling in het aantal uur vrijwilligerswerk en mantelzorg is dus het gevolg van het vinden van werk en dus geen negatief treatment effect op sociale participatie.

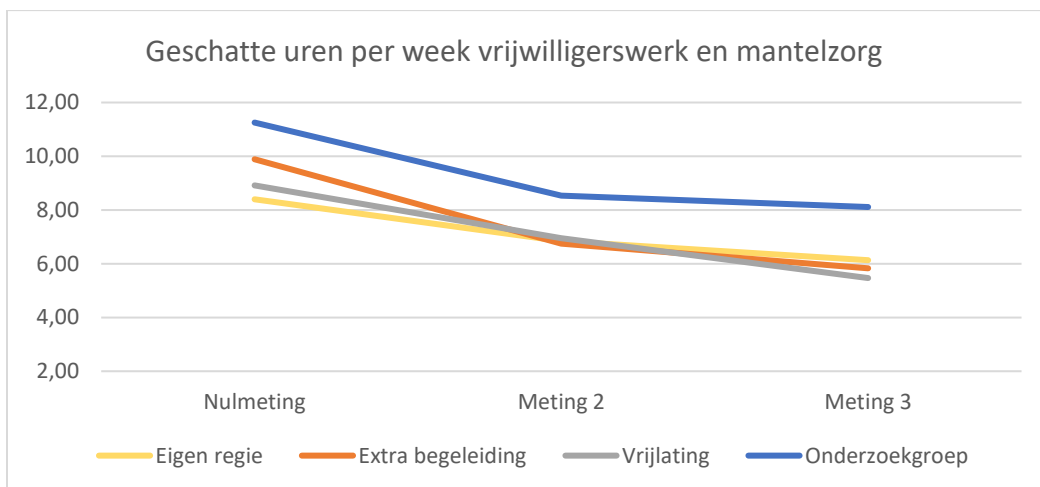


Fig. 3.9. Gecorrigeerde indicator uren besteed aan vrijwilligerswerk en mantelzorg naar treatment en meting, Wageningen 2017-2019

De effecten worden weergegeven in Tabel 3.12. Wat dit aspect betreft van sociale participatie vinden we over de hele linie een daling maar vooral bij de extra begeleidingsgroep en de vrijlatingsgroep. Deze daling over de tijd verwijst naar een treatment effect aangezien bij beide treatments we een significant negatief effect vinden zoals Tabel 3.12. laat zien. Het treatment effect blijkt echter zoals eerder uitgelegd verklaard te worden door de uitstroom naar werk, deeltijds en voltijds hetgeen de sociale participatie juist doet toenemen in plaats van afnemen.

Tabel 3.12. Geschatte uren per week besteed aan vrijwilligerswerk en mantelzorg gecorrigeerd voor compositieverschillen tussen treatmentgroepen, 2017-2019

Treatments	Nulmeting	Meting 2	Meting 3	Totaal	N=
Eigen regie	8.40	6.85 ¹⁾	6.13*	7.37	203
Extra begeleiding	9.89	6.75 ¹⁾	5.83	7.71*	168
Vrijlating	8.92	6.95*	5.47**	7.29**	225
Onderzoekgroep	11.25	8.54	8.11	9.35	170
Totaal	9.44	7.26*	6.35**	7.86	766

Noot¹: ***p<0,01; **p<0,05; *p<0,10, getest op basis van OLS regressie analyse (N=766); Model met treatment, cliëntvragenlijstmeting, klantmanager-nummer, nieuwe instroom/zittend bestand, en de controle variabelen.

Noot²: *Pseudo R*²=0.11; Bijdrage verklaarde variantie door klantmanager=27% (0.03).

Bron: Vragenlijsten deelnemers productevaluatie vertrouwensexperiment Wageningen, Tilburg University, 2020.

3.5.4 Zelfredzaamheid

De treatments zijn in belangrijke mate gestoeld op het idee dat het bevorderen van de zelfredzaamheid van de deelnemers uiteindelijk een belangrijk middel is om sociale participatie te verbeteren. De training die door Roland Blonk is aangeboden is gebaseerd op de theorie van gepland gedrag van Fishbein & Ajzèn (2010). In de training van de klantmanagers is door Blonk (2019) een zoekgedrag model gebruikt welke gebaseerd is op die theorie en die als basis diende voor de zogenaamde ‘skills’ methodiek waarin coaches/klantmanagers vaardigheden leren om de zelfredzaamheid van de deelnemers te bevorderen. Dat had niet alleen betrekking op de ‘eigen regie’ treatment maar ook op de ‘intensieve begeleiding’ treatment hoewel de wijze van coaching in deze laatste treatment anders, wellicht wat ‘meer sturend’ is als in de eerste. Maar ook in de eigen regie treatment wordt de deelnemer aangesproken op zijn of haar activiteiten en gedrag. In de vragenlijst is aan de mate van zelfredzaamheid daarom ook aandacht besteed. Er wordt onder andere gevraagd aan de deelnemer of men zichzelf zelfredzaam acht in verband met het zoeken en vinden van werk bijvoorbeeld. De lijst van capabilities eerder behandeld geeft ook inzicht in de belemmeringen en kansen om zelfredzaam te zijn, dat wil zeggen de dingen te kunnen doen die men belangrijk vindt voor het eigen leven. Wat zelfredzaamheid betreft op het gebied van werk zijn in de vragenlijst een viertal vragen gesteld daarover. Dat zijn:

- Ik kan betaald werk vinden als ik er echt moeite voor doe
- Ik denk dat ik in de toekomst betaald werk zal vinden
- Ik denk dat ik bij sollicitaties een goede indruk kan maken
- Ik denk dat ik een baan kan vinden die past bij mijn opleiding en ervaring.

De antwoordcategorieën vormen een 5-punts Likert schaal van ‘geheel mee oneens’ tot ‘geheel mee eens’. Deze vier items zijn vervolgens omgezet in een schaal van 0 (geheel oneens) tot 10 (geheel eens) en er is een gemiddelde bepaald over de vier items. Dat noemen we de indicator zelfredzaamheid. Vervolgens is het standaard regressiemodel gedraaid om de voor verschillen in samenstelling gecorrigeerde kansen op zelfredzaamheid voor elk van de groepen te bepalen. In Figuur 3.1.10 worden de resultaten in beeld gebracht.

Treatment effecten

De zelfredzaamheid neemt eerst licht toe tot aan meting 2 om daarna weer af te nemen tot iets onder het oorspronkelijke niveau. Opvallend is het relatief hoge niveau van zelfredzaamheid bij de nulmeting bij eigen regie hetgeen duidt op enige zelfselectie in deze groep. Het niveau is bij deze groep 0.6 punten hoger op een schaal van 0 tot 10 bij de start waarna het verschil hetzelfde blijft. De toe -en afname in zelfredzaamheid over de tijd die we bij alle groepen waarnemen is niet significant. De afname na meting 2 kan verband houden met het feit dat degenen met de grootste afstand tot de arbeidsmarkt overblijven. Zij ervaren en worden zich meer bewust van hun eigen (minder gunstige) positie op de arbeidsmarkt. Dat is dan selectie aan het einde in meting 3 omdat de groep met de kleinste kansen op werk is overgebleven en de mensen die werk hebben gevonden de vragenlijst niet meer invullen. Indien we in het model zonder consulentnummer het wel of niet werk hebben gevonden gedurende het experiment als variabele toevoegen aan het model is deze werk variabele significant ($p < 0.025$). Deze werk variabele blijft significant in het model met consulentnummer maar wordt iets minder sterk. Dit geeft aan dat het werk vinden deels te maken heeft met de begeleiding door de consulent.

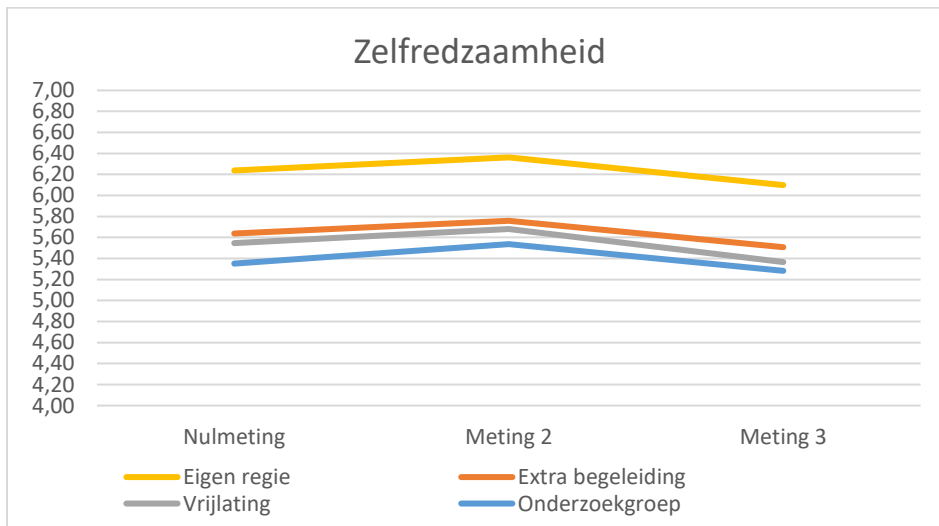


Fig. 3.10. Geschatte kans zelfredzaamheid bij zoeken naar en vinden van werk, naar treatment en meting

Causaal gezien leidt zelfredzaamheid tot uitstroom naar werk en niet andersom hetgeen betekent dat de meest zelfredzame deelnemers het eerst uitstromen. In interactie met de treatments is het effect van de werk variabele (deeltijds of voltijds) bijna significant bij de eigen regie groep ($p < 0.14$) waarin zelfredzaamheid het uitgangspunt van de treatment is. De meest zelfredzame mensen vinden het eerst werk maar in vergelijking met de onderzoekgroep vooral in de eigen regiegroep deels omdat er aan het begin al zelfselectie heeft plaatsgevonden. Tabel 3.13 geeft het gemiddelde van de geschatte kansen per treatment en meting en de significantieniveaus.

Tabel 3.13 Gemiddelde scores zelfredzaamheid (0-10) gecorrigeerd voor compositieverschillen tussen treatmentgroepen, 2017-2019

Treatments	Nulmeting	Meting 2	Meting 3	Totaal	N=
Eigen regie	6.24	6.36 ^{*1)}	6.10	6.23 ^{*1)}	203
Extra begeleiding	5.64	5.76	5.51	5.64	168
Vrijlating	5.55	5.68	5.37	5.54	225
Onderzoeksgroep	5.35	5.54	5.28	5.39	170
Totaal	5.73	5.82	5.56	5.71	766

Noot: *** $p < 0,01$; ** $p < 0,05$; * $p < 0,10$, getest op basis van OLS regressiemodel (N=766); Model met treatment, meting, klantmanagernummer en de controle variabelen. ¹⁾ Effect is significant in model zonder consulentnummer; ²⁾ *Pseudo R*²=0.34; Bijdrage verklaarde variantie door klantmanager = 3%.

Bron: Vragenlijsten deelnemers productevaluatie vertrouwensexperiment Wageningen, Tilburg University, 2020.

3.5.5 Sociaal en institutioneel vertrouwen

Het experiment is vertrouwensexperiment genoemd omdat het vertrekt vanuit het idee dat het opbouwen van een vertrouwensrelatie met de deelnemer uiteindelijk effectiever kan zijn om het gewenste doel te bereiken, namelijk bevordering van sociale participatie, dan de standaardaanpak met een strenge toepassing van de verplichtingen en bij niet nakoming sanctionering van de uitkering. De toetsing van rechtmatigheid blijft intact maar de toetsing aan de andere (informatie, sollicitatie, tegenprestatie) verplichtingen wordt op een meer relaxte manier vormgegeven in nauwe afstemming en overleg met de klant. Het is zeker geen 'loslaten' maar leren de eigen verantwoordelijkheid te nemen en zelfredzaam te zijn met het aanpakken van de problemen die men heeft en ervaart. Uit de experimentele economische literatuur weten we bijvoorbeeld dat het geven van vertrouwen leidt tot wederkerigheid, tot extra inspanning en verbeterde motivatie en productiviteit in reactie op het gestelde vertrouwen (Frey & Jegen 2001; Bohnet et al., 2001). Uit de psychologische motivatietheorie weten we dat extrinsieke motivatie ten koste kan gaan van de intrinsieke motivatie en dat sturing op intrinsieke motivatie beter kan werken. Vandaar ook de inbouw van de training van de klantmanagers om hen vaardigheden te leren die hen helpen mensen te ondersteunen in zelfredzaamheid. Wageningen heeft als artikel 83 gemeente gekozen voor een vrijlatingsgroep die de extra vrijlating krijgt maar overigens de reguliere dienstverlening. Deze groep staat daarom toe het 'zuivere' effect van de beloning in de vorm van extra vrijlating van de verdiensten te onderzoeken. Ook bij de andere treatments vindt er beloning plaats maar dan in de vorm van een positieve bejegening (aandacht, een luisterend oor), handelingsvrijheid (eigen regie) en intensieve begeleiding en maatwerk (extra begeleiding). Doordat de caseload van de klantmanager verlaagd is van 125 naar 75 tot 90, is er meer tijd voor de begeleiding van de klant. Er is tijd om een 'relatie' en of persoonlijke band met de deelnemer op te bouwen omdat er meer tijd voor contact is en meer coaching op basis van vertrouwen. Het wantrouwen onder de deelnemers ten opzichte van de gemeente is groot zo hebben alle klantmanagers ervaren. Om daar verandering in te brengen kost tijd (zie verslag Focusgroep gesprekken hierna). Vertrouwen komt te voet en gaat te paard. Dus het vertrouwen is ook weer snel weg indien het vertrouwen wordt geschaad; dat geldt overigens voor beide partijen.

Over dit aspect van vertrouwen hebben we vragen gesteld in de vragenlijst. Ten eerste betreft dat een vraag over vertrouwen in de instituties. De vraag luidde in welke mate men vertrouwen had in:

- De landelijke overheid (regering, parlement)
- De gemeente
- De afdeling werk en inkomen van de gemeente
- De eigen klantmanager/werkcoach.

De antwoord categorieën vormde een vier-puntsschaal van ‘helemaal geen vertrouwen’, ‘niet zo veel vertrouwen’ ‘een beetje vertrouwen’ tot ‘veel vertrouwen’. Deze is omgezet naar een 0-10 schaal. Daarnaast is gevraagd naar het vertrouwen dat men in andere mensen heeft. Dat wordt in de literatuur aangeduid als sociaal vertrouwen. Het wordt gezien als een vorm van sociaal kapitaal, mensen met veel vertrouwen hebben veelal een rijk sociaal netwerk dat hen ook helpt om vooruit te komen in het leven. Sociaal vertrouwen of vertrouwen in andere mensen is gemeten met een vraag die afkomstig is uit het “European Social Survey” of men vindt dat de meeste mensen te vertrouwen zijn of dat men niet voorzichtig genoeg kan zijn op een schaal van 0 (geen vertrouwen) tot 10 (volledig vertrouwen).

In Fig. 3.11 wordt de ontwikkeling in vertrouwen volgens de twee maatstaven per treatment groep weergegeven. Het vertrouwen in anderen krijgt bij aanvang van de deelnemers gemiddeld een zes over alle groepen en het vertrouwen in de instituties een kleine 7. Het vertrouwen in de ander stijgt daarna maar het vertrouwen in de instituties neemt sterk af. Opvallend is dat reeds bij de start het niveau van sociaal en institutioneel vertrouwen bij de onderzoeksgroep of controlegroep aanzienlijk hoger is dan in de andere treatments. Kennelijk heeft er enige zelfselectie plaatsgevonden bij de start.

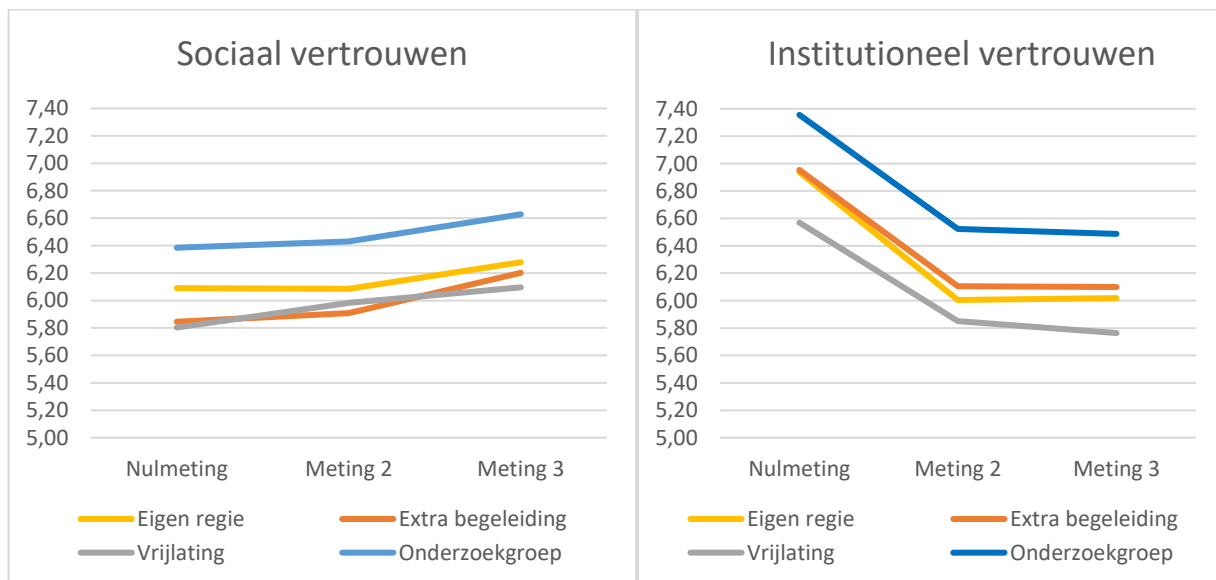


Fig. 3.11 Geschatte kans op sociaal vertrouwen en vertrouwen in instituties, naar treatment en meting, Wageningen 2017-2019

Het verschil bij de start blijft in stand nadien. Het niveau van sociaal vertrouwen stijgt in de controlegroep door terwijl het niveau van institutioneel vertrouwen verder daalt. Mogelijk is dit een indicatie dat de treatment niet helemaal hetzelfde is gebleven. De toename in sociaal vertrouwen vindt met name plaats na het eerste jaar bij alle groepen maar vooral ook bij de extra begeleidingsgroep (het effect is evenwel niet significant). Het niveau van institutioneel vertrouwen is bij aanvang hoger dan het niveau van sociaal vertrouwen maar neemt sterk af tijdens het experiment waardoor het aan het einde van het experiment lager is dan het niveau van sociaal vertrouwen. De daling vindt plaats bij alle groepen.

Het is onduidelijk waarom bij de start het niveau van sociaal en institutioneel vertrouwen bij de onderzoeksgroep hoger ligt dan bij de andere groepen. Het verschil is significant en negatief in het model zonder consulentnummer voor de drie treatments bij de start van het experiment. Om die reden is het geen treatment effect maar wellicht een selectie effect bij aanmelding. In het model met consulentnummer is het effect echter niet significant en statistisch dus 0. In Tabel 3.14 worden de bevindingen voor zowel institutioneel en sociaal vertrouwen in tabelvorm weergegeven. Bij institutioneel vertrouwen zien we een significant negatief effect bij alle groepen inclusief de controlegroep en bij de beide metingen Dit geeft de daling weer enerzijds bij de treatments en anderzijds over de tijd. Vooralsnog is onduidelijk waardoor deze daling wordt veroorzaakt.

Tabel 3.1.14 Gemiddelde scores institutioneel en sociaal vertrouwen (0-10) gecorrigeerd voor compositieverschillen tussen treatmentgroepen, 2017-2019

Treatments	Sociaal vertrouwen				Institutioneel vertrouwen				
	Nul-meting	Meting 2	Meting 3	Tot.	Nul-meting	Meting 2	Meting 3	Tot.	N=
Eigen regie	6.09	6.08	6.28	6.14	6.94	6.00***	6.02***	6.42	201
Extra begeleiding	5.84	5.91	6.20	5.96	6.95	6.10**	6.10**	6.44	167
Vrijlating	5.80	5.98	6.10	5.95	6.57	5.85**	5.76**	6.11	225
Onderzoeksgroep	6.38	6.43	6.63	6.48	7.36	6.52**	6.49**	6.80	170
Totaal	6.01	6.10	6.29 ¹⁾	6.12	6.92	6.11**	6.07**	6.42	763

Noot: *** $p < 0,01$; ** $p < 0,05$; * $p < 0,10$, getest op basis van OLS regressiemodel (N=763); Model met treatment, meting, klantmanagernummer en de controle variabelen. ¹⁾In model met klantmanagernummer zijn alle treatment effecten 0. Wel is er een 'net-niet' significant effect voor meting 3 die de positieve verandering in de tijd weergeeft ($p < 0,11$). Bij schatting model zonder consulentnummer vinden we voor alle drie treatments significante negatieve effecten in nulmeting, waarschijnlijk veroorzaakt door zelfselectie in de onderzoeksgroep.

Noot²: Pseudo R^2 : *strust*=0.17; *itrust*=0.21. Bijdrage verklaarde variantie door klantmanager *strust*=30%; *itrust*=32%

3.5.5 Financiële situatie en armoede

Alle AMvB gemeenten waaronder Wageningen hebben gekozen voor een vrijlatingsgroep als een van de treatments. De vrijlating houdt in dat in plaats van 25% van de verdiensten die niet worden ingehouden op

de uitkering in de eerste 6 maanden conform de huidige regels in de Participatiewet, nu gedurende de gehele 2-jarige experimenteer periode 50% van de verdiensten worden vrijgelaten tot een maximum van 199 euro per maand zoals nu ook al het geval is. Voor deze vrijlatingsgroep wordt de financiële situatie dus enigszins verbeterd waardoor dit mogelijk effect heeft op de ervaren stress die een slechte financiële situatie en armoede met zich meebrengt. Wanneer men door de vrijlating extra uren gaat werken is de kans op overstap naar een voltijdse baan mogelijk ook groter en wordt daardoor de financiële situatie extra verbeterd. Daarnaast kunnen er inkomenseffecten ontstaan ook in de andere treatments doordat een andere bejegening en aanpak ertoe leidt dat meer mensen voltijds werk vinden en volledig onafhankelijk worden van een uitkering of dat schuldproblemen worden aangepakt of dat mensen gestimuleerd worden om hun huishoudboekje beter bij te houden.

Een van de veronderstellingen die we op grond van de psychologische literatuur naar voren hebben gebracht is dat door ervaren financiële stress mensen cognitief en motivationeel daar zo mee bezig zijn dat er mentaal geen ruimte is voor activiteiten die de eigen situatie op iets langere termijn zouden kunnen verbeteren zoals zoeken naar werk, het volgen van een opleiding, het inschakelen van het sociale netwerk en dergelijke.

Inkomensarmoede

Er zijn twee maatstaven gebruikt om de financiële situatie en de armoedesituatie in beeld te brengen. De eerste is een maatstaf welke veel is gebruikt in onderzoek naar de inkomenssituatie van huishoudens in Nederland. Daarbij wordt direct gevraagd naar de financiële situatie van het huishouden waarin men leeft. De vraag luidt als volgt:

Wat is op dit moment de financiële situatie van uw huishouden? De antwoordcategorieën zijn:

- Ik maak schulden
- Ik gebruik mijn spaarmiddelen om van te leven
- Ik kan precies rondkomen
- Ik houd een beetje geld over
- Ik houd genoeg over om te sparen
- Ik weet het niet
- Ik wil dit niet zeggen

In de vraag wordt tevens vermeld wat het begrip rondkomen inhoudt. Rondkomen is net genoeg geld hebben om de absoluut minimaal noodzakelijke kosten van levensonderhoud te kunnen betalen.

De eerste twee items zijn nu gebruikt om een persoon inkomensarm te noemen. Er zijn in de literatuur veel armoedemaatstaven ontwikkeld en ook verschillende inkomensarmoede maatstaven. De bijstandsgrens is zo'n grens die het SCP bijvoorbeeld hanteert naast een aantal andere. Een daarvan is het "niet-veel-maar-toereikend budget" dat ongeveer op het niveau van de bijstand van een alleenstaande ligt. We verwijzen naar de armoederapporten van het SCP die regelmatig verschijnen (Hoff e.a., 2019). Ook het CBS publiceert cijfers over armoede. Het SCP heeft berekend dat in 2014 ongeveer 45% van de mensen in de bijstand onder de "niet-veel-maar-toereikend" budgetgrens leven. Recente cijfers laten zien dat dit gedaald is naar 25-35% in 2019. De maatstaf voor armoede op basis van de vraag naar de financiële situatie noemen we inkomensarmoede en geeft het percentage deelnemers dat niet kan rondkomen met het huishoudinkomen en ofwel de spaarmiddelen aanspreekt om rond te kunnen komen ofwel schulden moet maken.

Deprivatiearmoede

De tweede maatstaf die we hebben gebruikt noemen we deprivatiearmoede en is gebaseerd op twee vragen waarin gevraagd wordt bij elk van de 14 items of men deze absoluut noodzakelijk vindt om te hebben of te kunnen doen en of men voldoende geld heeft om het te kunnen kopen of te kunnen doen. De lijst van 14 items is een selectie uit een lijst van 21 items die ook is bevestigd in onderzoek van de Europese Commissie, EU-SILC genaamd. We hebben er de 14 voor de Nederlandse samenleving meest relevante items uit geselecteerd. Dat zijn de volgende.

- 1 keer per dag groenten en fruit
- 1 maaltijd per dag met vlees kip of vis
- Twee paar schoenen
- Laptop of desktop computer
- Vrijtijdsactiviteiten buiten het gezin
- Vieren van bijzondere gebeurtenissen
- Jaarlijks 1 week op vakantie gaan in het buitenland
- Vervangen van beschadigde meubels
- Vervangen van huishoudelijke apparaten
- Kleding vervangen
- 1 keer per maand uit eten gaan
- Regelmatig deelnemen aan vrijetijdsactiviteiten
- Toegang tot internet
- Gebruik kunnen maken van openbaar vervoer

Vervolgens is het aantal items berekend dat men noodzakelijk vindt maar niet heeft. Voor de maatstaf van deprivatiearmoede zijn er vervolgens vijf essentiële items geselecteerd: 1x per dag groenten en fruit; 1 maaltijd per dag met vlees, kip of vis; vervangen beschadigde meubels; vervangen van huishoudelijke apparaten en versleten kleding vervangen. Indien iemand elk van deze items noodzakelijk vindt maar minstens drie van deze vijf items mist beschouwen we deze persoon als arm. We noemen iemand dan deprivatiearm. In Fig. 3.12 wordt de ontwikkeling van inkomens -en deprivatiearmoede overeenkomstig beide maatstaven gedurende het experiment in beeld gebracht.

Dan blijkt dat 20 tot 30% van de deelnemers in inkomensarmoede leeft, en zo'n 30 tot 40% in deprivatiearmoede. Voor een deel zijn dit dezelfde mensen, voor een deel andere mensen. Zeker ook het cijfer over de deprivatie van deze basisbehoeften is hoog. Niettemin laat Fig. 3.12 een daling zien van inkomensarmoede tussen nulmeting en meting 2. Daarna neemt de inkomensarmoede weer toe. De daling geldt evenwel sterker voor deprivatiearmoede en over de gehele periode. Voorts vinden we een positief en significant effect op deprivatiearmoede voor het zittende bestand in alle groepen hetgeen betekent dat de kans op armoede in het zittende bestand het grootst is.

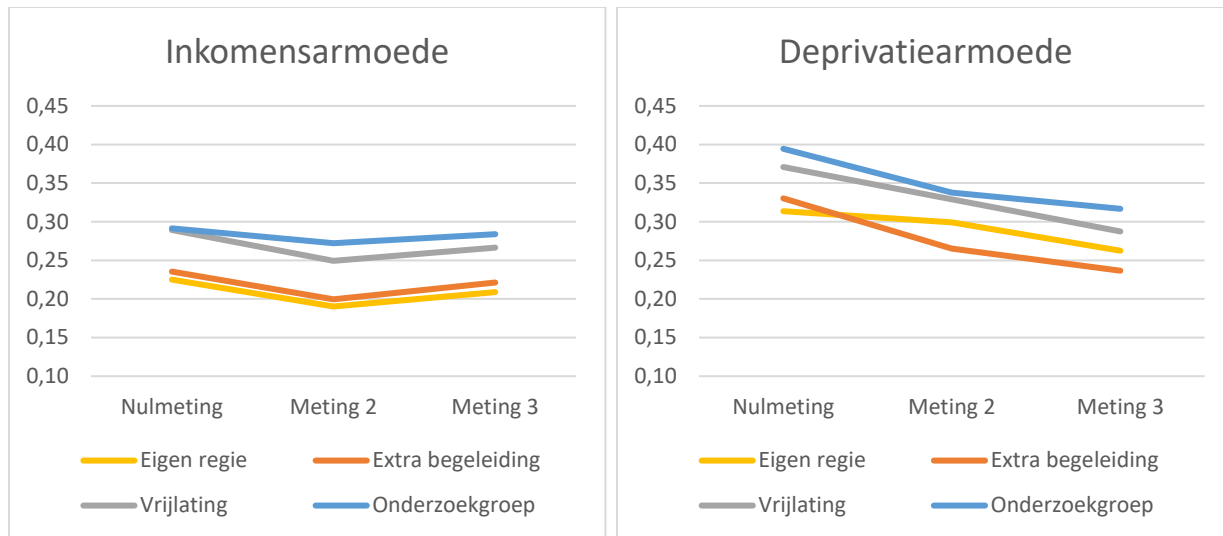


Fig. 3.12 Geschatte kansen op inkomensarmoede en deprivatiearmoede, Wageningen 2017-2019

De daling in armoede kan het gevolg zijn van de gunstige economische situatie maar het kan ook zijn dat de uitstroom naar werk van de deelnemers (19%) in Wageningen hier debet aan is geweest. Echter, zoals we eerder constateerden is deze uitstroomwinst niet beperkt tot de drie experimentgroepen maar realiseerde ook de onderzoekgroep gunstige uitstroomcijfers die enkele procentpunten hoger waren dan de uitstroomcijfers van de referentiegroep van niet-deelnemers. Controleren we in beide modellen voor uitstroom naar werk dan heeft dit geen effect op de inkomens -of deprivatiearmoede bij de groepen. Kennelijk zijn er andere factoren die het niveau van inkomensarmoede en deprivatiearmoede verklaren (zoals het inkomensbeleid).

Treatment effecten

De grafiek laat ook zien dat de inkomensarmoede vooral in het eerste jaar is gedaald met name bij de eigen regiegroep. Deze daling is zoals Tabel 3.14 laat zien significant ten opzichte van de onderzoekgroep (vierde kolom). We vinden tevens significante en negatieve treatment effecten (dus daling inkomensarmoede) voor alle groepen inclusief de controlegroep in meting 3 (geeft het effect weer ten opzichte van de nulmeting)¹². Het model verklaart echter weinig en laat vooral zien dat de onderzoekgroep bij de start reeds een hoger niveau van inkomensarmoede heeft hetgeen wijst op zelfselectie en niet op treatment effecten. Opvallend is dat we geen treatment effecten vinden bij deprivatiearmoede terwijl daar de daling groter is dan bij inkomensarmoede. Kennelijk zijn er andere redenen voor deze daling van deprivatiearmoede. De lage correlatie tussen beide maatstaven (0,27) bevestigt dat beide maatstaven iets anders meten. De eigen regiegroep heeft de laagste inkomensarmoede maar niet de laagste deprivatiearmoede. De controlegroep heeft wel de hoogste inkomensarmoede en deprivatiearmoede. Over het algemeen vinden we dus geen treatment effecten voor armoede en is ook de daling bij deprivatiearmoede niet significant na correctie voor compositieverschillen.

¹² De modellen voor inkomensarmoede zijn niet significant ($p > \chi^2 = .73$) en verklaren slechts 4% van de variantie. Het model zonder consulenteffecten verklaart slechts 2%.

Tabel 3.14 Gemiddelde scores inkomensarmoede en deprivatiearmoede (0-10) gecorrigeerd voor compositieverschillen tussen treatmentgroepen, 2017-2019

Treatments	Inkomensarmoede				Deprivatiearmoede				
	Nul-meting	Meting 2	Meting 3	Tot,	Nul-meting	Meting 2	Meting 3	Tot.	N=
Eigen regie	0.23***	0.19***	0.21**	0.21*	0.31	0.30	0.26	0.30	200
Extra begeleiding	0.24	0.20	0.22*	0.22	0.33	0.27	0.24	0.28	166
Vrijlating	0.29	0.25**	0.27	0.27	0.37	0.33	0.29	0.33	225
Onderzoekgroep	0.29	0.27	0.28**	0.28	0.39	0.34	0.32	0.35	168
Totaal	0.26	0.23	0.25	0.25	0.35	0.31	0.28	0.32	759

Noot¹: *** $p < 0,01$; ** $p < 0,05$; * $p < 0,10$, getest op basis van OLS regressie analyse $N=618$); Model met treatment, cliëntvragenlijstmeting, klantmanagernummer en de controle variabelen.

Noot²: *Pseudo R²: Inkomensarmoede 0.04; Deprivatiearmoede=0.10.* Bijdrage verklaarde variantie door klantmanager 48%;14%;

Bron: Vragenlijsten nulmeting en vervolgmetingen productevaluatie Vertrouwensexperiment Apeldoorn/Epe, Tilburg University, 2020

3.5.6 Samenvatting treatment effecten secundaire uitkomstmaten

In Tabel A1.1 in bijlage A presenteren we een overzicht van de gevonden treatment effecten ten opzichte van de onderzoek -of controlegroep en de effecten van de beide metingen voor alle drie groepen ten opzichte van de nulmeting op de niet-werk gerelateerde uitkomstmaten (ontwikkeling in de tijd). De meeste treatment effecten zijn niet significant (tabel A1.2). De meeste hypothesen worden daarom niet bevestigd met uitzondering van de verwachtingen ten aanzien van subjectieve gezondheid waarvoor we een positief en significant treatment effect vinden van extra begeleiding. Ook bij mentale gezondheid vinden we positieve en significante treatment effecten in meting 2 en meting 3 voor alle groepen ten opzichte van de controlegroep bij de start. Tevens vinden we voor mentale gezondheid een treatment effect in meting 3 voor de controlegroep, dus ten opzichte van de situatie bij deze groep bij de start. Kennelijk is er iets veranderd in de treatment van de controlegroep waardoor dit effect optreedt bij de deelnemer.

Bij welbevinden vinden we geen treatment effect ten opzichte van de controlegroep maar stijgt het welbevinden nochtans bij alle groepen betekenisvol. Datzelfde geldt ook voor mentale gezondheid. We vinden een betekenisvolle stijging in de tijd bij alle groepen. Dat er weinig treatmenteffecten gevonden worden bij deze uitkomstmaten heeft volgens ons te maken met twee aspecten. Ten eerste denken we dat we geen goede vergelijkingsgroep hebben omdat de standaard treatment misschien zelf is veranderd. Ten tweede denken we dat veel van de gekozen uitkomstmaten stabiele kenmerken zijn van mensen die niet zo snel veranderen, of het nu gaat om beleving en motivatie zoals bij welbevinden, subjectieve gezondheid, de perceptie van sociale integratie, vertrouwen of om gedrag zoals bij zoekgedrag, vrijwilligerswerk en sociale netwerken. Om deze kenmerken te veranderen zijn vaak ingrijpende gebeurtenissen nodig die het leven drastisch veranderen zoals een scheiding, langdurige werkloosheid, en verlies van kind partner of

vriend en vaak duurt het een lange tijd voor een effect zich openbaart door een andere beleving, motivatie of gedrag. Wat het eerste aspect betreft laten de grafieken ook zien dat de uitkomsten voor de controlegroep niet stabiel zijn en in dezelfde richting mee veranderen met de andere groepen. Bij de uitkomstmaat van mentale gezondheid vinden we bijvoorbeeld een positief significant treatment effect over de tijd bij de controlegroep in meting 3 (zie Tabel A1.2 in Bijlage A). Indien de standaard treatment niet was veranderd zou het vinden van een dergelijk effect onwaarschijnlijk zijn geweest. Daarnaast zagen we al op grond van de analyses van de uitstroom naar werk dat de standaard treatment naar alle waarschijnlijkheid is veranderd omdat er anders niet zo'n groot verschil zou zijn met de referentiegroep die de standaard treatment heeft gekregen. Met andere woorden, indien de standaard treatment inderdaad is veranderd, is een zuivere vergelijking met de treatment groepen lastig. Dan zullen we een vergelijking moeten maken met een vergelijkingsgroep die niet beïnvloed is door het experiment. Op dit moment is zo'n alternatieve vergelijkingsgroep niet beschikbaar. Daarvoor zullen we te rade moeten gaan bij andere databronnen.

Effecten over tijd

We vinden geen treatment effecten maar wel positieve en significante effecten over tijd in verband met welbevinden en mentale gezondheid maar dan voor alle groepen. Of dit het gevolg is van het experiment, dus een treatment effect, of veroorzaakt wordt door andere factoren, is vooralsnog causaal niet vast te stellen zonder zo'n vergelijkingsgroep. Het is wel opvallend omdat er geen evidentie is dat welbevinden bij bijstandsgerechtigden zou zijn toegenomen in de afgelopen twee jaar, integendeel, cijfers van het SCP suggereren (zie Van Echteld, 2019) dat voor mensen met een levenssatisfactiescore van 6 of lager (waartoe ook de bijstandsgerechtigden behoren) de score de afgelopen jaren is verslechterd.

Negatieve treatment effecten

Een uitkomst die benadrukt moet worden is dat we weinig significante positieve effecten vinden maar ook weinig significante negatieve effecten. Dat betekent dat de andere werkwijze in ieder geval niet slechter werkt op deze uitkomstmaten dan de reguliere begeleiding. Nochtans vinden we twee significante negatieve effecten. Dat is het geval bij vrijwilligerswerk en mantelzorg en institutioneel vertrouwen. Wat het institutionele vertrouwen betreft zijn er negatieve treatment effecten maar die doen zich bij alle treatment groepen voor inclusief de controlegroep. Bij eigen regie is dit effect al bij de nulmeting aanwezig en duidt het op een selectie effect en geen treatment effect. Onduidelijk is vooralsnog wat de achtergrond van deze daling in institutioneel vertrouwen is die zich bij alle groepen voordoet. De effecten blijven overigens hetzelfde in een model zonder klantmanagernummer. Het is dus geen effect van de klantmanager zelf.

Het tweede negatieve effect doet zich voor bij vrijwilligerswerk en mantelzorg en vooral bij de extra begeleidingsgroep en de vrijlatingsgroep in vergelijking met de controlegroep. Nadere analyse laat evenwel zien dat de daling in het aantal uren vrijwilligerswerk en mantelzorg volledig voor rekening komt van de toename in aantal uren betaald werk. De daling is dus een gevolg van de treatment doordat meer mensen vanwege de begeleiding werk vinden maar daardoor minder tijd hebben voor vrijwilligerswerk en mantelzorg. Om die reden is dit effect eigenlijk geen negatief effect maar juist een positief effect van de treatment, althans indien uitstroom naar betaald werk het primaire doel is. Er treedt dus substitutie op

waardoor de grotere arbeidsparticipatie ertoe leidt dat mensen minder tijd hebben voor maatschappelijke participatie.

Resumerend..

We vinden slechts enkele significante treatment effecten zoals bij subjectieve gezondheid, mentale gezondheid en vrijwilligerswerk; anderen worden ogenschijnlijk weerlegd. Toch ontwikkelen een aantal belangrijke uitkomstmaten zich in de tijd gunstig zoals welbevinden en sociaal vertrouwen maar niet beter dan in de controlegroep. Opvallend is dat er reeds bij de start verschillen zijn in de uitgangssituatie hetgeen duidt op zelfselectie. Die doet zich met name voor bij de eigen regie groep en de controle -of onderzoekgroep. We vinden dus vaak geen causaal verband met de treatments met uitzondering van de eerder genoemde. Nader onderzoek moet uitwijzen of onderliggend bijvoorbeeld de stijging in welbevinden, sociaal vertrouwen en mentale gezondheid niet toch een treatment effect is dat niet zichtbaar wordt omdat de reguliere begeleiding in de controlegroep gedurende het experiment is veranderd en is gaan lijken op de andere treatments. We vinden over het algemeen weinig verschillen tussen de treatments. Wel is het zo dat als er verschillen zijn deze zich vooral voordoen bij eigen regie en extra begeleiding maar niet bij de vrijlatingsgroep met uitzondering van het substitutie-effect van betaald werk op uren besteed aan vrijwilligerswerk.

Hierna bespreken we de uitkomsten van de procesevaluatie en analyseren we het uitvoeringsproces zelf aan de hand van de vragenlijsten gesteld aan de klantmanager. Wellicht krijgen we hiermee meer inzicht in de wijze waarop de standaard treatment zich heeft ontwikkeld.

4. Procesevaluatie aan de hand van de procesevaluatie vragenlijsten en de focusgroep gesprekken

4.1. Het profiel van de deelnemers

In de inleiding is beschreven hoe op basis van latente klasse analyse een profiel is gemaakt van de deelnemers die de vragenlijst in de nulmeting hebben ingevuld (n=279). Dat is gebeurd op basis van het zoek -en sollicitatiegedrag van de deelnemer plus de belemmeringen die zij ervaren om werk te vinden. Op grond van die analyse werd vastgesteld dat een aanzienlijk deel van de deelnemers, meer dan 40%, zichzelf arbeidsgeschikt acht en nog actief zoekt naar betaald werk maar dat bijna 60% van de deelnemers zich ongeschikt acht voor werk vanwege gezondheidsproblemen, sociale problemen of belemmeringen om werk te vinden (o.a. te lage kwalificaties) en daardoor zo ontmoedigd zijn geraakt dat ze niet meer actief op zoek zijn naar werk.

In de gemeente Wageningen omvat het experiment meer dan de helft van alle bijstandsgerechtigden en zijn voorts alle klantmanagers betrokken bij het experiment. Zoveel als mogelijk zijn de klantmanagers toegewezen aan een van de treatments en behouden deelnemers dezelfde klantmanager gedurende deelname aan het experiment. Dit is evenwel in de praktijk niet volledig waargemaakt. De onderzoeksgroep is bijvoorbeeld begeleid door 13 verschillende klantmanagers waarbij vier van hen 73% van de deelnemers onder hun hoede hadden en 11 ook andere treatments hebben begeleid. Datzelfde geldt voor de andere treatments. Tegen deze achtergrond bespreken we in deze paragraaf de bevindingen van de procesevaluatie welke zoals gezegd bestaat uit de analyse van de drie procesevaluatievragenlijsten en de analyse van de focusgroep gesprekken met de klantmanagers. Eerst bespreken we de analyse van de procesevaluatievragenlijsten en daarna de focusgroep gesprekken.

4.2. Analyse van de procesevaluatie vragenlijsten

Er zijn drie procesevaluatie vragenlijsten die door de betreffende klantmanager/consulent worden ingevuld voor de deelnemers die hij of zij in begeleiding heeft. Deze noemen we de nulmeting, de eerste vervolgmeting (meting 2) en de tweede vervolgmeting (meting 3). Indien een deelnemer uitstroomt vult de klantmanager de vragenlijst niet meer in. De klantmanager heeft dus voor al zijn of haar deelnemers een of meerdere keren de vragenlijst ingevuld. De resultaten van deze vragenlijsten geven inzicht in het proces van begeleiding vanuit het perspectief van de klantmanager zelf. Er wordt een beeld geschetst van de ontwikkeling die de deelnemer heeft doorgemaakt volgens de klantmanager en in hoeverre er sprake was van een effectieve begeleiding en of er zich leereffecten hebben voorgedaan voor zowel deelnemer als klantmanager of consulent.

4.2.1. Gehanteerde uitkomstmaten

Om een inzicht te krijgen in het proces van begeleiding en de ontwikkeling van de deelnemer door de ogen van de consulent, wordt er naar de volgende aspecten gekeken:

- Hoe beoordeelden de klantmanagers de **motivatie** van de deelnemers voor de treatment en de te behalen doelen en hoe veranderde dit over de tijd?

- Hoe beoordeelden de consultants de deelnemer ten aanzien van een viertal aspecten van de begeleiding: of de deelnemer **realistische eisen** stelt ten aanzien van de (hoogst bereikbare) **doelen**; het **zelf in actie** komen; het **zelf doelen stellen** en het verwachte **effect van de treatment** gedurende het experiment?
- Hoe beoordeelden de consultants de mate van **sociale participatie** en de **afstand tot de arbeidsmarkt** voor de deelnemer (de plaats op de participatieladder) en hoe veranderde dit over de tijd?
- Hoe beoordeelden de consultants de **leereffecten** bij de deelnemer (mate van **zelfredzaamheid**) en bij zichzelf en hoe veranderde dit over de tijd?
- In hoeverre beoordeelden de consultants de geboden treatment als geschikt voor de deelnemer (alleen in de derde vragenlijst bevraagd)?

Uiteraard is een belangrijk doel van de procesevaluatie om te bekijken in hoeverre er volgens de klantmanagers verschillen zijn in uitkomsten tussen deelnemers die verschillende treatments hebben ondergaan. Hierdoor zal duidelijk worden of en welke verschillen er op een bepaald moment zijn in gepercipieerde uitkomsten door de klantmanager tussen de treatments, in hoeverre er een ontwikkeling plaatsvindt over de tijd en in hoeverre deze zijn toe te schrijven aan het effect van de treatment.

4.2.2. Analysemethode vragenlijst

Op basis van de verzamelde data van de procesevaluatie vragenlijsten is er een gewone regressieanalyse (OLS) uitgevoerd, waarbij het verband tussen de uitkomsten van de verschillende treatments en de treatments wordt geanalyseerd. In deze analyse is een aantal onafhankelijke variabelen toegevoegd om de schatting van de uitkomstmaat (afhankelijke variabele) zo zuiver mogelijk te maken, namelijk: de meting waarin de score is bepaald op de uitkomstmaat, of iemand zich wel of niet heeft afgemeld en een dummy variabele die aangeeft of de deelnemer uit het zittende bestand komt, of uit de nieuwe instroom. In de rapportage van de analyses, wordt er gebruik gemaakt van een beschrijvende tabel waarin de gemiddelde scores voor de treatment groepen op een uitkomstvariabele worden getoond per meting en in totaal. Deze gemiddelden zijn geschat op basis van de eerdergenoemde regressieanalyse. Een nadere uitleg over hoe de tabellen gelezen moeten worden, volgt in de volgende paragraaf¹³.

¹³ Er zijn ook regressiemodellen getest met toevoeging van een aantal achtergrondkenmerken. In deze modellen werd er constant gehouden voor de volgende kenmerken van de deelnemer: leeftijd, geslacht, soort huisvesting, leefvorm, duur van uitkering en nationaliteit. Interessante bevindingen uit deze analyses worden steeds bij de resultaten besproken.

4.2.3. Resultaten

In deze paragraaf zullen de resultaten op de vragen uit paragraaf 4.2.1 worden beschreven. De tabellen die worden getoond zijn als volgt te lezen. In de tabellen zijn de gemiddelde scores op de uitkomstvariabele (bijvoorbeeld motivatie, plaats op participatieladder, zelfredzaamheid, etc.) per treatmentgroep en per meting te zien. Er kan dus worden gezegd dat de gemiddelden die hieronder worden weergegeven, de gemiddelden zijn onafhankelijk van (of gecorrigeerd voor) het effect van of iemand is afgemeld en of iemand afkomstig is uit het zittend bestand of nieuwe instroom.¹⁴ Hoe moeten de tabellen worden gelezen? De kolom 'totaal' geeft de gemiddelde score op de 3 metingen voor elke treatmentgroep. De sterretjes geven significantieniveaus weer (zie ook Bijlage B voor bijbehorende geschatte treatmenteffecten). Bij drie sterretjes weten we met 99% betrouwbaarheid dat het verschil tussen de treatments (deels)¹⁵ wordt veroorzaakt door de treatment zelf. Bij twee sterretjes weten we dit met 95% betrouwbaarheid en bij een ster met 90% betrouwbaarheid. Ontbreken de sterretjes, dan wordt het verschil waarschijnlijk door iets anders veroorzaakt dan door het effect van de treatment. De onderste rij met 'Totaal', geeft de gemiddelden over de groepen per meting, dus over de tijd. Sterretjes in deze rij laten zien dat er een significante of betekenisvolle toename of afname plaatsvindt over de tijd ten opzichte van de *nulmeting*. Kijken we naar de resultaten voor de treatmentgroepen in de nulmeting dan kunnen we sterretjes interpreteren als betekenisvolle effecten van de treatments eigen regie, extra begeleiding en vrijlating vergeleken met de *reguliere begeleiding* in die meting. Met eigen regie bedoelen we hier 'zelf in actie', extra begeleiding is 'maatwerk' en vrijlating is 'verdiene loont'. Bij de resultaten voor metingen 2 en 3 geven de sterretjes de effecten (een betekenisvolle toe- of afname) weer van een treatment *bovenop* het effect in de *nulmeting* voor die treatment. Een sterretje bij de onderzoekgroep bijvoorbeeld betekent dan een betekenisvolle toe- of afname in de tijd vergeleken met de score in de nulmeting. Bij de beschrijving van de resultaten zal de nadruk liggen op betekenisvolle verschillen in scores die toe te schrijven zijn aan het *effect van de treatments* althans in de perceptie van de klantmanager. In totaal zijn er 860 vragenlijsten ingevuld: 342 in meting 1, 275 in meting 2 en 243 in meting 3. Echter, in de tabellen die volgen kan het aantal vragenlijsten van dit totaal afwijken omdat er dan missende antwoorden zijn op bepaalde variabelen in de analyse. De totale aantallen staan in de tabellen vermeld.

Motivatie voor eigen treatment

In de vragenlijsten zijn de klantmanagers gevraagd om aan te geven in hoeverre zij de deelnemer gemotiveerd achten om zich in te zetten voor respectievelijk de eigen treatment en de (hoogst bereikbare)

¹⁴ In totaal zijn er voor elke uitkomstvariabele twee regressiemodellen geschat: een zonder interactie tussen de treatments en meting en een met de interacties. De regressie coëfficiënten (effectmetingen) voor de effecten van de treatments en de metingen van het eerste model zijn opgenomen in Bijlage B, tabel 1. De regressie coëfficiënten van het tweede model met interacties zijn te vinden in Bijlage B, tabel 2. Voor het gemak van interpretatie laten we in deze paragraaf de geschatte gemiddelden zien, en niet de volledige uitkomsten van de regressieanalyse. De regressieanalyse (met alle eerdergenoemde variabelen) is wel gebruikt als basis om de gemiddelden in tabel 4.1 t/m 4.10 te schatten.

¹⁵ Waarom 'deels'? Als het effect van een treatment significant is, verklaart dit effect nog niet 100% van de verschillen in gemiddelden tussen groepen. Ook de variabelen afmelder en zittend bestand verklaren een deel en daarnaast zijn er altijd niet-waargenomen factoren die een invloed kunnen uitoefenen.

doelen. De resultaten worden getoond als een rapportcijfer van 0-10, waarbij 0 totaal niet gemotiveerd betekent en 10 heel gemotiveerd.

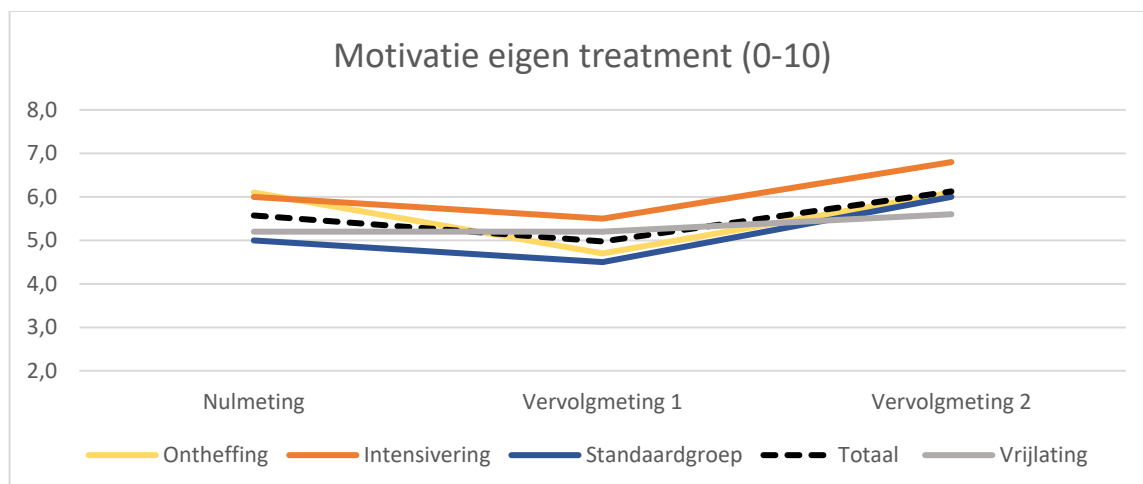
Tabel 4.1: Gemiddelde scores klantmanagers op motivatie deelnemer eigen treatment (0-10)

	Nulmeting		Meting 2		Meting 3		Totaal	
		N		N		N		N
Eigen regie	6,1**	93	4,7	54	6,1*	55	5,6**	202
Extra begeleiding	6,0**	76	5,5	57	6,8	50	6,1**	183
Vrijlating	5,2	83	5,2	82	5,6	70	5,3	235
Onderzoekgroep	5,0	70	4,5	46	6,0*	54	5,2	170
Totaal	5,6	322	5,0**	239	6,1*	229	5,6	790

Noot: *** $p < 0,01$; ** $p < 0,05$; * $p < 0,10$, getest op basis van een OLS regressiemodel (N=790)

Bron: Vragenlijsten nulmeting en vervolgmetingen procesevaluatie Vertrouwensexperiment Wageningen, Tilburg University, 2020.

In tabel 4.1 kunnen we zien dat gemiddeld genomen de motivatie voor de eigen treatment door de klantmanager vrij laag wordt ingeschat met een rapportcijfer van 5,6. Ook zien we (in de kolom 'Totaal') dat de klantmanagers in de onderzoekgroep en de vrijlatingsgroep relatief het laagste scoren op de ingeschatte motivatie van de deelnemer en dat de klantmanagers in de eigen regie groep en extra begeleidingsgroep significant hogere scores geven aan de motivatie van de deelnemer dan in de onderzoekgroep, met de groep extra begeleiding als hoogste score. Als we dan kijken naar de gemiddelde verandering over de tijd (rij 'Totaal'), dan zien we een significante afname in de door de klantmanager veronderstelde motivatie bij de deelnemer in de tweede meting ten opzichte van de nulmeting en een kleine significante toename in de derde meting. Deze resultaten zijn ook terug te vinden in Tabel B1 van Bijlage B. Ook in figuur 4.1 is een tijdelijke dip te zien in de ingeschatte motivatie. Echter, als we dan in tabel 4.1 kijken naar significante verschillen tussen groepen op verschillende meetmomenten, dan zien we juist in de tweede meting geen significante verschillen ten opzichte van de nulmeting die we kunnen toewijzen aan het effect van de treatments. In de nulmeting zien we wel een positief effect bij de eigen regie en extra begeleidingsgroep ten opzichte van de onderzoekgroep. Omdat hier nog nauwelijks begeleiding had plaatsgevonden, zijn deze initiële verschillen wellicht een weerspiegeling van de verwachtingen van de klantmanagers bij de start. Ook als we in ons model controleren voor een uitgebreidere set van achtergrondkenmerken van deelnemers, dan blijven de gevonden effecten bestaan. De verschillen in de nulmeting kunnen dus niet verklaard worden door een verschil in de samenstelling van de deelnemersgroepen. Verder valt op dat de klantmanager de motivatie bij de deelnemer in de onderzoekgroep in de derde meting hoger inschat ten opzichte van de nulmeting, terwijl het omgekeerde (dus een betekenisvol lager ingeschatte motivatie) bij de eigen regie groep (zie Tabel 2, Bijlage B). Hoewel we in tabel 3.2.1 bij extra begeleiding en vrijlating in de derde meting een stijging in ingeschatte motivatie zien, is dit geen treatment effect. Deze stijging heeft een andere oorzaak.



Figuur 4.1. Gemiddelde scores klantmanagers op motivatie deelnemer voor eigen treatment per treatmentgroep en meting

Motivatie voor hoogst bereikbare doel

Daarnaast hebben de klantmanagers ook ingevuld wat hun inschatting is van de motivatie van de deelnemer om zich in te zetten voor de in het werkplan afgesproken hoogst bereikbare doel of doelen, zoals betaald werk, vrijwilligerswerk, het volgen van een opleiding, etc.

Tabel 4.2: Gemiddelde scores klantmanagers op motivatie deelnemer hoogst bereikbare doel (0-10)

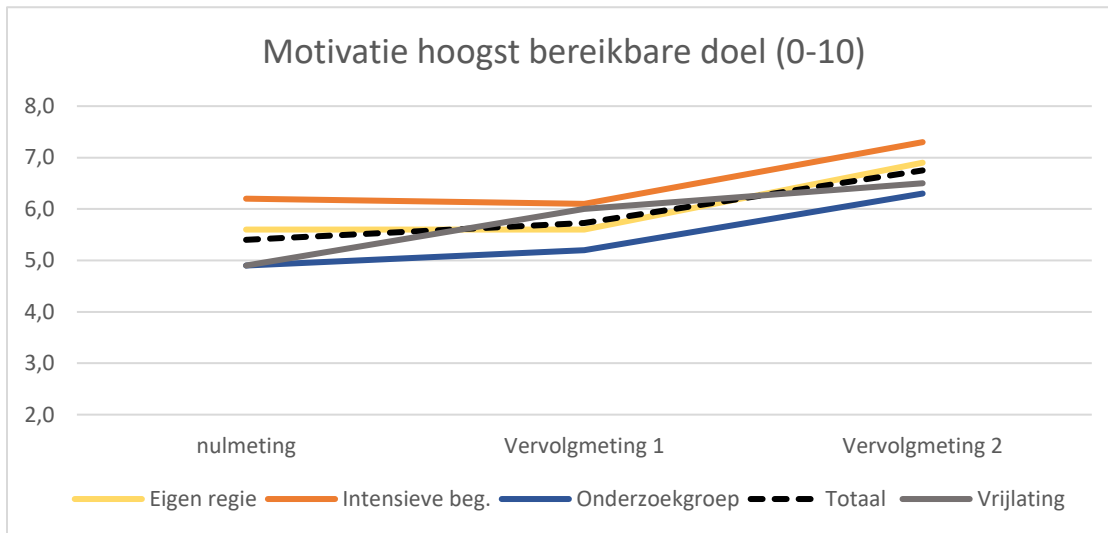
	Nulmeting	N	Meting 2	N	Meting 3	N	Totaal	N
Eigen regie	5,6**	91	5,6	59	6,9	55	6,0**	205
Extra begeleiding	6,2**	74	6,1	59	7,3	52	6,5***	185
Vrijlating	4,9	89	6,0	82	6,5	70	5,8	241
Onderzoekgroep	4,9	69	5,2	51	6,3**	54	5,5	174
Totaal	5,4	323	5,7	251	6,8***	231	6,0	805

Noot: *** $p < 0,01$; ** $p < 0,05$; * $p < 0,10$, getest op basis van een OLS regressiemodel (N=805)

Bron: Vragenlijsten nulmeting en vervolgmetingen procesevaluatie Vertrouwensexperiment Wageningen, Tilburg University, 2020.

In tabel 4.2 zien we dat de gemiddelde score voor de inschatting van de motivatie van de deelnemer voor het hoogst bereikbare doel net voldoende is (6). Ook zien we in de kolom 'totaal' dat volgens de klantmanagers de deelnemers in zowel de eigen regiegroep als de extra begeleidingsgroep significant hoger scoren dan de deelnemers in de onderzoekgroep als gevolg van de treatment. In de rij 'totaal' zien we geen significant verschil tussen de nulmeting en meting 2, maar wel een significante stijging in ingeschatte motivatie in meting 3 ten opzichte van de nulmeting. Deze resultaten zijn ook te zien in tabel 1 van Bijlage B. Als we dan dieper ingaan op de verschillen tussen groepen per meting dan zien we dat in de nulmeting de eigen regie -en extra begeleidingsgroep significant hoger scoort dan de onderzoekgroep. Het is aannemelijk dat de verwachtingen bij de start van het experiment hier een rol speelt. Ten slotte zien we

alleen voor de onderzoeksgroep een significante toename over tijd in de ingeschatte motivatie bij de deelnemer, terwijl er bij de andere treatments geen significante effecten zijn, ondanks dat er wel een stijging zichtbaar is in de ingeschatte motivatie bij de deelnemer in de derde meting (zie ook tabel B2, Bijlage B). Deze stijging is ook duidelijk zichtbaar in figuur 4.2, maar kan dus niet verklaard worden door een treatment effect. Ook het model waarin we controleren voor een uitgebreide lijst van achtergrondkenmerken, levert geen andere resultaten op.



Figuur 4.2. Gemiddelde scores klantmanager op motivatie deelnemer voor hoogst bereikbare doel per treatmentgroep en meting

Realistische eisen

Het derde onderzochte aspect van de begeleiding is hoe realistisch de klantmanagers de eisen inschatten die de deelnemer aan de te behalen doelen stelt. Ook hier hebben de klantmanagers een score gegeven van 0-10 waarin 0 staat voor zeer onrealistisch en 10 voor zeer realistisch.

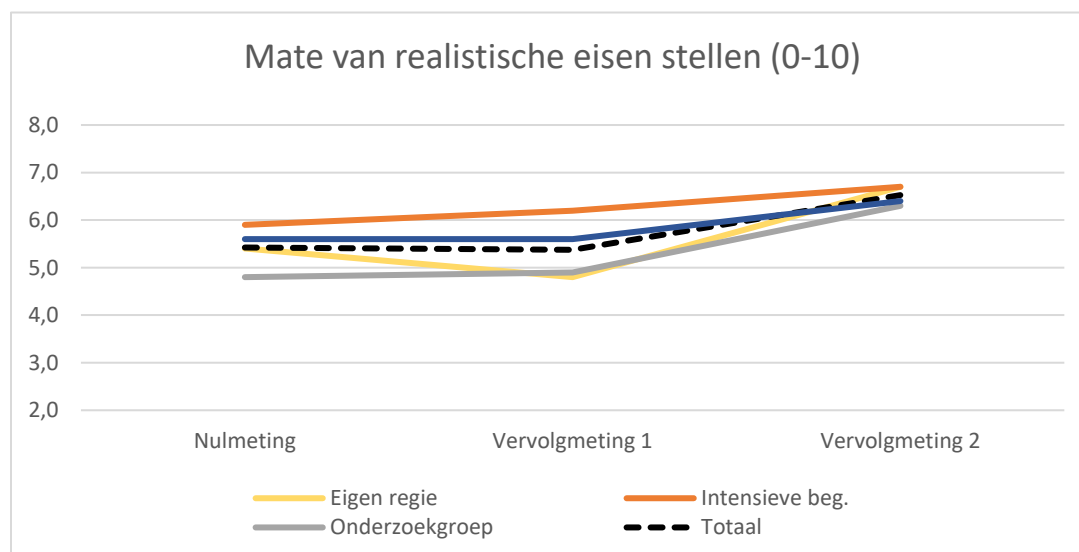
Tabel 4.3: Gemiddelde scores klantmanagers op de mate van realistisch eisen stellen deelnemer (0-10)

	Nulmeting	N	Meting 2	N	Meting 3	N	Totaal	N
Eigen regie	5,4*	93	4,8	57	6,7	50	5,6	200
Extra begeleiding	5,9**	74	6,2	55	6,7	51	6,3***	180
Vrijlating	5,6**	86	5,6	72	6,4	62	5,9**	220
Onderzoeksgroep	4,8	69	4,9	48	6,3**	48	5,3	165
Totaal	5,4	322	5,4	232	6,5***	211	5,8	765

Noot: *** $p < 0,01$; ** $p < 0,05$; * $p < 0,10$, getest op basis van een OLS regressiemodel (N=765)

Bron: Vragenlijsten nulmeting en vervolgmetingen procesevaluatie Vertrouwensexperiment Wageningen, Tilburg University, 2020.

In tabel 4.3 zien we dat de klantmanagers relatief lage rapportcijfers geven voor het stellen van realistische eisen door de deelnemer, gemiddeld een 5,8. Er zijn wel verschillen tussen de treatmentgroepen: zowel de extra begeleidingsgroep als de vrijlatingsgroep scoren significant hoger op dit aspect dan de onderzoeksgroep (kolom 'gemiddelde totaal'). In de onderste rij zien we een significante toename tussen de nulmeting en meting 3 in de mate waarin klantmanagers inschatten dat deelnemers realistisch eisen stellen, terwijl we geen verschil vinden in meting 2. Bij de start van het experiment zijn er al duidelijke verschillen zichtbaar tussen de treatments eigen regie, extra begeleiding en vrijlating enerzijds en de onderzoeksgroep anderzijds. De eerste drie groepen scoren hoger, hoewel het verschil bij de eigen regie groep statistisch gezien zwak is. Wederom zou een verklaring voor deze verschillen de hooggespannen verwachtingen bij aanvang van het experiment kunnen zijn voor het effect van de treatment, want ook als we constant houden voor achtergrondkenmerken van de deelnemers, blijven de gevonden effecten overeind. We vinden voor de metingen daarna bij geen enkele groep een treatment effect behalve in meting 3 bij de onderzoeksgroep (zie ook tabel B2, Bijlage B).



Figuur 4.3. Gemiddelde scores klantmanagers op de mate van realistisch eisen stellen door deelnemer per treatmentgroep en meting

Zelf in actie komen

Een vierde aspect van de begeleiding is de mate waarin de klantmanager de deelnemer in staat acht om zelf in actie te komen om de gestelde doelen te bereiken. Dit zegt iets over de zelfredzaamheid van de deelnemer en de ontwikkeling daarin. In tabel 4.4. worden de resultaten weergegeven. Hier is weer een schaal van 0-10 gebruikt, waarbij 0 staat voor zeer slecht en 10 voor zeer goed.

De gemiddelde score die klantmanagers geven aan de deelnemers voor het zelf in actie komen is laag, slechts een 4,5. We vinden wel significante treatment effecten, namelijk bij de extra begeleidingsgroep en de vrijlatingsgroep, die significant hoger scoren dan de onderzoeksgroep. Over de tijd zien we alleen maar een significante toename van dit aspect van zelfredzaamheid bij de onderzoeksgroep in meting 3 ten

opzichte van de nulmeting maar niet bij de andere treatments. Opvallend is nog dat in de nulmeting het effect van de extra begeleidingsgroep positief is ten opzichte van de onderzoeksgroep. De verwachtingen voor deze groep waren kennelijk hoog bij de start.

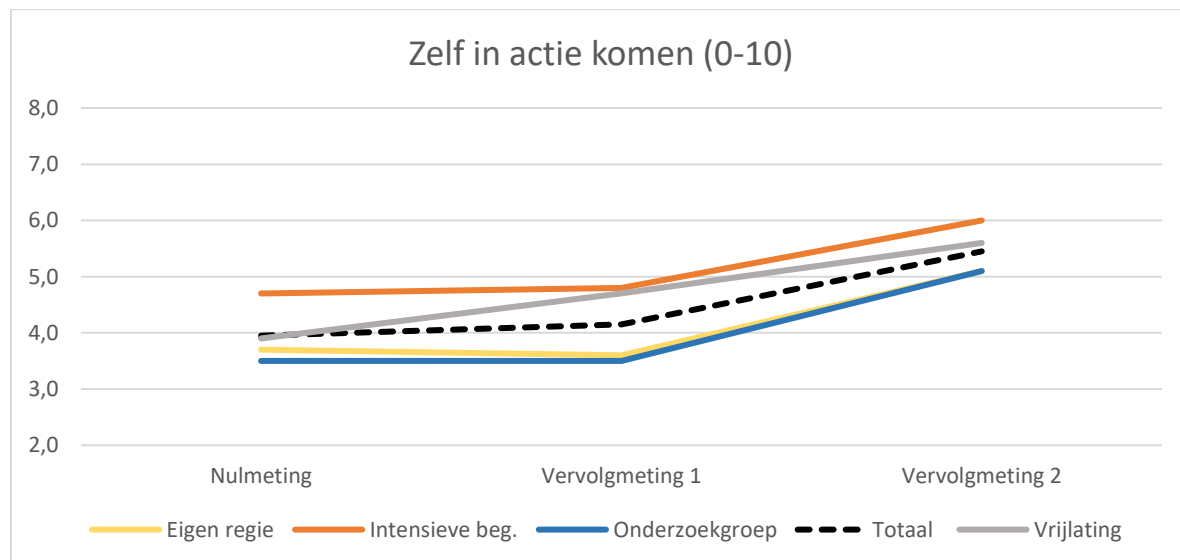
Tabel 4.4: Gemiddelde scores klantmanagers op zelf in actie komen deelnemer (0-10)

	Nulmeting	N	Meting 2	N	Meting 3	N	Totaal	N
Eigen regie	3,7	85	3,6	61	5,1	55	4,1	201
Extra begeleiding	4,7**	73	4,8	58	6,0	51	5,2***	182
Vrijlating	3,9	88	4,7	79	5,6	72	4,7**	239
Onderzoeksgroep	3,5	58	3,5	58	5,1**	53	4,0	169
Totaal	4,0	304	4,2	256	5,5***	231	4,5	791

Noot: *** $p < 0,01$; ** $p < 0,05$; * $p < 0,10$, getest op basis van een OLS regressiemodel (N=791)

Bron: Vragenlijsten nulmeting en vervolgmetingen procesevaluatie Vertrouwensexperiment Wageningen, Tilburg University, 2020.

Voor de andere groepen zien we geen significante verschillen in de verschillende metingen hoewel de klantmanagers wel hogere rapportcijfers geven over de tijd. De toename in de tijd voor de andere groepen, wordt kennelijk door iets anders veroorzaakt dan de treatment zelf. Ook als we controleren voor achtergrondkenmerken, dan veranderen de resultaten niet.



Figuur 3.2.4. Gemiddelde scores klantmanagers over zelf in actie komen bij de deelnemer naar treatmentgroep en meting

Zelf doelen stellen

Een vijfde aspect van de begeleiding is in hoeverre deelnemers zelf in staat zijn doelen te stellen, volgens de klantmanagers. Ook dit aspect zegt iets over de ingeschatte zelfredzaamheid van de deelnemer en de ontwikkeling daarin. Ook hier is weer een score van 0 tot 10 gebruikt, waarbij 0 betekent geheel niet in

staat zelf doelen te stellen en 10 geheel wel. Het gemiddelde rapportcijfer van de klantmanager op dit aspect van zelfredzaamheid is laag: gemiddeld een 5,2. Echter, de onderzoeksgroep scoort het laagst, gemiddeld een 4,5, terwijl alle andere groepen significant hoger scoren, met het hoogste rapportcijfer voor de extra begeleidingsgroep van gemiddeld een 5,9, nog steeds niet erg hoog. De doelgerichte aanpak in Wageningen heeft volgens de klantmanagers dus wel resultaat bij de deelnemers in de drie treatment groepen.

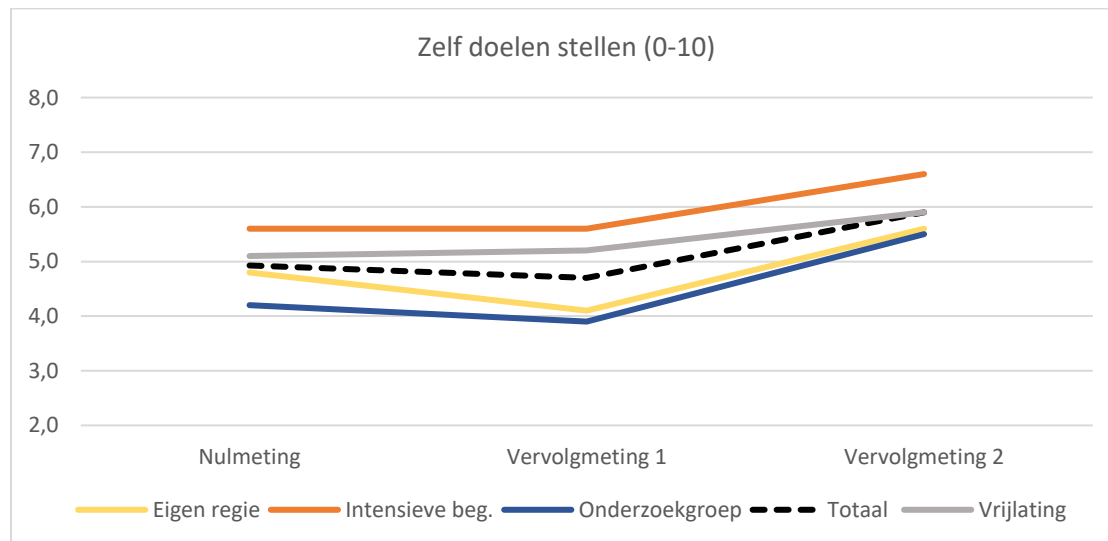
Tabel 4.5: Gemiddelde scores klantmanagers op zelf doelen stellen deelnemer (0-10)

	Nulmeting	N	Meting 2	N	Meting 3	N	Totaal	N
Eigen regie	4,8*	93	4,1	61	5,6	56	4,8*	210
Extra begeleiding	5,6**	76	5,6	57	6,6	51	5,9***	184
Vrijlating	5,1**	91	5,2	81	5,9	71	5,4***	243
Onderzoeksgroep	4,2	72	3,9	59	5,5**	58	4,5	189
Totaal	4,9	332	4,7	258	5,9***	236	5,2	826

Noot: *** $p < 0,01$; ** $p < 0,05$; * $p < 0,10$, getest op basis van een OLS regressiemodel (N=826)

Bron: Vragenlijsten nulmeting en vervolgmetingen procesevaluatie Vertrouwensexperiment Wageningen, Tilburg University, 2020.

Nadere analyse laat zien dat de ingeschatte treatment effecten voornamelijk voorkomen in de nulmeting, wanneer er nog weinig begeleiding heeft plaatsgevonden en dus verwijzen naar de verwachtingen bij de klantmanagers aan het begin. We vinden geen significante effecten in de metingen daarna, behalve bij de onderzoeksgroep zelf in meting 3. Voor de andere groepen zien we wel een significante stijging in de tijd maar geen treatment effecten. Het significante effect in het rij-totaal in meting 3 verwijst dus naar een stijging in de tijd voor alle vier treatments tezamen (zie ook tabel 2, Bijlage B). Als we controleren voor achtergrond kenmerken, dan verdwijnt het (reeds zwakke) effect bij de eigen regie groep.



Figuur 4.5. Gemiddelde scores klantmanagers van zelf doelen stellen door de deelnemer naar treatmentgroep en meting

Verwacht effect van treatment op deelnemer

Ook is de klantmanagers gevraagd hoe groot zij verwachten dat het effect van de treatment is voor de deelnemer op een schaal van 0 tot 10. Hierbij staat 0 voor totaal geen effect en 10 voor een zeer groot effect. De inschatting van het verwachte effect van de treatments bij de deelnemer is laag, gemiddeld slechts een 4,5. Alleen de score van de extra begeleidingsgroep wijkt significant en positief af van de score van de onderzoeksgroep en scoort een punt hoger met een 5,5. Kennelijk verwachten de klantmanagers van de andere twee treatments weinig effect (eigen regie en vrijlating).

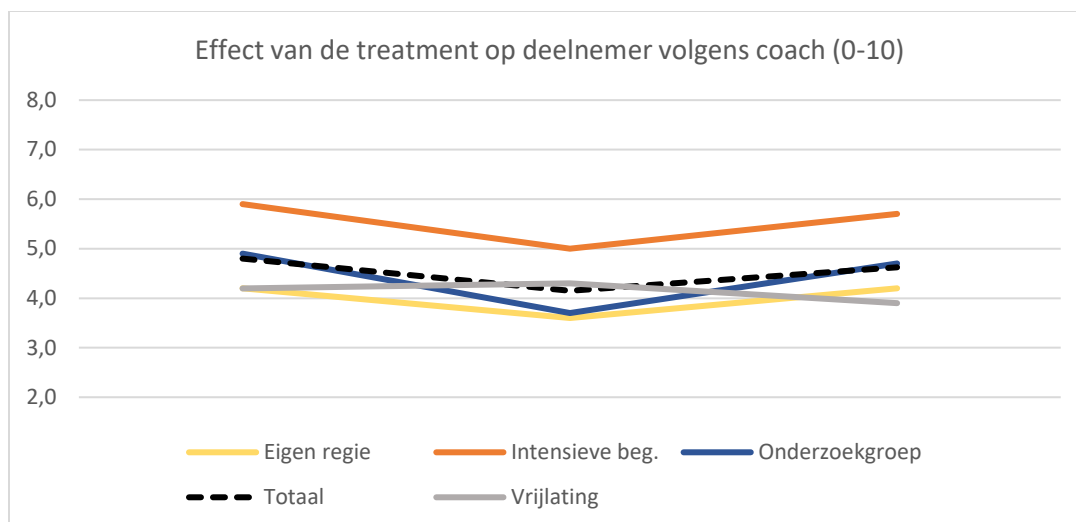
Tabel 4.6: Gemiddelde scores klantmanagers op het verwachte effect van de treatment (0-10)

	Nulmeting	N	Meting 2	N	Meting 3	N	Totaal	N
Eigen regie	4,2	96	3,6	61	4,2	54	4,0	211
Extra begeleiding	5,9**	77	5,0	57	5,7	51	5,5***	185
Vrijlating	4,2*	93	4,3**	82	3,9	73	4,1	248
Onderzoeksgroep	4,9	76	3,7**	58	4,7	52	4,4	186
Totaal	4,8	342	4,2**	258	4,6	230	4,5	830

Noot: *** $p < 0,01$; ** $p < 0,05$; * $p < 0,10$, getest op basis van een OLS regressiemodel (N=830)

Bron: Vragenlijsten nulmeting en vervolgmetingen procesevaluatie Vertrouwensexperiment Wageningen, Tilburg University, 2020.

Echter als we naar de effecten over de tijd kijken zien we dat het verschil tussen de extra begeleidingsgroep en de onderzoeksgroep significant is in de nulmeting maar niet in meting 2 en 3. De verwachtingen voor extra begeleiding waren dus hoog aan het begin. Daarna daalt het verwachte effect in meting 2 om weer te herstellen in meting 3 (zie ook tabel B1, Bijlage B). We vinden geen treatment effect voor de vrijlatingsgroep. Over tijd is er wel een negatief effect van vrijlating in de nulmeting ten opzichte van de onderzoeksgroep in de nulmeting en een positief effect in meting 2 ten opzichte van de onderzoeksgroep in die meting. Als we controleren voor achtergrondkenmerken, dan veranderen de resultaten enigszins. Zo wordt het verschil van de eigen regie groep met de onderzoeksgroep significant negatief en vinden we significante negatieve effecten voor eigen regie en vrijlating in de nulmeting, ten opzichte van de onderzoeksgroep. De gevonden verschillen op dit aspect lijken wederom te maken te hebben met de hoge verwachtingen van de consultants bij de start van het experiment.



Figuur 4.6. Gemiddelde scores klantmanagers van het verwachte effect van de treatment naar treatmentgroep en meting

Sociale participatie

Wat betreft sociale participatie, zijn consultants gevraagd om een inschatting te geven in de mate waarin deelnemers in staat zijn te participeren. De klantmanager is opnieuw gevraagd scores te geven tussen 0 en 10, waar 0 staat voor participeert totaal niet en 10 staat voor participeert totaal wel.

Tabel 4.7: Gemiddelde scores klantmanagers van verwachte effect op sociale participatie (0-10)

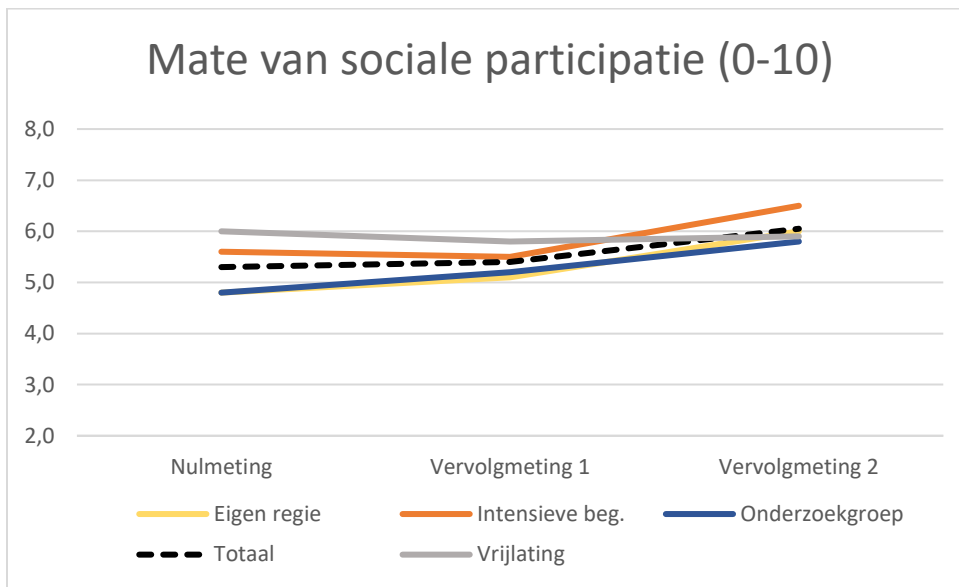
	Nulmeting	N	Meting 2	N	Meting 3	N	Totaal	N
Eigen regie	4,8	95	5,1	57	6,0	43	5,3	195
Extra begeleiding	5,6**	77	5,5	61	6,5	41	5,9**	179
Vrijlating	6,0***	92	5,8	83	5,9*	54	5,9***	229
Onderzoeksgroep	4,8	76	5,2	56	5,8**	44	5,3	176
Totaal	5,3	340	5,4	257	6,1***	182	5,6	779

Noot: *** $p < 0,01$; ** $p < 0,05$; * $p < 0,10$, getest op basis van een OLS regressiemodel (N=779)

Bron: Vragenlijsten nulmeting en vervolgmetingen procesevaluatie Vertrouwensexperiment Wageningen, Tilburg University, 2020.

In tabel 4.7 zien we dat het gemiddelde rapportcijfer door de klantmanager een magere voldoende is, gemiddeld een 5,6. Ook wordt duidelijk dat de deelnemers van de extra begeleidings- en vrijlatingsgroep significant beter scoren volgens de klantmanagers op sociale participatie dan de deelnemers in de onderzoeksgroep. En wederom laat de tabel zien dat er geen significant verschil is tussen de nulmeting en de metingen daarna hetgeen de hoge verwachtingen aan het begin bevestigt. In de onderzoeksgroep krijgen de deelnemers een significant hoger cijfer in meting 3 ten opzichte van de nulmeting (zie ook tabel 1, Bijlage B). Datzelfde geldt voor de vrijlatingsgroep (zie tabel 2, Bijlage B). Daardoor wordt het verschil van de onderzoeksgroep met de andere treatmentgroepen bij de start kleiner in meting 3 (zie Fig. 4.7). In het model

met achtergrond kenmerken verandert er niet veel maar wordt het ingeschatte positieve effect van de standaard treatment in de onderzoeksgroep in meting 3 zwakker.



Figuur 3.2.7. Gemiddelde scores op sociale participatie per treatmentgroep en meting

Participatieladder

Naast een inschatting van de sociale participatie is de consulenten ook gevraagd een score te geven voor de afstand van de deelnemer tot de arbeidsmarkt op een schaal van 1 tot 6. Dit wordt ook wel de participatieladder of tredeladder genoemd, 1 staat hier voor een zeer grote afstand tot de arbeidsmarkt, terwijl 6 staat voor het verrichten van betaald werk. Een score van 5 of 6 duidt op een kleine afstand tot de arbeidsmarkt, een score van 3 of 4 vraagt tussenstapjes en een score lager dan 3 op een grote afstand tot de arbeidsmarkt.

Tabel 4.8: Gemiddelde positie op participatieladder (1-6)

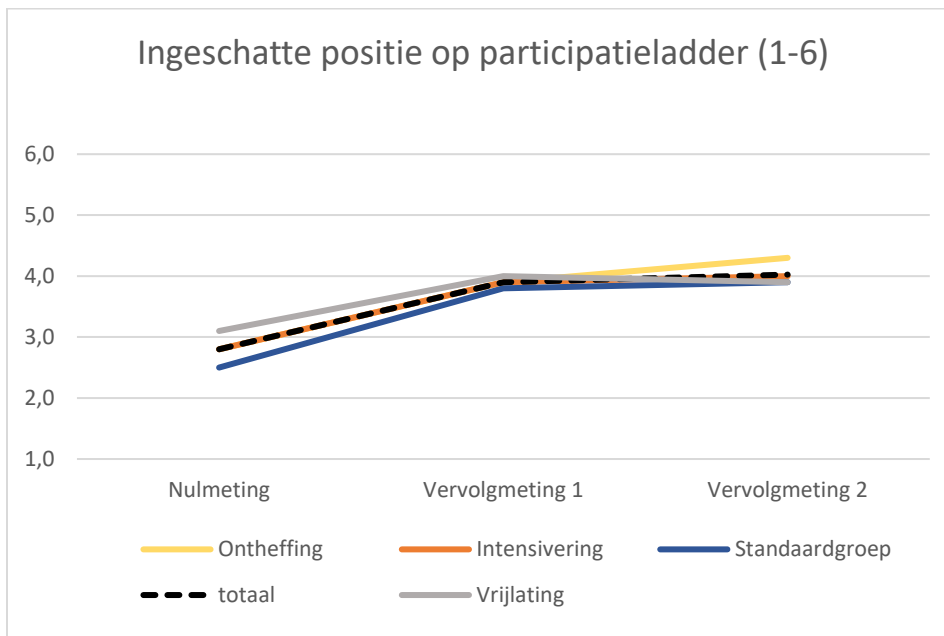
	Nulmeting	N	Meting 2	N	Meting 3	N	Totaal	N
Eigen regie	2,8*	91	3,9	62	4,3	55	3,7**	208
Extra begeleiding	2,8	77	3,9	61	4,0	53	3,6	191
Vrijlating	3,1**	93	4,0	85	3,9*	73	3,7**	251
Onderzoekgroep	2,5	69	3,8***	62	3,9***	59	3,4	190
Totaal	2,8	330	3,9***	270	4,0***	240	3,6	840

Noot: ***p<0,01; **p<0,05; *p<0,10, getest op basis van een OLS regressiemodel (N=840)

Bron: Vragenlijsten nulmeting en vervolgmetingen procesevaluatie Vertrouwensexperiment Wageningen, Tilburg University, 2020.

De gemiddelde score die klantmanagers geven aan de positie van de deelnemers op de participatieladder ligt net onder de 3 bij de start (2,8) maar neemt nadien toe tot bijna 4 in meting 2 en 3,6 in meting 3. Dit is

een significante toename zoals de onderste rij laat zien zowel in meting 2 als 3 ten opzichte van de nulmeting. De eigen regie- en vrijlatingsgroep laten significant hogere scores zien dan de onderzoeksgroep. Voor beiden is dat effect al aanwezig bij de start, kennelijk waren de verwachtingen bij de klantmanagers hoog gespannen. De vrijlatingsgroep heeft dat niet kunnen waarmaken. Het positieve treatment effect aan het begin bij vrijlating wordt in de derde meting tenietgedaan door een negatief effect (zie tabel 2, Bijlage B). Voor de onderzoeksgroep zien we echter een sterke significante en positieve toename in meting 2 en 3, waardoor de verschillen met de andere treatmentgroepen worden verkleind. Dit is ook zichtbaar in figuur 4.8. Als we controleren voor achtergrond kenmerken, dan zien we dat het positieve effect van de eigen regie groep verdwijnt maar dat er een positief treatment effect is bij de extra begeleidingsgroep. Al met al zien we een flinke stijging op de participatieladder voor alle treatments, die na correctie voor achtergrondkenmerken vooral bij de extra begeleidingsgroep en de onderzoeksgroep veroorzaakt wordt door een significant treatment effect. De stijging bij de onderzoeksgroep doet vermoeden dat de treatment niet hetzelfde is gebleven.



Figuur 4.8. Gemiddelde positie op participatieladder per treatmentgroep en meting

Leereffecten

Het begeleiden van de deelnemers met behulp van de treatments is een leerproces voor zowel de deelnemer als voor de klantmanager. Om dit leerproces in kaart te brengen gedurende het experiment, is aan de klantmanagers de vraag gesteld in hoeverre zij vinden dat de deelnemer leert om zelfredzaam te zijn en in hoeverre zij zelf leren van de treatment om zo de deelnemer beter te begeleiden. Bij beide vragen is gevraagd scores te geven van 0 tot 10, waarbij voor het leereffect van de deelnemer 0 staat voor totaal niet zelfredzaam en 10 voor volledig zelfredzaam en voor het leereffect van de klantmanager 0 staat voor niets bijgeleerd en 10 voor heel veel bijgeleerd.

Leereffecten deelnemer: zelfredzaamheid

In tabel 4.9 staan de resultaten voor de ingeschatte leereffecten van de deelnemers door de klantmanagers. Gemiddeld schatten de klantmanagers de leereffecten voor de deelnemers in op gemiddeld een 4,9, wat in termen van een rapportcijfer onvoldoende is. Nochtans is sprake van een significante stijging in meting 3 (5,4) ten opzichte van de nulmeting (4,8).

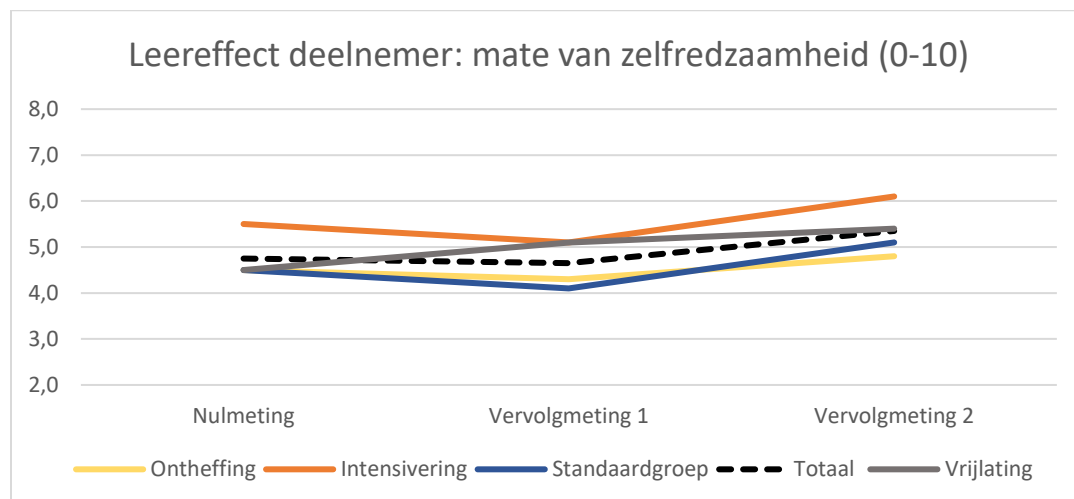
Tabel 4.9: Gemiddelde leereffecten deelnemer (mate van zelfredzaamheid) (0-10)

	Nulmeting	N	Meting 2	N	Meting 3	N	Totaal	N
Eigen regie	4,5	96	4,3	61	4,8	54	4,5	211
Extra begeleiding	5,5**	77	5,1	56	6,1	49	5,6***	182
Vrijlating	4,5	93	5,1**	77	5,4	73	5,0**	243
Onderzoeksgroep	4,5	76	4,1	53	5,1	55	4,6	184
Totaal	4,8	342	4,7	247	5,4**	231	4,9	820

Noot: *** $p < 0,01$; ** $p < 0,05$; * $p < 0,10$, getest op basis van een OLS regressiemodel (N=820)

Bron: Vragenlijsten nulmeting en vervolgmetingen procesevaluatie Vertrouwensexperiment Wageningen, Tilburg University, 2020.

Tevens vinden we significante treatment effecten bij de extra begeleidingsgroep en de vrijlatingsgroep die gemiddeld met een 5,6 en 5,0 een half tot 1 punt hoger scoren op het leereffect bij de deelnemer van zelfredzaamheid dan de onderzoeksgroep. Het positieve effect van de extra begeleidingsgroep ten opzichte van de onderzoeksgroep was er al in de nulmeting hetgeen de hoge verwachtingen bij de klantmanagers van deze treatment bij de start toont. Het positieve treatment effect volgens de klantmanagers bij de groep vrijlating is er al in meting 2. Het toevoegen van achtergrondkenmerken aan het model levert geen andere uitkomsten op.



Figuur 4.9. Gemiddelde ingeschatte mate van zelfredzaamheid (leereffect deelnemer) per treatmentgroep en meting

Leereffecten klantmanager

Vervolgens kijken we naar het ingeschatte leereffect van de klantmanager bij hem of haarzelf. In tabel 4.10 zien we dat het gemiddelde leereffect voor de klantmanager zelf laag wordt ingeschat, gemiddeld op 3,3 maar wel significant hoger in alle drie treatment groepen (3,6) dan in de onderzoeksgroep (2,4). De klantmanagers zelf geven wel aan dat het lastig is deze vraag te beantwoorden voor elke deelnemer apart maar dat het leereffect vooral ontstaat in de gezamenlijke ervaring met alle deelnemers. Dit aspect is daarom ook besproken in de focusgroep gesprekken.

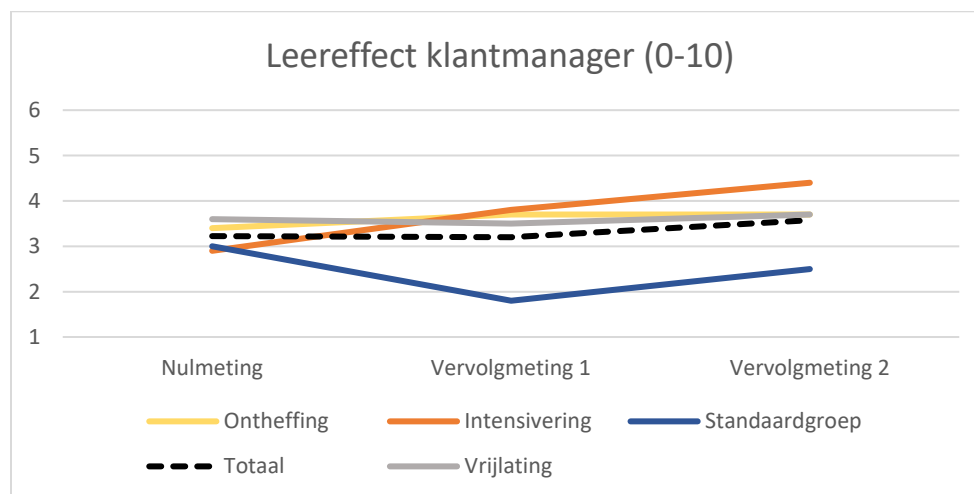
Tabel 4.10: Gemiddelde score op leereffecten bij de klantmanager (0-10)

	Nulmeting	N	Meting 2	N	Meting 3	N	Totaal	N
Eigen regie	3,4	96	3,7**	62	3,7	56	3,6***	214
Extra begeleiding	2,9	77	3,8***	61	4,4**	53	3,7***	191
Vrijlating	3,6	93	3,5*	86	3,7	75	3,6***	254
Onderzoeksgroep	3	76	1,8**	62	2,5	59	2,4	197
Totaal	3,2	342	3,2	271	3,6	243	3,3	856

Noot: ***p<0,01; **p<0,05; *p<0,10, getest op basis van een multiple regressieanalyse (N=856)

Bron: Vragenlijsten nulmeting en vervolgmetingen procesevaluatie Vertrouwensexperiment Wageningen, Tilburg University, 2020.

De onderzoeksgroep scoort dus duidelijk het laagst. Over de tijd zien we bij alle drie treatment groepen positieve significante treatment effecten, ofwel ingeschatte leereffecten, reeds in meting 2 (zie ook tabel 1, Bijlage B). Daarentegen zien we een significant negatief treatment of ingeschat leereffect bij de onderzoeksgroep in meting 2. Ondanks de iets lagere score voor vrijlating in meting 2 is het treatment effect (zwak) positief. Ten slotte zien we een positief effect van extra begeleiding ook in meting 3. De klantmanagers van de extra begeleidingsgroep lijken dus het leereffect naarmate het experiment vordert hoger in te schatten, terwijl dit voor de onderzoeksgroep juist lager wordt. De ingeschatte leereffecten bij de onderzoeksgroep blijven duidelijk achter bij de andere treatmentgroepen (zie figuur 4.10). Het model met achtergrond kenmerken geeft geen andere uitkomsten.



Figuur 3.2.10. Gemiddelde score op leereffect van klantmanager per treatmentgroep en meting

Vergelijking eigen regie en extra begeleidingsgroep

In voorgaande tabellen is niet beschreven of de verschillen tussen de ingeschatte effecten voor de deelnemer bij de groepen eigen regie, extra begeleiding en vrijlating betekenisvol konden worden toegeschreven aan het effect van de treatment, omdat we primair steeds een vergelijking hebben gemaakt tussen deze drie groepen en de onderzoeksgroep. Echter, om meer inzicht te bieden in de verschillen tussen deze drie groepen, is een kort overzicht gemaakt in onderstaande tabel. De resultaten zijn gebaseerd op de OLS regressieanalyses, waarbij dezelfde controlevariabelen zijn gebruikt als in de tabellen B1 en B2 in Bijlage B. Het gaat hier om de kolom 'Totaal' in de tabellen 4.1 t/m 4.10.

In tabel 4.11 is ten eerste een vergelijking gemaakt tussen de extra begeleidingsgroep en de eigen regiegroep. De eigen regie groep werd geacht meer eigen regie te pakken en minder begeleiding te krijgen, terwijl de extra begeleidingsgroep juist een intensievere vorm van begeleiding heeft gehad. Uit de resultaten blijkt dat dit op 7 van de 10 uitkomstmaten in de visie van de klantmanager positieve resultaten heeft voor de extra begeleidingsgroep. Met name scoort deze groep volgens de klantmanager hoger op de aspecten die de zelfredzaamheid van de deelnemer weerspiegelen zoals het zelf in actie komen en het zelf doelen stellen, maar ook het verwachte effect van de treatment. Vergelijken we de extra begeleidingsgroep met de vrijlatingsgroep, waarbij de eerste groep dus intensieve begeleiding op maat kreeg en de tweede groep een ruimere inkomensvrijlating, dan zien we dat bij de helft van de uitkomstmaten de ingeschatte effecten van intensieve begeleiding voor de deelnemer positiever worden ingeschat dan de effecten voor de deelnemer bij de vrijlatingsgroep. Met name het effect van de treatment op de motivatie voor de eigen treatment en het hoogst bereikbare doel en het verwachte effect van de treatment scoren volgens de klantmanagers in de extra begeleidingsgroep hoger.

Tabel 4.11: Verschillen in scores op de uitkomstmaten van de procesevaluatie tussen eigen regie, extra begeleiding en vrijlating (niet-gestandaardiseerde regressie coëfficiënten)

Uitkomstmaten	Extra begeleiding vs. eigen regie	Extra begeleiding vs. vrijlating	Eigen regie vs. vrijlating
Motivatie eigen treatment	0	0,8**	0,5*
Motivatie hoogst bereikbare doel	0,5*	0,7**	0
Realistische eisen	0,6**	0	0
Zelf in actie komen	1***	0	-0,6**
Zelf doelen stellen	1**	0,5*	-0,5**
Verwacht effect treatment	1,5***	1,4***	0
Sociale participatie	0,6**	0	-0,7***
Participatieladder	0	0	0
Leereffect deelnemer	0,9***	0,6**	-0,4*
Leereffect consulent	0	0	0

Noot: *** $p < 0,01$; ** $p < 0,05$; * $p < 0,10$; 0 (niet significant)

Bron: Vragenlijsten nulmeting en vervolgmetingen procesevaluatie Vertrouwensexperiment Wageningen, Tilburg University, 2020.

Ten slotte hebben we de eigen regie groep vergeleken met de vrijlatingsgroep. In die vergelijking vinden we op 5 uitkomstmaten significante verschillen tussen de groepen, waarvan 4 in het nadeel zijn van de eigen regiegroep. Alleen op de motivatie voor de eigen treatment scoort de eigen regie groep volgens de klantmanagers hoger. De grootste negatieve verschillen voor de eigen regie groep vinden we bij het ingeschatte effect op sociale participatie en het zelf in actie komen van de deelnemer.

Geschiktheid treatment

Ten slotte gaan we in op een vraag die in laatste meting is gesteld aan de klantmanagers over de treatments. In deze meting is gevraagd in hoeverre de consultants de treatment geschikt vonden voor deze klant, met een score 0 voor geheel ongeschikt en 10 voor totaal geschikt.

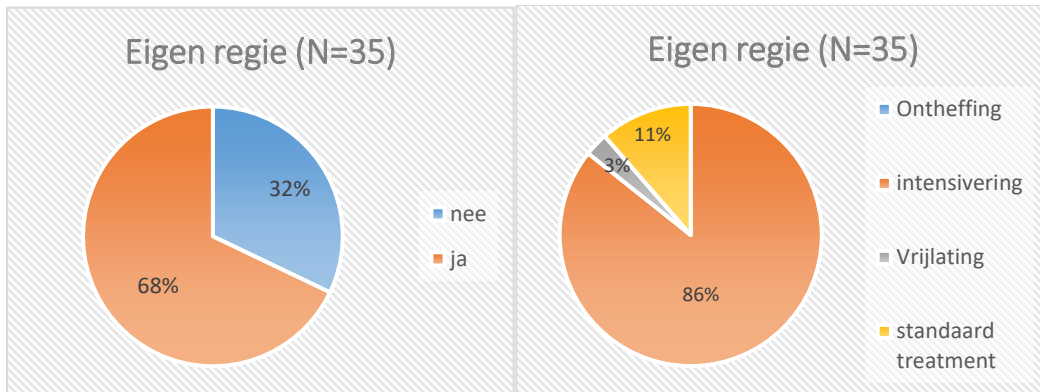
Tabel 3.2.12: Rapportcijfer over geschiktheid treatment voor deelnemer

Treatment geschikt (0-10)?	Gemiddelde score	N
Eigen regie	4	53
Extra begeleiding	7,3	50
Vrijlating	3,8	69
Onderzoekgroep	5,1	49
Totaal	4,9	221

Bron: Vragenlijst tweede vervolgmeting procesevaluatie Vertrouwensexperiment Wageningen, Tilburg University, 2020.

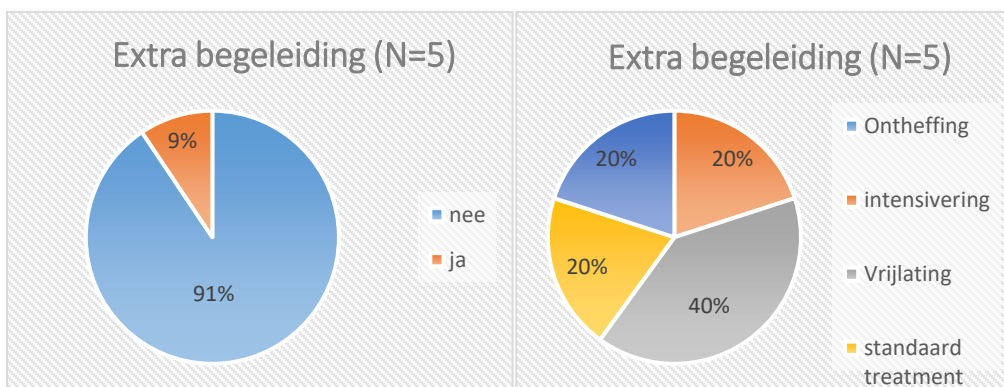
Uit deze scores wordt duidelijk dat de treatment extra begeleiding de hoogste score krijgt, daarna de onderzoekgroep en het laagste scoren de eigen regie- en vrijlatingsgroep.

Vervolgens is na deze vraag de vraag gesteld of de klantmanagers vinden dat een andere treatment meer passend was geweest voor de deelnemer. Hieronder staan de resultaten samengevat in een aantal grafieken. In figuur 4.11 zien we dat in 68% van de gevallen de klantmanagers van de eigen regie groep zou hebben gekozen voor een andere treatment. Dit komt overeen met het eerder vermelde lage geschiktheidscijfer voor eigen regie. Daarna zien we in de figuur rechts voor welke andere treatment de klantmanagers dan zouden hebben gekozen. In de overgrote meerderheid van de gevallen (86%) zou er zijn gekozen voor de treatment van extra begeleiding, in 11% voor de reguliere treatment en in 3% van de gevallen voor vrijlating.



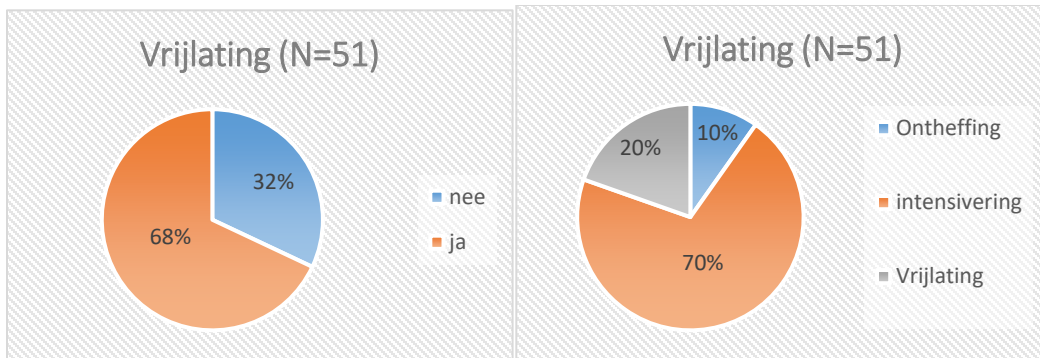
Figuur 4.11. *Andere treatment meer passend? Percentage ja/nee voor treatment eigen regie (links). En: Indien andere treatment meer geschikt, welke is dan meer geschikt? (rechts)*

Bij de treatment extra begeleiding blijkt juist dat slechts in 9% van de gevallen gekozen zou zijn voor een andere treatment. Voor welke treatment ze dan zouden hebben gekozen geeft een minder eenduidig beeld dan voor de eigen regiegroep. De percentages zeggen echter niet zoveel omdat ze zijn berekend op zeer kleine aantallen (N=5).



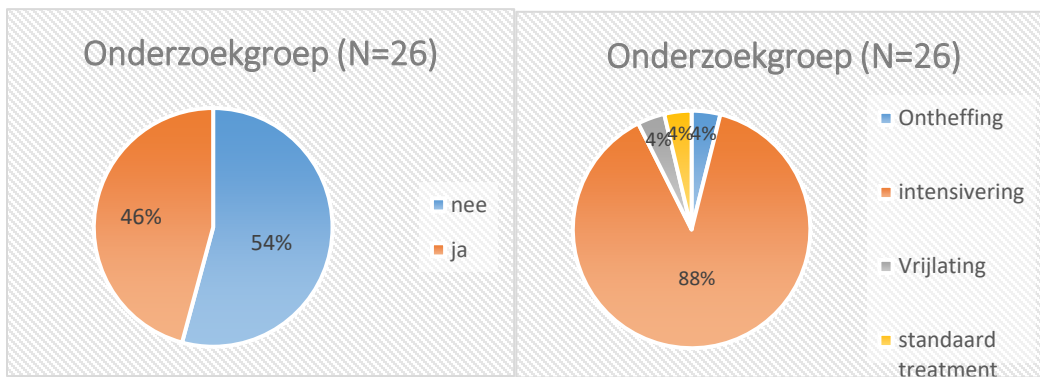
Figuur 4.12. *Andere treatment meer passend? Percentage ja/nee voor treatment extra begeleiding (links). En: Indien andere treatment meer geschikt, welke is dan meer geschikt (rechts)*

Kijken we naar de vrijlatingsgroep in figuur 4.13 dan blijkt dat 68% kiest voor een andere treatment die volgens de klantmanager meer passend zou zijn geweest. Ook dit correspondeert weer met het eerder gevonden geschiktheidscijfer. Van deze 51 gevallen waarvoor een andere treatment meer geschikt zou zijn, zou voor 70% gekozen zijn voor extra begeleiding (intensivering), 20% voor vrijlating en 10% voor eigen regie (ontheffing). Ook hier blijkt, net als bij de eigen regiegroep een grote voorkeur te zijn voor extra begeleiding.



Figuur 4.13. *Andere treatment meer passend? Percentage ja/nee voor treatment intensieve begeleiding met vrijlating (links). Indien andere treatment meer geschikt, welke is dan meer geschikt? (rechts)*

Ten slotte zien we in figuur 4.14 de uitkomsten voor de onderzoekgroep. Hier blijkt dat in iets minder dan de helft van de gevallen een andere treatment meer geschikt zou zijn geweest. En wederom zien we bij de keuze voor een meer geschikte treatment de treatment intensivering het meest gekozen wordt (88%). Slechts in steeds 4% van de gevallen zou er respectievelijk gekozen zijn voor vrijlating, standaardtreatment (wat opvallend is) en ontheffing.



Figuur 4.14. *Andere treatment meer passend? Percentage ja/nee voor treatment standaard groep (links). En: Indien andere treatment meer geschikt, welke is dan meer geschikt? (rechts)*

Al met al kan er geconcludeerd worden dat de treatment extra begeleiding in alle andere groepen het hoogste scoort op de ingeschatte geschiktheid van de treatment voor de deelnemer. De klantmanagers van de treatments eigen regie en vrijlating zouden in bijna 70% van de gevallen gekozen hebben voor een andere treatment, en dan met name de treatment extra begeleiding. Bij de onderzoekgroep zou er iets minder vaak gekozen zijn voor een andere treatment, maar als er dan wordt gekozen, is dit ook hoofdzakelijk weer voor extra begeleiding.

4.2.4. Conclusies

In paragraaf 4 zijn de resultaten besproken van de analyses op de procesevaluatievragenlijsten op basis van een aantal uitkomstmaten. De verwachting was dat de ingeschatte uitkomsten van de treatmentgroepen eigen regie, extra begeleiding en vrijlating door de klantmanagers in die groepen andere en op sommige

aspecten zoals zelfredzaamheid hogere scores zouden laten zien op de uitkomstmaten dan de onderzoeksgroep, vanwege de andere werkwijze die deel uit maakte van deze treatments. De analyses laten zien dat vooral de uitkomsten bij de extra begeleidingsgroep in de visie van de klantmanagers positief afwijken van de uitkomsten bij de onderzoeksgroep op de besproken uitkomstmaten. Op 9 van de 10 uitkomstmaten scoort de extra begeleidingsgroep (sterk) significant hoger dan de onderzoeksgroep, alleen bij de score op de participatieladder kon er geen verschil worden gevonden (wel na correctie voor verschillen in achtergrondkenmerken). Met name bij het aspect van zelfredzaamheid, het zelf doelen stellen, komt de extra begeleidingsgroep beter uit de bus, omdat daar het treatmenteffect het grootst is. Vervolgens vonden we ook voor de vrijlatingsgroep op 7 van de 10 uitkomstmaten een significant positieve effect ten opzichte van de onderzoeksgroep, met name op het zelf doelen stellen en het leereffect voor de klantmanager. Voor de eigen regie groep vonden we op 5 van de 10 uitkomstmaten een positief effect, hoewel deze minder sterk significant en minder groot waren dan bij de andere twee treatmentgroepen. Voornamelijk het leereffect bij de klantmanager blijkt positief af te wijken van dat van de onderzoeksgroep. Een kanttekening bij deze bevindingen is wel dat nadere analyses getoond hebben dat veel verschillen tussen de treatments reeds worden gevonden in de nulmeting en in onze visie deels verwijzen naar de hoge verwachtingen die klantmanagers hadden aan het begin van het experiment. In de vervolgmetingen werden vaak geen significante treatment effecten gevonden. Opvallend was dat we voor de onderzoeksgroep in meting 3 soms positieve treatmenteffecten vonden die suggereren dat de treatment gedurende het experiment is veranderd. We hebben getest of deze initiële verschillen verklaard konden worden door compositieverschillen in de samenstelling van de deelnemers groepen, maar dat bleek in de meeste gevallen niet zo te zijn. De precieze reden van deze verschillen in de nulmeting is op basis van de huidige analyses niet met zekerheid te geven.

In meting 2 bleken een aantal effecten significant negatief te worden, namelijk bij de ingeschatte motivatie voor de eigen treatment en het verwachte effect van de treatment. Blijkbaar zijn de hoge verwachtingen aan het begin sterk getemperd gedurende het eerste jaar waardoor er een soort dip is ontstaan en de scores door de klantmanagers naar beneden zijn bijgesteld. Opvallend is dat dit met name geldt voor de onderzoeksgroep. De vrijlatingsgroep laat juist in de tweede meting een stijging zien op het verwachte effect van de treatment en de leereffecten. In de derde meting vond er wel in de meeste gevallen een stijging plaats ten opzichte van de nulmeting, maar vaak was deze alleen voor de onderzoeksgroep toe te schrijven aan een positief effect van de treatment. Ten slotte hebben we de drie treatmentgroepen eigen regie, extra begeleiding en vrijlating onderling vergeleken en geconcludeerd dat op veel punten de treatment extra begeleiding het in de ogen van de klantmanagers beter doet dan eigen regie en vrijlating en dat vrijlating het weer beter doet dan eigen regie. Ook uit de mening over de geschiktheid van de treatment voor de deelnemer blijkt dat de klantmanagers extra begeleiding het hoogste cijfer geven. Indien zij bij de overige drie treatments zouden kiezen voor een andere treatment, dan zou dit in 70-88% van de gevallen ook extra begeleiding zijn. Gezien de ingeschatte effecten door de klantmanagers op de besproken uitkomstmaten is dit ook geen verrassing.

Ten slotte hebben we bij de bespreking van de resultaten ook aangegeven dat we modellen hebben getest waaraan een aantal achtergrondkenmerken van de deelnemers zijn toegevoegd. Hieruit bleek dat de gevonden resultaten over het algemeen niet of zeer weinig veranderen als we controleren voor meer

achtergrond kenmerken van de deelnemers. Daaruit bleek wel dat deelnemers met kortere verblijfsduren in de bijstand op de gekozen uitkomstmaten hoger scoren in de ogen van de klantmanagers dan mensen met langere duren, afmelders lager scoren, vrouwen hoger dan mannen, mensen met een buitenlandse nationaliteit lager dan mensen met een Nederlandse nationaliteit, ouderen lager dan jongeren en echtparen soms lager dan alleenstaanden.

4.3 Focusgroep gesprekken

4.3.1 Inleiding

Het eerste focusgroep gesprek van Wageningen vond plaats op 23 januari 2018, bijna 4 maanden na de start van het experiment op 1 oktober 2017. Het eerste gesprek vormde als het ware de nulmeting van de procesevaluatie maar dan zijn de eerste gesprekken met de deelnemers reeds gevoerd en de eerste ervaringen verwerkt. Het tweede gesprek vond bijna 8 maanden later plaats op 17 september 2018. Het derde gesprek vond weer 6 maanden nadien plaats op 21 maart 2019. Dat was tevens het laatste gesprek. Nadien zijn door de projectleider exit gesprekken gevoerd met de klantmanagers die zijn gedeeld met de onderzoekers en is er nog een korte vragenlijst ingevuld door de klantmanagers na afloop. Op 23 januari 2018 werden de eerste ervaringen besproken. Op 17 september 2018 werden in het bijzonder de ervaren problemen en knelpunten aan de hand van de casuïstiek besproken en hoe en in welke mate deze al dan niet konden worden opgelost. Op 21 maart 2019 werden de tot dan toe opgedane ervaringen besproken en vooruit gekeken naar de activiteiten in het laatste half jaar van het project. De gesprekken zijn in audio opgenomen en er is van elke gesprek een transcript gemaakt. Deze transcripten zijn beluisterd door twee onderzoekers op basis waarvan een uitgebreid verslag is gemaakt. Een beknopte samenvatting van dit verslag is in de laatste tussenrapportage van april 2019 opgenomen. In dit eindverslag zijn de bevindingen uit deze laatste tussenrapportage grotendeels overgenomen.

Onderwerpen

In de focusgroep gesprekken wordt besproken wat de kansen en eventuele belemmeringen zijn van de andere bejegening van de deelnemer en welke resultaten men verwacht en al dan niet bereikt. In het bijzonder wordt ingegaan op de handelwijzen van de coach in de verschillende interventies en in welk opzicht ze van elkaar dienen te verschillen. Betekent 'eigen regie' dat men de deelnemer volledig moet loslaten of juist niet omdat het 'leren van zelfredzaamheid' meer contacten met de deelnemer vraagt. Ten tweede wordt ingegaan op de casuïstiek van de begeleiding waarbij specifieke problemen in de begeleiding zoals fraude, schulden en ongemotiveerd of onaangepast gedrag aan de orde wordt gesteld om ervan te leren. Ten derde wordt ingegaan op de lokale context waarbinnen het VEX-experiment plaatsvindt, dat is de gemeentelijke organisatie, de politiek, de media en alle betrokkenen daaromheen. Om deze redenen wordt in de gesprekken ook ingegaan op de bijzondere opzet en werkwijze in de uitvoering van het Wageningse experiment en de dilemma's die mogelijk hieruit zijn voortgekomen. Het bijzondere in Wageningen is de methodiek van het gezamenlijk met de deelnemer opstellen van een werkplan met overeengekomen ('hoogst bereikbare') doelen, de focus op bevordering deeltijdwerk alsmede de schaal van het experiment waarbij de helft van de bijstandsgerechtigden in Wageningen deelneemt aan het experiment. De focusgroep gesprekken vinden plaats in het bijzijn van de projectleider die het experiment

uitvoeringstechnisch aanstuurt en coördineert. In de drie gesprekken werden de ervaringen besproken alsook de ervaren problemen en knelpunten in de coaching van de deelnemers in de verschillende treatments maar ook in de uitvoering binnen de gemeentelijke context. Wat coaching betreft werd een sterke behoefte gevoeld om casuïstiek te bespreken waarin men belangrijke dilemma's had ervaren en hoe deze al dan niet zijn opgelost.

Opzet van de focusgroep gesprekken

De focusgroep gesprekken vonden plaats op basis van een half-gestructureerde vragenlijst die in een PowerPoint presentatie werd toegelicht door de onderzoekers. Per onderwerp werden een aantal vragen gesteld die de input vormde voor het gesprek. Het proces van experimenteren in de PW is in de vragenlijst onderverdeeld in 5 stappen. Bij elke stap hebben de onderzoekers vragen geformuleerd die de leidraad vormden bij het gesprek. De volgende onderwerpen zijn aan de orde gesteld:

Stap 1. De beleidsuitgangspunten en het doel van het experiment.

Stap 2. Achterliggende theorie/filosofie en wetenschappelijke vragen.

Stap 3. Van theorie naar RCT-design voor een niet-AMvB experiment.

Stap 4. Van design naar implementatie: de drie treatment groepen; de relatie met de cliënt; de eigen rol als coach etc.

Stap 5. Van implementatie naar uitkomsten: de lokale en politieke context, de randvoorwaarden en de uitkomsten voor de cliënt.

Algemene ervaringen

Nagenoeg alle klantmanagers waren bij de drie gesprekken aanwezig. Klantmanagers zijn in Wageningen in beginsel verantwoordelijk voor een treatment maar soms moet daarvan worden afgeweken. Dat vindt men over het algemeen lastig. Men komt wekelijks bij elkaar om over de invulling van de treatments te spreken en met elkaar af te stemmen. De treatments zijn qua inhoud en aantal contacten bij de start van het experiment op papier gezet. Men probeert zich daar zo goed mogelijk aan te houden maar in de praktijk wordt daar soms van afgeweken in verband met de casuïstiek waardoor om uiteenlopende praktische redenen anders wordt gehandeld. Zo vindt men bijvoorbeeld het aantal van 6 contacten bij de extra begeleidingsgroep te weinig om goed contact met de deelnemer te kunnen houden. Daarnaast was bij de eigen regiegroep zeker in het begin (veel) meer contact nodig dan was afgesproken (2 keer per jaar). De belangrijkste vraag die in de gesprekken aan de orde is gekomen is welke lessen er zijn getrokken uit de praktijk van de begeleiding en wat is er geleerd van het proces van experimenteren en de 'nieuwe' aanpak tot nu toe bij de klantmanagers?

Sfeer in de gesprekken

De drie gesprekken verliepen in een prettige en geanimeerde sfeer. Er bestond duidelijk behoefte om de ervaringen met de onderzoekers te delen en vragen te kunnen stellen aan de onderzoekers. Men dacht ook duidelijk na over hoe de handelwijze van henzelf van invloed kan zijn op de uitkomsten van het onderzoek. Men was zich zeer goed bewust van het gevaar dat de treatments te veel op elkaar gaan lijken waardoor het onderzoek niet meer goed de verschillen in beeld kan brengen. In alle drie gesprekken kwam de caseload ter sprake die men te hoog vindt om de treatments goed te kunnen uitvoeren. Daardoor ontbreekt de tijd om aandacht aan de deelnemers te geven die nodig is om hen goed van dienst te zijn. Men zou ook

graag meer ondersteuning zien vanuit de organisatie voor het experiment. Men stelde gaande het experiment ook vragen over of de treatments zoals uitgevoerd nog wel stroken met de handelwijze zoals deze aan het begin op papier is gesteld. Een andere vraag die naar voren kwam is ook of de treatments bij iedereen zullen werken, iedereen heeft individuele aanpak nodig, en hoe men moet handelen bij ongewenst gedrag, fraude, gebrekkige motivatie of bij een persoonlijke problematiek. Hoe ga je als klantmanager om met deze zaken, hoe streng of hoe soepel moet en mag je zijn, en hoe ga je om met persoonlijke problemen/verhalen van deelnemers. In het kort worden de bevindingen van de gesprekken hieronder nog eens puntsgewijs weergegeven.

4.3.2 Focusgroep gesprekken: het proces van experimenteren

Verkenningfase

In het eerste half jaar van het experiment hebben de klantmanagers het experiment praktisch moeten vormgeven in de spreekkamer. In Wageningen hebben ze als team die uitdaging op zich genomen onder strakke begeleiding van de projectleider. Veel vragen die men had zijn in het wekelijkse overleg maar ook in de focusgroep gesprekken aan de orde gekomen. Hieronder geven we eerst een verslag van de bevindingen van de focusgroep gesprekken in de verkenningfase aangezien het uitvoeren van het experiment ook een leerproces is bij de klantmanager die zich de theorie, methodiek en praktische handvatten eigen moet maken. Dat doen we aan de hand van de hierboven geschetste stappen in de uitvoering.

Uitgangspunten

- Opvallend is het enthousiasme bij en de steun die het experiment in deze fase krijgt van de klantmanagers wat betreft de uitgangspunten, de filosofie en de uitwerking en juist ook vanwege de eerste ervaringen met de gesprekken met de deelnemer. Men omarmde met name de focus in het experiment op het loslaten van de verplichtingen en het zelf in actie laten komen van de deelnemer, de andere bejegening en de extra vrijlatingsmogelijkheden om de cliënt te stimuleren aan het werk te gaan.
- Allereerst wordt erkend dat het opbouwen van een vertrouwensrelatie met de cliënt tijd vraagt door het grote wantrouwen van de deelnemer naar de gemeente toe maar dat het wel veel beter werkt in de praktijk van de begeleiding. De deelnemers zijn in eerste instantie achterdochtig en wantrouwend en zijn afwachtend. Later tonen ze zich dankbaar voor de andere houding van de klantmanager. Men toont zich wel bezorgd over de werkdruk en caseload die nauwelijks minder is dan in de reguliere begeleiding waardoor er geen tijd is voor extra begeleiding en contact.
- De klantmanagers vinden de manier van begeleiding beter omdat de mens met zijn of haar vragen en wensen en problemen centraal staat. Er ontstaat een vertrouwensband die kan helpen effectiever te begeleiden. Sommige klantmanagers vragen zich wel af of de extra tijd die de begeleiding vraagt zich altijd uitbetaalt in effectiviteit. Men wijst op de grote verschillen tussen de verschillende groepen van mensen in de bijstand, dat is de diversiteit van de deelnemers en het culturele aspect (bijvoorbeeld tussen ouderen en jongeren en tussen migranten en autochtonen).

Interventies/treatments

- Bij de klantmanagers leven veel vragen bij het onderscheidend vermogen van de treatments vooral tussen eigen regie of zelf in actie en extra begeleiding. Het idee van mensen leren zelfredzaam te zijn wordt omarmd maar men vraagt zich af wat dat betekent in de praktijk wanneer de cliënt vraagt om extra begeleiding en tijd in de vorm van contacten. Er bestaat behoefte aan intervisie en overleg over 'moeilijke' cases. In het wekelijkse overleg tussen de klantmanagers wordt dit besproken.
- Nochtans hebben de klantmanagers die betrokken zijn bij de 'eigen regie' treatment soms problemen om de grens te bepalen tussen 'leren zelfredzaam te worden' hetgeen aan het begin meer begeleidingstijd vraagt en 'afstand nemen' of zo veel mogelijk loslaten en het aan de cliënt overlaten om actie te nemen ('zelf aan het stuur'). Het loslaten vraagt ook veel van de klantmanager. Anderzijds wordt erkend dat het leren van de cliënt de 'eigen regie' te pakken wellicht wel de beste manier is om hen te helpen.
- Men vraagt zich af of de aanpak van het experiment bij iedere deelnemer zal werken aangezien mensen sterk verschillen en dus verschillende soorten interventies nodig hebben. Veel mensen in de bijstand hebben te maken met ernstige beperkingen in de fysieke en/of mentale sfeer zeker ook in Wageningen. Daardoor krijgt de treatment van 'eigen regie' een andere betekenis dan bij mensen die geen beperkingen ervaren. Hoewel het experiment mensen met ernstige fysieke of mentale beperkingen uitsluit is dit soms lastig te bepalen, waar ligt de grens? Afgezien daarvan zijn er andere groepen die niet zijn uitgesloten van deelname maar een soortgelijke gezondheid problematiek kennen. Wageningen heeft ervoor gekozen zo min mogelijk mensen uit te sluiten dus ook statushouders niet, zelfs niet als ze de taal maar matig machtig zijn. De klantmanagers werken met tolken in dat geval om de deelnemer uit te leggen wat de bedoeling is.

Uitvoering/implementatie

- De klantmanagers van Wageningen overleggen regelmatig met elkaar om van elkaar te leren en een soort gemeenschappelijk gedragskader te ontwikkelen. Er bestaat behoefte aan aanvullende training omdat men het toepassen van de filosofie in de praktijk lastig vindt. Men vraagt zich af of het experiment veel verandert omdat Wageningen al lang een gemeente is met wethouders die een sterk sociaal gezicht laten zien. Er was daardoor in de reguliere begeleiding al veel in maatwerk voorzien. Dit zou er volgens de klantmanagers toe kunnen leiden dat later in het onderzoek er niet veel verschillen tussen de treatments (eigen regie en extra begeleiding) worden gevonden.
- In relatie tot de training wordt opgemerkt dat het integratieve gedragsmodel en de 'skills' methodiek als nuttig en zinvol wordt ervaren maar dat deze specifieker zou kunnen worden toegespitst op de afzonderlijke treatments in plaats van op de theorie en methodiek van zelfredzaamheid in het algemeen. Door de klantmanagers wordt meer aandacht hiervoor gevraagd om meer inzicht te krijgen in de verschillen tussen de interventies en het bijzondere van elke treatment.
- Een derde aspect dat in dit verband werd genoemd is de invloed van verschillen tussen de deelnemers maar ook tussen klantmanagers zelf; de een begeleidt nu eenmaal anders dan de andere. Opgemerkt wordt dat de invloed van de klantmanager kan worden gemeten bijvoorbeeld middels de informatie in de procesevaluatievragenlijsten. Volgens het RCT-design zouden de klantmanagers ook willekeurig moeten zijn toebedeeld aan de interventies maar in de praktijk is

dat moeilijk gebleken en is rekening gehouden met de wensen, motivatie en capaciteiten en persoonlijkheid van de klantmanager.

- Het gevoel bij klantmanagers bestaat dat men 'al doende leert' en beter snapt wat de bedoeling is en wat dat in de begeleiding betekent en vraagt. Dat vraagt echter tijd. Het wekelijkse overleg biedt de mogelijkheid ervaren knelpunten te bespreken en een gezamenlijke aanpak te vinden. Het schriftelijk vastleggen van de inhoud van de treatments heeft ertoe bijgedragen dat men bewust wordt van de grenzen van de verschillende treatments en dat men dus weet wat wel en niet werkt. Het vraagt wel voortdurend afstemming en aanpassing hetgeen de betrokkenen kenschetsen als een geleidelijk leerproces voor henzelf.
- Wat de uitvoering binnen de organisatie betreft wordt opgemerkt dat de organisatie moest wennen aan het experiment. Het is een leerproces voor henzelf maar ook voor de organisatie die moet leren en nog moet wennen aan het onderliggende gedachtegoed en de implementatie daarvan in de praktijk. De schaal van het experiment in Wageningen heeft daarbij zeker invloed gehad.

Leerproces

In het eerste half jaar bleek er een grote behoefte te bestaan aan het bespreken van de casuïstiek omdat men in de uitvoering aanliep tegen vragen waar men graag nader over van gedachten wilde wisselen om meer inzicht en praktische handvatten te krijgen voor de uitvoering van de treatments. Het focusgroep gesprek was ook bedoeld om lessen te trekken uit de ervaringen voor het experiment en dus te leren hiervan. Het eerste half jaar werd ervaren als een leerproces zowel voor henzelf als de organisatie waarbij men met vallen en opstaan een min of meer voor de praktische uitvoering van de treatments een gemeenschappelijke werkwijze en methodiek praktisch heeft vorm gegeven.

Groei en ontwikkelingsfase

Na de eerste 3 tot 6 maanden van verkenning en aftasting welke werkwijze het beste past is in het tweede focusgroep gesprek na bijna een jaar de balans opgemaakt van de ervaringen in dat jaar. Deze tweede fase is de fase van groei en ontwikkeling waarbij men in gezamenlijkheid de gekozen werkwijze en methodiek verder heeft uitgewerkt en die zo goed mogelijk uitgevoerd in de praktijk. Hieronder worden de bevindingen kort samengevat in deze groeifase.

Uitgangspunten en bejegening

- De klantmanagers ervaren in het eerste jaar dat de andere bejegening positieve uitkomsten heeft op de beleving door de cliënt, en dat het zijn of haar welbevinden en de vertrouwensrelatie met de cliënt ten goede komt. De klantmanagers ervaren dat aandacht geven en vertrouwen opbouwen (veel) tijd vraagt waardoor er meer contact nodig is met de cliënt ook in de eigen regie groep. Soms geven deelnemers in de eigen regie groep echter aan dat eigen regie in hun ogen betekent dat er juist minder contact nodig is omdat ze het zelf moeten doen. Daarin zit een zekere spanning.
- Men weet in bepaalde casussen niet goed hoe te handelen omdat een duidelijk handelingskader ontbreekt. Men vaart op het eigen gevoel maar soms ontbreekt inzicht in de achterliggende factoren die het gedrag van de cliënt kunnen verklaren. Vooral wanneer de cliënt zich weinig constructief opstelt of zelfs obstructie pleegt vraagt men zich af hoe te handelen. Wanneer is het gerechtvaardigd om de cliënt te verwijderen uit het experiment of de klassieke "stok" van

maatregelen en sancties te gebruiken terwijl het experiment uitgaat van vertrouwen en het loslaten van verplichtingen en dus minder toepassen van maatregelen. Dat speelt in elke treatment maar vooral bij de eigen regiegroep kwam die vraag naar voren.

Interventies/treatments

- In de praktijk is het soms lastig om de treatment zuiver toe te passen omdat acties van de klantmanager om bijvoorbeeld vacatures open te stellen niet beperkt zijn tot bepaalde treatments terwijl eigen regie ervan uitgaat dat men zelf in actie komt. Ook komt het incidenteel voor dat cliënten werk weigeren of niet meewerkt aan proefplaatsingen en duidelijk is dat de cliënt niet echt wil meewerken aan de treatment of dat zelfs sprake is van vermoeden van mogelijke fraude. De normale reflex zou zijn een maatregel op te leggen maar dat is in het experiment niet de intentie. Wanneer is met andere woorden de grens bereikt en hoe ga je daarmee om in de praktijk vormt een belangrijk dilemma in de praktijk?
- Ook de juiste dosering van het aantal contacten met de cliënt vooral in de eigen regiegroep werd als problematisch ervaren. Als het doel is cliënten te leren zelfredzaam te zijn of te worden kan dat betekenen dat in het begin veel contacten nodig zijn enerzijds om vertrouwen te kweken anderzijds om duidelijk te maken dat de klant zelf de eigen regie moet oppakken en niet niets kan doen en achterover kan leunen. In de intensieve treatment groep is er veel contact met de cliënt en kan gemakkelijker een vertrouwensrelatie wordt gekweekt waardoor ook sneller duidelijk is of men echt gemotiveerd is of de kantjes ervan afloopt.

Uitvoering/implementatie

- Opnieuw wordt verwezen naar de grote verschillen tussen de deelnemers en hoe daarmee om te gaan. Eigen regie lijkt voor bijvoorbeeld de groep van migranten per definitie niet goed te werken. De meervoudige problematiek waar deze mensen veelal mee te maken hebben vraagt eigenlijk om intensieve integrale begeleiding. Daarnaast is het cultuurverschil erg groot hetgeen ook de treatment van intensieve begeleiding bemoeilijkt. Het uitgangspunt van de begeleiding is dat gezamenlijk een plan wordt gemaakt maar dat is voor deze doelgroep onhaalbaar.
- Lang is gesproken over de voorwaarden binnen de organisatie om het experiment goed uit te voeren. Er verandert veel in het werk, er worden veel ambities vanuit de organisatie geformuleerd zoals het werken in multidisciplinaire clusters hetgeen veel aandacht opeist van de klantmanagers waardoor ze minder tijd hebben voor de deelnemers in het vertrouwensexperiment. De ondersteuning vanuit de organisatie is volgens de klantmanagers nog niet optimaal. Ook de organisatie maakt een leerproces door.

Implementatie

Het eerste jaar van implementatie wordt als een leerproces ervaren om het experiment eerst goed vorm te geven en vervolgens uit te voeren conform de opzet en het onderliggende gedachtegoed. Daarbij zijn een aantal knelpunten en dilemma's geformuleerd die deels te maken hebben met de praktische uitvoering van de treatments zoals de dosering van het aantal contacten en hoe om te gaan met moeilijke gevallen, waarbij vooral bij eigen regie men onzeker was over hoe deze uit te voeren wanneer de deelnemer zich niet gemotiveerd toont of wanneer deze vanuit een andere culturele achtergrond een andere aanpak vraagt en deels met de lokale uitvoeringscontext die verandert en die soms ambivalent omgaat met het vertrouwensexperiment. Het gevaar wordt onderkend dat men bij niet gemotiveerde deelnemers snel

terugvalt op de klassieke strenge aanpak wanneer een alternatieve aanpak ontbreekt of niet het gewenste effect heeft.

De bezinningsfase

In het derde half jaar volgt de bezinning op wat men heeft gedaan en wat er is bereikt. Men heeft inzicht gekregen in wat wel en niet werkt en wat beter kan. In het derde focusgroep gesprek van 21 maart 2019 komen alle eerder genoemde punten daarom ook weer terug. Men vindt het belangrijk om de ervaren dilemma's en hoe daarmee is omgegaan in de praktijk met de onderzoekers te delen. Wat heeft men ervan geleerd, welke lessen zijn getrokken? Hierbij wilde men ook een blik vooruit werpen naar de toekomst, na afloop van het experiment. Gesproken is over de overgangperiode wanneer het experiment afloopt (per 1 oktober 2019) maar nog geen besluiten zijn genomen door het Kabinet over de consequenties van het experiment voor de Participatiewet. Wat betekent dat voor de deelnemers en hoe moet ik als klantmanager omgaan met de onzekerheid die dat met zich meebrengt bij de deelnemers? Aan het Kabinet is gevraagd om uitstel te verlenen tot 1 januari 2020 om te voorkomen dat deelnemers vanwege de beëindiging van het experiment niet meer wensen deel te nemen aan het onderzoek of hun gedrag op de arbeidsmarkt veranderen omdat men geconfronteerd wordt met de gevolgen van de beëindiging van het experiment. Dat zou de effectmeting kunnen beïnvloeden. Daarnaast is gesproken over de lessen die zijn getrokken uit de ervaringen met de andere aanpak in het experiment. Hieronder worden de bevindingen in deze bezinningsfase kort samengevat.

Uitgangspunten

- De uitgangspunten van het experiment om uit te gaan van vertrouwen en aandacht, maatwerk en integraal werken worden door de klantmanagers gedeeld en ondersteund. Vanuit de historie van Wageningen is er wat maatwerk betreft niet veel veranderd. Weliswaar was er wel een stevig en strikt beleid ten aanzien van handhaving maar het uitgangspunt was wel maatwerk te leveren. In de ogen van de klantmanagers is er wat dat betreft met het experiment niet zo veel veranderd. Het experiment ondersteunt de oude filosofie van maatwerk en gaat een stapje verder wat betreft zelfredzaamheid.
- De meerwaarde van vertrouwen en aandacht geven voor het begeleiden van mensen wordt door veel klantmanagers erkend. Men heeft wel moeite met het idee van "loslaten" gezien de ervaringen met sommige niet gemotiveerde deelnemers in de eigen regie groep. Men hecht eraan om in voorkomende gevallen gebruik te kunnen blijven maken van de toepassing van maatregelen wanneer het gedrag van de cliënt daar aanleiding toe geeft.
- Het experiment werd als nuttig ervaren omdat het handvatten geeft om als team een meer methodische aanpak van begeleiding tot stand te brengen. Het doelgericht begeleiden van mensen naar zelfredzaamheid werd als methodiek als zinvol ervaren. Klantmanagers vinden het dan wel belangrijk dat de organisatie hen daarin steunt en ruimte geeft om de nieuwe aanpak ook daadwerkelijk te realiseren. Dat betekent dat meer tijd en aandacht nodig is en dus de caseload lager.
- In het algemeen vindt men het willekeurig toewijzen van deelnemers aan de treatmentgroepen in de praktijk lastig. Men begrijpt dat dit voor het onderzoek nodig is maar men ervaart een rolconflict

aangezien men van mening is dat een treatment moet passen bij de persoon voor wie deze bestemd is. Na het experiment zouden de verschillende interventies gericht ingezet kunnen worden afhankelijk van de situatie en noden van de cliënt.

Interventies/treatments

- De drie treatments verschillen in de mate waarin de klantmanagers hier heil in zien. Men ziet niet veel nut in het volledig loslaten van de deelnemer in eigen regie. Juist door frequent contact ontstaat er een band die nodig is om vertrouwen tot stand te brengen. Pas dan kan een appel worden gedaan op de zelfredzaamheid van een persoon.
- Voor veel deelnemers is deeltijdwerk het maximaal haalbare. De ruimere vrijlating kan ertoe bijdragen dat deeltijdwerken aantrekkelijker wordt. Deeltijdwerk mag natuurlijk niet ertoe leiden dat voltijdswerk wordt verdrongen. Overigens wordt in deze groep voor het overige de reguliere dienstverlening aangeboden. Voor veel deelnemers zal de baan tevens een flexbaan zijn (oproep contract, tijdelijke baan, uitzendbaan).
- Men is het meest enthousiast over de intensieve begeleidingsgroep in het experiment omdat deze de mogelijkheid geeft om meer tijd en contact te besteden aan de begeleiding van de cliënt. Indien tevens maatwerk kan worden geboden is de verwachting dat dit de cliënt ten goede kan komen en bijdraagt aan het succes van re-integratie.

Uitvoering/implementatie

- De uitvoering heeft voor de andere manier van werken behoefte aan begeleiding en training waarin klantmanagers bijstandsgerechtigden leren zelfredzaam te zijn. Een systematische begeleiding en training van de klantmanagers kan in het experiment leiden tot een meer gestructureerde ofwel methodische manier van werken.
- De implementatie van de uitgangspunten van het experiment zou zo snel mogelijk moeten plaatsvinden. Indien deelnemers weer terugvallen in het oude regime na deelname aan het experiment wordt het vertrouwen aangetast. Een overgangsregime kan hierin soelaas bieden.
- Klantmanagers verwachten positieve effecten van het experiment mits de organisatie de uitvoering en implementatie ondersteunt door de randvoorwaarden te scheppen voor succesvolle invoering van de nieuwe aanpak.

Uitbouw en toekomst

Kortom, in deze bezinningsfase spreken de klantmanagers uit dat zij op grond van hun ervaringen gedurende de experimenteerperiode positief staan tegenover de filosofie en inhoud van het experiment. Men verwacht dat dit kan bijdragen aan een meer methodische en professionelere manier van begeleiding. Men hecht vooral veel waarde aan het aspect van het opbouwen van een vertrouwensband, extra begeleiding, meer tijd voor het 'echte gesprek' en een individuele maatwerk gerichte aanpak afgestemd op de situatie en persoonlijke omstandigheden van de cliënt. Over de treatment van eigen regie is men over het algemeen kritisch en wordt men niet erg enthousiast om de cliënt het zelf te laten uitzoeken en los te laten. Men erkent wel het belang van zelfredzaamheid maar daarvoor is wel contact met de cliënt voor nodig. Men is over het algemeen ook kritisch over de vraag in hoeverre de organisatie de andere aanpak ondersteunt middels een lagere caseload alsook in de sfeer van randvoorwaarden.

Gespreksverslagen klantmanagers met projectleider

De onderzoekers hebben de gespreksverslagen van de projectleider met de klantmanagers kunnen inzien. In grote lijnen bevestigen deze het beeld dat uit de focusgroep gesprekken en de procesevaluatie vragenlijsten is gekomen. Men staat over het algemeen positief tegenover het gedachtegoed van het experiment om te werken vanuit vertrouwen, tijd voor het echte gesprek, aandacht voor wat de betrokkene zelf wil inclusief het bevorderen van de zelfredzaamheid en extra begeleiding in de vorm van maatwerk. Er werd in het exit gesprek regelmatig gerefereerd naar de gevlugelde uitspraak “alles wat aandacht krijgt groeit”. Men erkent dat een ‘mensgerichte’ aanpak voor de deelnemers de spanning vermindert en ruimte biedt om met de eigen doelen of wensen aan de slag te gaan. Men heeft wat minder op met ‘zelf in actie’ omdat men vindt dat loslaten vaak niet de beste methode is en een ‘stok’ achter de deur bij sommige mensen nodig is om hen in beweging te krijgen. In Wageningen zijn er relatief veel statushouders onder de deelnemers. Daarvoor is het experiment volgens de klantmanagers minder geschikt. Door de culturele verschillen en de taalproblemen vereist dit een specifieke aanpak. De klantmanagers staan over het algemeen positief tegenover een doelgerichte methodische aanpak hoewel ze denken dat dit voor sommige deelnemers een te abstracte benadering is. Soms is een weg van geleidelijkheid volgens hen nodig om duidelijk te krijgen wat de beste aanpak voor deze persoon is. Over vrijlating zijn de meningen verdeeld, men ziet wel het belang van deeltijdwerk aangezien voor veel deelnemers voltijdse uitstroom een onhaalbare kaart is maar men vindt het lastig te bepalen hoeveel uur men daadwerkelijk zou kunnen werken en of voltijds werk wel of niet haalbaar zou zijn. Men benoemt ook het probleem van de inkomensberekening van de uitkering bij vrijlating en de afbouw van toeslagen en de moeite die de gemeentelijke administratie hiermee heeft. Wat het onderzoek zelf betreft ziet men met belangstelling uit naar de resultaten maar vindt men anderzijds de belasting ervan (vragenlijsten, administratie) groot naast de bestaande taken.

5 Bevindingen en conclusies

5.1. Opzet van het experiment en werving

Per 1 oktober 2017 is de gemeente Wageningen gestart met een experiment in het kader van artikel 83 van de Participatiewet met als titel en adagium “Vertrouwen werkt!”. Wageningen behoorde tot de vier initiatiefgemeenten die samen met Utrecht, Tilburg en Groningen in 2015 een andere aanpak voorstelden vanwege twijfels over de effectiviteit van de Participatiewet die per 1 januari 2015 van kracht was geworden. In de gekozen opzet is gedurende 2 jaar conform de opzet in de andere experimenteergemeenten, Tilburg, Groningen, Deventer, Nijmegen en Utrecht) een ‘random control trial’ (RCT) experiment uitgevoerd waarbij deelnemers aselekt zijn toegewezen aan vier ‘treatments’ of interventies. Dit zijn eigen regie of ‘zelf in actie’ (onthefing en ‘leren zelfredzaamheid’), extra begeleiding (intensieve maatwerkbegeleiding), vrijlating (‘verdienen loont’ plus reguliere begeleiding) en de onderzoeksgroep met de reguliere dienstverlening (‘care as usual’).

Werving en onderzoek: invulling van de vragenlijsten

Alle bijstandsgerechtigden in Wageningen met uitzondering van enkele specifieke groepen die zijn uitgesloten van deelname aan het experiment hebben een uitnodiging gekregen voor deelname. Uitgesloten werden ouderen boven 65 jaar, dak- en thuislozen, arbeidsongeschikten, statushouders die de taal niet machtig zijn en mensen verblijven in inrichting. Er is veel aan gedaan om de werving tot een succes te maken met voorlichtingsmateriaal, filmpjes, voorlichtingsbijeenkomsten etc. alsook een persoonlijke aanpak om mensen te bewegen mee te doen. Het oorspronkelijke streefaantal van 300 aanmeldingen werd ruimschoots gehaald. Uiteindelijk hebben 410 mensen zich aangemeld waarvan in de loop van de tijd 18 mensen zich alsnog afmeldden waardoor er 392 deelnemers resteerden. Na aanmelding werden mensen aselekt ingeloot in een van de vier groepen. Deelname was vrijwillig, na indeling in een groep mocht men alsnog weigeren maar niet meer meedoen aan het experiment.

De invulling van de Qualtrics online vragenlijsten door de deelnemers bleek vooral na de nulmeting lastig maar uiteindelijk is een respons gehaald van 85% in de nulmeting en 60% in de vervolgmetingen. Door de gunstige arbeidsmarkt stroomde 20% uit de uitkering naar werk waardoor het aantal deelnemers in begeleiding redelijk voortvarend slonk maar tegelijkertijd ook de respons naar beneden ging. Eenmaal uitgestroomd werd de motivatie minder om de vragenlijst in te vullen. De klantmanagers vulden drie keer een vragenlijst in (aan het begin, na 1 jaar en na 2 jaar) en deze is voor 90% voor alle deelnemers ingevuld. Na 2 jaar was 40% uitgestroomd naar werk of anderszins (verhuizing, studie etc.).

5.2 Theoretische uitgangspunten en inhoud treatments

Het experiment beoogt voor deze deelnemers een alternatieve aanpak en bejegening te ontwikkelen uitgaande van vertrouwen, eigen regie, keuzevrijheid, zelfredzaamheid en vraaggericht maatwerk. Hieraan ligt een theoretisch kader ten grondslag ontleend uit de gedragseconomie (vertrouwen en belonen) en de sociale en psychologische literatuur (zelfredzaamheid, intrinsieke motivatie, vertrouwen en wederkerigheid, gezondheid en welbevinden). Daarnaast werd verwezen naar Sen’s capability benadering in verband met keuzevrijheid om de dingen te kunnen doen men voor het eigen leven belangrijk vindt. In

de aanvraag is de centrale probleemstelling van het Wageningse Vertrouwensexperiment als volgt verwoord: “In hoeverre kunnen de doelstellingen van bevordering van (arbeids)participatie beter bereikt worden door een andere bejegening en een andere handelingsvrijheid enerzijds (eigen regie), gecombineerd met positieve financiële prikkels ofwel het belonen van eigen initiatief door middel van een hogere inkomstenvrijlating (vrijlating), alsmede door een vraaggerichte, intensieve begeleiding op maat (extra begeleiding)”. Ten behoeve van de uitvoering hebben de klantmanagers een training gekregen op basis van de zogenaamde “skills” methodiek en het integratieve gedragsmodel waarin vaardigheden zijn geleerd om de zelfredzaamheid van de deelnemers te bevorderen. Eigen regie is dus geen ‘loslaten’ maar ‘leren zelfredzaam te zijn of te worden’ door het gesprek aan te gaan, het initiatief bij de deelnemer te laten en handelingsvrijheid te geven om zelf in actie te komen.

Opzet onderzoek

Het onderzoek bestaat uit twee onderdelen: de product –en procesevaluatie (zie Muffels & van der Klink 2017). De *productevaluatie* is erop gericht na te gaan wat beter werkt voor de deelnemende bijstandsgerechtigde; de nieuwe aanpak of de reguliere aanpak. Er wordt voor de evaluatie een onderscheid gemaakt tussen primaire en secundaire uitkomstmaten. De primaire uitkomstmaat betreft uitstroom naar voltijd (27 uur of meer) of deeltijdwerk (meer dan 8 of 12 uur). De secundaire uitkomstmaten betreffen zelfredzaamheid, gezondheid (fysiek, mentaal), welbevinden, en sociale participatie. Over een periode van 2 jaar worden deelnemers aan het experiment gevolgd door hen op drie momenten een vragenlijst te laten invullen. De *procesevaluatie* is erop gericht het proces van begeleiding in de uitvoering beter in beeld te krijgen en te onderzoeken op welke wijze en in hoeverre de treatments zijn uitgevoerd zoals beoogd. Op deze manier kan worden onderzocht of bij het ontbreken van effecten sprake is van een theorie of uitvoering ‘failure’. Ook de klantmanagers vullen drie keer een vragenlijst in over het proces van begeleiding. Onderdeel van de procesevaluatie was tevens het voeren van drie focusgroep gesprekken met de klantmanagers gedurende het experiment. Naast de data uit de vragenlijsten wordt gebruik gemaakt van de lokale BUS (bijstandsuitkeringsstatistiek)-data. De gegevens van het experiment worden daarna gekoppeld aan de micro-data gegevens van het CBS. Hierdoor zijn er meer gegevens over de deelnemers bekend, zoals inkomen, en kan ook gekeken worden naar deeltijd uitstroom en uitstroom naar tijdelijk werk. Bovendien kunnen mensen dan op de arbeidsmarkt worden gevolgd na afloop van het experiment. In dit eindverslag zijn alleen de lokale BUS-gegevens gebruikt.

Verandering tijdens de uitvoering

De wereld staat tijdens het experiment evenwel niet stil. De politieke en beleidscontext verandert maar er zijn ook leereffecten bij de coaches en deelnemers waardoor de treatments kunnen veranderen. Dat maakt het experiment complex want voor een juiste vergelijking zouden de treatments zoveel mogelijk hetzelfde moeten blijven. De verandering in uitvoeringspraktijk en beleidscontext beïnvloedt mogelijk de standaard of reguliere treatment waardoor een zuivere vergelijking tussen de reguliere aanpak en de treatments in het onderzoek werd bemoeilijkt. Het eindverslag gaat daar op verschillende plaatsen nader op in omdat er redenen zijn om aan te nemen dat dat het geval kan zijn geweest.

Wageningen heeft bij de start gekozen voor een bepaalde methodiek waarbij op basis van een werkplan een hoogst bereikbare doel of doelen werden gesteld in overleg met de deelnemer. Dat gold ook voor de

onderzoeksgroep met reguliere dienstverlening. De klantmanagers in Wageningen hebben daarnaast vaak om praktische redenen meerdere treatments begeleid. De klantmanagers hebben als team geopereerd waarbij wekelijks werd overlegd over de aanpak, de invulling van de treatments en de casuïstiek. Tevens hebben ze een training gehad en zijn gezamenlijke casuïstiek besprekingen gehouden. Om deze redenen kunnen “cross-over” effecten tussen de treatments zijn ontstaan waardoor ze meer op elkaar zijn gaan lijken. Voorts vinden we significante treatment effecten voor de onderzoeksgroep in de procesevaluatie voor verschillende uitkomstmaten en in de productevaluatie bij bepaalde uitkomsten. Dit duidt erop dat er over de tijd iets is veranderd in de standaard treatment zelf. Dat is deels ook vanzelfsprekend omdat beleid meevolueert maar deels heeft het ook te maken met een soort “Hawthorne” effect, een gedragseffect van deelname aan het experiment en het onderzoek door de klantmanagers en de deelnemers waardoor zij zich anders gaan gedragen. De vergelijking tussen de treatments en de controlegroep werd daarmee problematischer.

5.3 Uitstroom naar werk en matching

De waargenomen uitstroomkansen naar werk (voltijds) zijn voor de 391 deelnemers uit het zittende bestand 4-5% hoger voor de treatment groepen eigen regie en extra begeleiding in vergelijking met de controlegroep of onderzoeksgroep. Voor de vrijlatingsgroep is er geen of zelfs een klein negatief verschil met de controlegroep. Na 1 oktober 2019 (het einde van het experiment) zijn tot aan het einde van het jaar evenwel nog relatief veel mensen uitgestroomd. De resultaten voor de vrijlatingsgroep zijn over deze langere periode dan gunstiger. De geschatte en voor een aantal achtergrondvariabelen (leeftijd, geslacht, duur uitkering bij start, nationaliteit, aantal keer bijstand bij start, gezinssamenstelling of leefvorm, huisvestingssituatie) gecorrigeerde kansen (logit regressiemodel) liggen dicht bij de waargenomen kansen. Dit geeft aan dat de vertekening die mogelijk is opgetreden erg klein is.

Nochtans vinden we geen significante verschillen ($p < 0,10$) in de ‘intention-to-treat’-analyse in vergelijking met de controlegroep (de analyse inclusief de afmeldingen). Dat betekent dat er geen betekenisvolle verschillen zijn in voltijdse uitstroom naar werk tussen elk van de drie treatmentgroepen en de controlegroep en dat het effect van elk van de treatments ten opzichte van de controlegroep dus nul is.

Power probleem

Dit is opvallend en kan te maken hebben met een power (kleine aantallen) probleem waardoor schattingen niet significant zijn op bijvoorbeeld 90% ($p < 0,10$) maar wel op 85% betrouwbaarheidsniveau ($p < 0,15$). Een indicatie daarvoor is ook dat als we eigen regie en extra begeleiding samen nemen en de uitstroomkansen vergelijken met de controlegroep dat we dan wel een significant effect vinden ($p < 0,10$ op basis van een ttest). Om die reden geven we ook de effecten weer bij significantieniveaus groter dan 85% maar lager dan 90% ($p < 0,15$).

Vergelijking met niet-deelnemers

In het model waarin we de uitstroom naar voltijds werk vergelijken met de referentiegroep van niet-deelnemers zien we wel significante positieve verschillen bij eigen regie. Het verschil in percentage tussen eigen regie en extra begeleiding en de referentiegroep is met 11-13% ook hoger dan in vergelijking met de

controlegroep. Het effect van extra begeleiding is net niet significant op 90% maar wel op 85% betrouwbaarheidsniveau ($p < 0.15$). De referentiegroep is echter niet gerandomiseerd waardoor we niet weten of het verschil causaal geduid kan worden. Daarnaast heeft er zelfselectie plaatsgevonden bij de aanmeldingen. Zo bleek dat in de onderzoekpopulatie langere duren waren ondervertegenwoordigd. Ook naar nationaliteit was de groep niet helemaal representatief. Om hiervoor te corrigeren is 'matching' toegepast. Daartoe zijn er twee matching methoden gebruikt: 'nearest neighbour' (wegingsmethode) en 'propensity score' matching (modelmatige schatting van de kans op treatment). Omdat we geen verschillen vonden tussen eigen regie en extra begeleiding hebben we beiden samen genomen in de analyse. Dat betekent dat als we uitgaan van het geschatte gemiddelde treatment effect (ATE) over deelnemers en niet-deelnemers er volgens beide methoden een 4,8% tot 5,6% hogere kans bij de treatmentgroepen is om volledig uit te stromen in vergelijking met de referentiegroep van niet-deelnemers. Dat is lager dan we vonden in de ruwe data of bij de logit regressie resultaten (11-13%) maar niettemin een substantieel effect.

Deeltijdwerk en flexwerk

Uit de gegevens van de vragenlijsten (meting 2 en 3) bleek dat ongeveer een op de drie deelnemers (35%) zegt te werken als werknemer of als zelfstandige. Indien 20% van deelnemers voltijds werk heeft gevonden (27 uur of meer per week) dan betekent dit dat ongeveer 15% in deeltijd werkt (minder dan 27 uur). We hadden verwacht dat vanwege de vrijlating de meeste deeltijdbanen in de vrijlatingsgroep zouden voorkomen. Dat bleek ook het geval te zijn, maar opvallend genoeg niet veel meer dan in de controle -of onderzoeksgroep. Vier van de vijf deelnemers die werk vonden tussen 1 oktober 2017 en 1 oktober 2019 is uitgestroomd naar een tijdelijke baan. We vonden geen grote verschillen tussen de treatment groepen behalve voor eigen regie die een iets hogere uitstroom realiseerde naar vaste banen.

5.4 Niet-werk gerelateerde uitkomstmaten

De meeste treatment effecten die we onderzochten zijn niet significant (tabel A1.2). De daarbij horende hypothesen worden daarom dus weerlegd. We vinden nochtans enkele significante treatment effecten zoals bij extra begeleiding voor subjectieve gezondheid en bij eigen regie en extra begeleiding in meting 2 en 3 voor mentale gezondheid. Daarnaast vinden we een negatief significant effect bij extra begeleiding en vrijlating voor het aantal uren besteed aan vrijwilligerswerk en mantelzorg. Het effect is negatief in vergelijking met de controlegroep. Nadere analyse laat evenwel zien dat de daling in het aantal uren vrijwilligerswerk en mantelzorg in beide treatments volledig voor rekening komt van de toename in aantal uren betaald werk. De daling is dus een gevolg van de treatment doordat meer mensen vanwege de vrijlating of de extra begeleiding in beide groepen werk vinden maar daardoor minder tijd hebben voor vrijwilligerswerk en mantelzorg. Om die reden is dit effect eigenlijk geen negatief effect maar juist een positief effect van de treatment, althans indien uitstroom naar betaald werk het primaire doel is. Er treedt dus substitutie op waardoor de grotere arbeidsparticipatie ertoe leidt dat mensen minder tijd hebben voor maatschappelijke participatie. Uit de gegevens in de vragenlijsten over de uitstroom naar deeltijdarbeid kwam het beeld naar boven dat bij de vrijlatingsgroep maar juist ook de onderzoek -of controlegroep de uitstroom naar deeltijdwerk relatief groot is ten opzichte van de andere groepen.

We vinden weinig significante positieve effecten vinden maar ook weinig negatieve effecten. Dat betekent dat de andere werkwijze in ieder geval niet slechter werkt op deze uitkomstmaten dan de reguliere

begeleiding. Nochtans vinden we naast het eerdere genoemde negatieve effect voor vrijwilligerswerk een negatief effect voor institutioneel vertrouwen. Deze doet zich bij alle treatment groepen voor inclusief de controlegroep. Bij eigen regie is dit negatieve effect al bij de nulmeting aanwezig en duidt het op een selectie effect en geen treatment effect. Tot slot vinden we bij eigen regie een negatief treatment effect voor inkomensarmoede. Dat zou erop kunnen wijzen dat de inkomensarmoede dankzij de treatment eigen regie is gedaald maar nadere analyse leerde evenwel dat dit negatieve effect bij eigen regie zich reeds bij de start van het experiment voordeed en dus geen echt treatment maar vermoedelijk een selectie-effect is. Over het algemeen vinden we dus geen treatment effecten voor armoede en is ook de daling bij deprivatiearmoede in de tijd die we signaleerden niet significant na correctie voor compositieverschillen.

Ontwikkeling in de tijd

Toch ontwikkelen een aantal belangrijke uitkomstmaten zich in de tijd gunstig (positief effect) zoals bij welbevinden, subjectieve gezondheid, mentale gezondheid, sociaal vertrouwen en armoede maar niet beter dan in de controlegroep. We vinden ook negatieve effecten zoals bij ervaren keuzevrijheid, sociaal netwerk, vrijwilligerswerk, institutioneel vertrouwen en inkomensarmoede. Bij vrijwilligerswerk bleek dat samen te hangen met het extra aantal uren dat men ging werken welke ten koste gingen van het aantal uren vrijwilligerswerk en mantelzorg. Bij keuzemogelijkheden en sociaal netwerk vermoeden we dat door uitstroom naar betaald werk er een selectieve groep van deelnemers overblijft met weinig perspectief op een betaalde baan. Bij inkomensarmoede bleek het een selectie effect en geen treatment effect. Opvallend is dat er niet alleen bij inkomensarmoede maar ook bij de andere maatstaven reeds bij de start verschillen zijn in de uitgangssituatie hetgeen duidt op zelfselectie. Die doet zich met name voor bij de eigen regie groep en de controle -of onderzoekgroep. Nader onderzoek moet uitwijzen of onderliggend bijvoorbeeld de stijging in welbevinden, sociaal vertrouwen en mentale gezondheid niet toch een treatment effect is dat niet zichtbaar wordt omdat de reguliere begeleiding in de controlegroep gedurende het experiment is veranderd en is gaan lijken op de andere treatments. De stijging van welbevinden is opvallend omdat er geen evidentie is dat welbevinden bij bijstandsgerechtigden zou zijn toegenomen in de afgelopen twee jaar, integendeel cijfers van het SCP (Wennekers e.a. 2019) suggereren dat voor mensen met een score op tevredenheid met het leven of geluk van 6 of lager (waartoe ook de bijstandsgerechtigden behoren) de score de afgelopen jaren – dus na de economische crisisjaren - is verslechterd. We vinden over het algemeen weinig verschillen tussen de treatments. Wel is het zo dat als er verschillen zijn deze zich vooral voordoen bij eigen regie en extra begeleiding maar niet bij de vrijlatingsgroep met uitzondering van het substitutie-effect van betaald werk op uren besteed aan vrijwilligerswerk

Stabiele kenmerken

Dat er weinig treatmenteffecten gevonden worden bij deze uitkomstmaten heeft volgens ons te maken met twee aspecten. Ten eerste denken we dat we geen goede vergelijkingsgroep hebben omdat de standaard treatment misschien zelf is veranderd. Ten tweede denken we dat veel van de gekozen uitkomstmaten stabiele kenmerken zijn van mensen die niet zo snel veranderen, of het nu gaat om beleving en motivatie zoals bij welbevinden, subjectieve gezondheid, de perceptie van sociale integratie, vertrouwen of om gedrag zoals bij zoekgedrag. Om deze kenmerken te veranderen zijn vaak ingrijpende gebeurtenissen nodig die het leven drastisch veranderen zoals een scheiding, langdurige werkloosheid, en verlies van kind

partner of vriend en vaak duurt het een lange tijd voor een effect zich openbaart door een andere beleving, motivatie of gedrag.

Verandering van de reguliere begeleiding

De regressieanalyses zowel van de procesevaluatie als van de productevaluatie gaven naar onze mening aanwijzingen voor een verandering van de treatment in de controlegroep. Bij de uitkomstmaten van mentale gezondheid en welbevinden vonden we een positief significant treatment effect over de tijd (meting 3) bij de controlegroep. Indien de reguliere treatment niet was veranderd zou het vinden van een dergelijk effect onwaarschijnlijk zijn geweest. Met andere woorden, indien de normale dienstverlening inderdaad is veranderd, is geen zuivere vergelijking van de controlegroep met de treatment groepen meer mogelijk. Dan zullen we een vergelijking moeten maken met een vergelijkingsgroep die niet beïnvloed is door het experiment. Op dit moment is zo'n alternatieve vergelijkingsgroep niet beschikbaar.

Verschillen tussen de treatments extra begeleiding en zelf in actie

Opvallend is dat we op de secundaire uitkomstmaten weinig tot geen verschillen vinden tussen de treatments. Dit heeft verschillende oorzaken. De eerste is dat in het experiment met opzet is gekozen voor een 'eigen regie' groep die mensen leert zelfredzaam te worden. Dat betekent dus geen 'loslaten' zoals stricto sensu onder het begrip 'eigen regie' zou kunnen worden verstaan. De training aan de consulenten was gericht op zelfredzaamheid waarbij het onderscheid tussen eigen regie en intensieve begeleiding minder aandacht heeft gekregen. De treatments zijn aan het begin van het experiment nader omschreven en op schrift gesteld. Daarin werd ook aangegeven hoeveel contactmomenten er gemiddeld over de tijd zullen zijn per jaar met de deelnemer. Het was duidelijk dat bij eigen regie het aantal contactmomenten minder zou zijn en dat de regie voor contact bij de deelnemer zou behoren te liggen. De afgesproken 'case load' was voor eigen regie ook hoger vastgesteld waardoor er minder tijd zou zijn dan bij intensieve begeleiding per deelnemer. De ervaring leerde echter al snel dat de op papier gestelde uitgangspunten in de praktijk anders uitpakt. Reeds bij het begin van het experiment bleek dat 'leren zelfredzaamheid' in het begin juist veel contact vraagt met de deelnemer en dat de meeste deelnemers het juist prettig vonden in het experiment dat ze veel contact hadden met de klantmanager. Doordat vanwege de focus op zelfredzaamheid het onderscheid tussen beide treatments niet geheel duidelijk was en ook de contactintensiteit aan het begin niet verschilde was het voor klantmanagers in het begin niet duidelijk wat de verschillen waren tussen de treatments. Geleidelijk aan is er een praktische invulling gegeven aan dit onderscheid tussen eigen regie en extra begeleiding; later bleek ook dat er minder contact met de deelnemer nodig was. Ook al vinden we geen significante verschillen, de klantmanagers zelf zijn meer gerapporteerd van de extra begeleiding als treatment. Dat bleek heel duidelijk uit de analyse van de procesevaluatie vragenlijsten wanneer gevraagd werd naar welke treatment het meest geschikt werd geacht voor de deelnemer. Zij zijn van mening dat de deelnemers hier meer baat bij hebben dan de eigen regie treatment. De proactieve rol die klantmanagers hebben bij extra begeleiding past wellicht beter bij de rol die van hen werd en wordt verwacht als professional, 'loslaten' hoorde en hoort daar minder bij. Mogelijk heeft het rolconflict dat consulenten ervoeren bij het uitvoeren van de treatments - zie het verslag van de focusgroep gesprekken- ook een rol gespeeld. Ook al werd van hen gevraagd om de treatments zo zuiver mogelijk uit te voeren en niet te veranderen, vanuit ethisch oogpunt vonden consulenten het lastig om de deelnemer iets te onthouden wat op dat moment door de consulent als professional nodig wordt

geacht. Niet uitgesloten kan worden dat dit een effect heeft gehad op het zuiver uitvoeren van de treatments.

Procesevaluatie

In de procesevaluatie worden de opinies en verwachtingen van de klantmanagers gepeild van de uitkomsten voor de deelnemer in de verschillende treatmentgroepen. De verwachting was dat klantmanagers van de treatmentgroepen eigen regie, extra begeleiding en vrijlating op aspecten zoals motivatie, zelfredzaamheid, stijging op participatieladder en het effect van de treatment hogere scores zouden laten zien op de uitkomstmaten dan de onderzoekgroep, vanwege de andere werkwijze die deel uit maakte van deze treatments. De analyses laten zien dat vooral de uitkomsten bij de extra begeleidingsgroep in de visie van de klantmanagers positief afwijken van de uitkomsten bij de onderzoekgroep op de besproken uitkomstmaten. Op nagenoeg alle tien uitkomstmaten scoort de extra begeleidingsgroep volgens de klantmanager (sterk) significant hoger dan de onderzoekgroep. Met name bij het aspect motivatie en zelfredzaamheid, zoals zelf in actie komen en het zelf doelen stellen, komt de extra begeleidingsgroep beter uit de bus, omdat daar het treatment effect het grootst is. Vervolgens vonden we ook voor de vrijlatingsgroep op veel uitkomstmaten een significant positief effect ten opzichte van de onderzoekgroep, met name op het zelf doelen stellen en het leereffect voor de klantmanager. Voor de eigen regie groep vonden we enkele positieve effecten, hoewel deze minder sterk significant en minder groot waren dan bij de andere twee treatmentgroepen. Vooral bij het aspect motivatie en leereffect bij de klantmanager bleek deze groep positief af te wijken van de onderzoekgroep. Opvallend is dat veel lijntjes die het verloop van de uitkomsten in de tijd weergeven een stijgende lijn laten zien zeker in meting 3 ten opzichte van de nulmeting. Dat we bij de onderzoekgroep veel positieve significante effecten waarnemen op veel uitkomstmaten in meting 2 en meting 3 geeft te denken, en suggereert dat deze treatment in de loop van de tijd is gewijzigd. Eerder vermeldden we al dat klantmanagers betrokken zijn bij meer dan een treatment en dat geldt ook voor klantmanagers van de controlegroep. Cross-over effecten zijn dan aannemelijk.

4.5 Tot besluit

Op grond van de ruwe data vinden we substantiële verschillen in uitstroomcijfers naar voltijds werk tussen de treatments eigen regie en extra begeleiding en de controlegroep maar zeker ook met de referentiegroep van niet-deelnemers. Na correctie voor verschillen in achtergrondkenmerken worden deze verschillen niet significant behalve wanneer we extra begeleiding en eigen regie samen nemen, dan is het effect net significant op 90% niveau. We vermoeden dat we hier deels te maken hebben met een power (kleine aantallen) probleem. Op grond van de regressiemodellen worden vooralsnog de meeste hypothesen over de verschillen tussen de treatments en de standaardgroep weerlegd. Toch is er ook positief nieuws. Het gevonden positieve treatment effect op subjectieve gezondheid maar ook op mentale gezondheid is zeker niet onbelangrijk tegen de achtergrond van de bijstandpopulatie in Wageningen waarvan bijna 60% zich arbeidsongeschikt acht vanwege fysieke of mentale gezondheid problemen. Niet onbelangrijk zijn ook scores die klantmanagers geven aan de sociale participatie van de deelnemers en de positie van de deelnemers op de participatieladder welke een significante en sterke stijging laten zien voor veel

deelnemers waardoor de afstand tot de arbeidsmarkt is verkleind. Deze stijging in sociale participatie was voor de gemeente ook een belangrijk doel.

De lijntjes van de secundaire uitkomstmaten over de tijd laten bijna allemaal een stijging zien bij alle treatments en dus ook bij de onderzoeksgroep. Zo blijkt ten aanzien van de uitkomstmaat welbevinden dat de tevredenheid met leven van mensen die een 6 of lager scoren (waaronder bijstandsgerechtigden) de laatste jaren afneemt terwijl het niveau van welbevinden in het experiment stijgt voor alle deelnemers. Ook bij sociaal vertrouwen zien we een stijging bij alle groepen. Nader onderzoek met een andere vergelijkingsgroep is dus nodig om tot definitieve uitspraken te komen over de uitkomsten op deze uitkomstmaten. De komende maanden willen we daartoe een poging wagen wanneer we de cijfers gaan analyseren op de CBS-microdata bestanden.

Beleidsvragen

Niet alle beleidsvragen die aan het begin zijn geformuleerd kunnen definitief worden beantwoord op dit moment. We laten ze hieronder kort de revue passeren:

- *De eerste beleidsvraag had betrekking op het effect van eigen regie en ontheffing van de re-integratieverplichtingen op uitstroom naar werk.*

Op grond van de ruwe data vonden we hogere uitstroompercentages naar voltijds werk bij de ontheffingsgroep, die in Wageningen eigen regiegroep of 'zelf in actie' wordt genoemd, maar ook bij extra begeleiding in vergelijking met de onderzoek -of controlegroep (4-5%) en de referentiegroep van niet-deelnemers (11-13%)¹⁶. Na correctie voor verschillen in achtergrondkenmerken zijn de verschillen met de controlegroep niet meer significant. Het verschil tussen eigen regie en extra begeleiding samen en de controlegroep is net wel significant. In vergelijking met de referentiegroep vinden we wel een significant effect bij eigen regie en bij extra begeleiding. De verschillen tussen de drie treatment groepen zijn niet significant. Voor mensen die in het zogenaamde 'granieten' bestand zitten is in feite al sprake van ontheffing van verplichtingen. Er is nog maar sporadisch contact met de (werk)consulent en velen hebben ook ontheffing van de sollicitatie -en arbeidsplicht. In de eigen regiegroep meten we dus niet zozeer ontheffing als wel de combinatie van ontheffing en van begeleiding op zelfredzaamheid (het leren van zelfredzaamheid). Alle uitkomstmaten wijzen in de goede richting maar zoals eerder vastgesteld zijn er geen significante effecten van elk van de treatments in vergelijking met de controlegroep, mogelijk omdat de reguliere begeleiding in de onderzoeksgroep zelf is veranderd en meer is gaan lijken op de andere treatments. De voltijdse uitstroom naar werk is van alle drie groepen het hoogste in de eigen regiegroep, bijna 13% hoger dan in de referentiegroep die mogelijk een betere vergelijkingsgroep is. Het verschil in uitstroom naar voltijds werk is na controle op achtergrondkenmerken significant in de ontheffingsgroep vergelijking met de referentiegroep. We hebben vervolgens berekend dat als iemand in de referentiegroep de treatment eigen regie zou hebben gekregen het extra effect op uitstroom naar voltijds werk 5,6% zou zijn geweest (berekening van het marginale effect).

¹⁶ Deze percentages hebben betrekking op de deelnemers uit het zittende bestand dus met een lopende uitkering per 1 oktober 2017 en niet de nieuwe instroom daarna tot 1 oktober 2018.

- *De tweede beleidsvraag had betrekking op het effect op de arbeidsinschakeling wanneer mensen op een positieve vraaggerichte manier worden ondersteund en intensiever en op maat worden begeleid (met name extra begeleidingsgroep).*

Deze onderzoeksvraag is hierboven deels reeds beantwoord. Er is geen verschil in uitstroomkansen naar voltijdswerk tussen elk van de treatment groepen apart en de controlegroep maar wel in vergelijking met de referentiegroep die mogelijk een betere vergelijkingsgroep is. De extra begeleidingsgroep realiseert samen met de eigen regie groep de hoogste uitstroom naar voltijd werk. De uitstroom is in de extra begeleidingsgroep in de ruwe data 11% hoger en na correctie voor selectiviteit middels logistische regressie 4,8% hoger (berekening marginaal effect) dan bij de niet-deelnemers. Nemen we eigen regie en extra begeleiding samen en controleren we voor selectiviteit in vooral de langere duren middels matching, dan is de uitstroom naar werk nog steeds zo'n 5 tot 6% hoger in beide treatment groepen samen vergeleken met de niet-deelnemers die de reguliere dienstverlening kregen.

- Een derde vraag die de gemeente stelde was in hoeverre met vrijlating van verdiensten conform artikel 83 van de Participatiewet mensen meer voltijds dan wel meer in deeltijdwerk aan de slag zullen gaan en/of het aantal uren uitbreiden. De ruwe data (BUS-data en gegevens uit de vragenlijsten van de deelnemers) laten zien dat de vrijlatingsgroep minder goed scoort op uitstroom naar voltijdwerk maar wel goed op deeltijdwerk. In de vrijlatingsgroep en de onderzoeksgroep is het aandeel deeltijdwerk het hoogst is maar het aandeel voltijdwerk het laagst. Echter, het aandeel voltijdwerk is niet significant verschillend van de onderzoeksgroep¹⁷. Dat is ook logisch want de vrijlatingsgroep heeft verder de reguliere dienstverlening gekregen. Vrijlating lijkt dus wel positief uit te werken op het vinden van deeltijdwerk. In hoeverre er substitutie tussen deeltijd en voltijd werk is opgetreden is zonder nadere analyse niet duidelijk. Door zelfselectie bleek de deelnemersgroep niet helemaal vergelijkbaar met de niet-deelnemers vooral naar duur in de bijstand. Daarom hebben we 'matching' technieken toegepast om daarvoor te corrigeren. Dan blijkt dat de eigen regiegroep en de extra begeleidingsgroep samen een 5-6% hogere uitstroom naar voltijdwerk realiseren dan de groep van niet-deelnemers die de reguliere dienstverlening kregen. We vinden voorts een negatief treatment effect voor vrijwilligerswerk en mantelzorg in deze vrijlatingsgroep, waarbij door toename van het aantal uren besteed aan betaald werk minder uren werden besteed aan vrijwilligerswerk en mantelzorg. Indien uitstroom naar werk het uitgangspunt is, is dit echter een positief treatment effect en geen negatief effect.

- *Een vierde beleidsvraag voor de gemeente was of mensen die deelnemen aan het experiment meer ruimte ontwikkelen om eigen initiatief te ontplooiën, meer in brede zin gaan participeren, en wat het effect hiervan is op de arbeidsinschakeling (eigen regie, extra begeleiding, vrijlating).*

Deze vraag verwijst naar het effect op maatschappelijke participatie. We vinden op grond van de analyse van de deelnemersvragenlijsten geen significante effecten op de beleving van sociale participatie of op de contacten in het sociaal netwerk maar wel op de deelname aan vrijwilligerswerk en mantelzorg. Het effect is negatief, hetgeen betekent dat mensen hun uren besteed aan

¹⁷ In het eindverslag van de projecteider is meer in detail ingegaan op deeltijdwerk op basis van informatie over inkomsten in de BUS. Uitstroomcijfers naar deeltijd zullen in de aanvullende rapportage worden opgenomen op basis van analyses op de micro-data.

vrijwilligerswerk en mantelzorg afbouwen wanneer men voltijds of in deeltijd aan de slag gaat, maar dit is vanuit het perspectief van sociale participatie positief omdat hun sociale participatie daardoor juist toeneemt in plaats van afneemt. Voorts vinden we significante en positieve treatment effecten op subjectieve en mentale gezondheid. Eerder constateerden we reeds dat bijna 60% van de deelnemers aangeven een zwakke fysieke en vooral mentale gezondheid te bezitten waardoor de kloof tot de reguliere arbeidsmarkt erg groot is. De andere bejegening helpt kennelijk de deelnemer om hierin stappen te zetten. Dit is vanuit het oogpunt van sociale participatie minstens zo belangrijk omdat een zwakke gezondheid ertoe leidt dat men minder gemakkelijk het sociale netwerk kan onderhouden. Op grond van de procesevaluatie (consulent vragenlijst) vindt er bovendien voor de meeste deelnemers een stijging plaats op de participatieladder. We vinden geen verschillen tussen de treatments eigen regie en extra begeleiding ten opzichte van de onderzoeksgroep maar wel een significante toename in de tijd voor alle groepen. Consulents zijn zelf het meeste positief over de extra begeleidingsgroep en het minst over eigen regie. De extra begeleidingsgroep scoort volgens hen op bijna alle aspecten positief ten opzichte van de reguliere begeleiding, zoals op zelfredzaamheid, zelf in actie komen, het verwachte effect op de deelnemer, het leereffect voor de deelnemer en het leereffect voor de consulent. In de consulent vragenlijsten en de focusgroep gesprekken komt vooral terug dat deelnemers het meest baat hebben in de ogen van de consulents bij individuele op de persoonlijke mogelijkheden en noden afgestemde maatwerkbegeleiding zoals dat in de treatment extra begeleiding ook werd beoogd. Van eigen regie verwacht men niet veel en zelfs negatieve effecten op zelfredzaamheid, op sociale participatie en op de leereffecten voor de deelnemer.

- *Een belangrijke vraag voor de gemeente was wat de neveneffecten zijn van de ruimere handelingsvrijheid op maatschappelijke participatie, maar ook op zelfredzaamheid, welbevinden en gezondheid (eigen regie, intensieve begeleiding).*

De bevindingen op maatschappelijke participatie, welbevinden en gezondheid zijn hiervoor reeds aan de orde geweest. In verband met zelfredzaamheid vinden we bij de analyse van de deelnemersvragenlijsten geen effect van de treatments op zelfredzaamheid. In de procesevaluatie geven klantmanagers aan dat zij vooral bij extra begeleiding en vrijlating positieve effecten op de zelfredzaamheid van de deelnemer ervaren (zelf in actie komen, zelf doelen stellen) ten opzichte van de controlegroep. Ook vinden we dat het welbevinden en het sociaal vertrouwen in de tijd toeneemt maar bij alle groepen. Gelet op de resultaten van ander onderzoek dat het welbevinden bij mensen met vergelijkbare scores op tevredenheid leven en geluk eerder is afgenomen de laatste jaren dan toegenomen kan sprake zijn van een treatment effect maar daarvoor is nader onderzoek nodig met een andere vergelijkingsgroep.

- *In het experiment wordt onderzocht wat de uitkomsten over de tijd gezien zijn van meer handelingsvrijheid en een intensievere begeleiding op maat. Dat vraagt een andere manier van ondersteuning en andere kwalificaties van de consulents. De gemeente wil graag weten hoe dit proces verloopt. Worden de treatments uitgevoerd zoals beoogd en in welke mate zijn ze onderscheidend? Wat werkt in de uitvoering en wat werkt niet? Welke condities, aspecten maken het verschil en wat betekent dit voor de implementatie van de uitkomsten (procesevaluatie).*

Het bleek lastig om de treatments zoals beoogd uit te voeren. Het onderscheid tussen eigen regie en intensieve begeleiding was bij de start niet duidelijk voor de klantmanagers. Of dit een ‘theory failure’

of 'implementation failure' is, is moeilijk te zeggen. Wij denken beide. In de oratie van Blonk (2018) wordt opgemerkt dat er weinig literatuur is over het integratieve gedragsmodel wat betreft handvatten voor de professional. Wellicht heeft dat een rol gespeeld en waren we nog niet zo ver in het uitkristalliseren van dit model in de praktijk. Ook in de uitvoering was er geen methodiek om de treatments inhoud en vorm te geven. Het onderscheidende van de treatments was daardoor minder zeker in het eerste half jaar. We denken dat met name de verandering bij de reguliere treatment duidt op een experiment effect waarvoor we niet konden corrigeren. De 'andere bejegening' maakt ook volgens de onderzoekers het verschil hoewel er geen meting daarvan bij de deelnemers heeft plaatsgevonden. Op grond van de analyse van de deelnemers en consultant vragenlijsten, de gesprekken met consultants en onze eigen waarnemingen is er wel een duidelijk beeld ontstaan. Veel deelnemers, vooral degenen die langer in de bijstand zitten, zijn niet op zoek naar betaald werk maar vooral bezig met 'overleven' en er het beste van te maken gezien de zwakke gezondheid, en hun sociale en financiële problemen. Ze zijn op zoek naar 'bijstand' voor deze problemen, naar aandacht voor hun mogelijkheden en noden, een 'luisterend oor' en het 'goede gesprek' waardoor bij hen zelf beweging en perspectief ontstaat. De deelnemers zelf lijken vooral vertrouwen en een andere persoonlijke en 'mensgerichte' bejegening in plaats van een controlerende benadering belangrijk te vinden en te waarderen in het experiment. Tijd of aandacht alleen is niet voldoende, het gaat om de intenties en kwaliteit van het gesprek, en het perspectief dat kan worden geboden. Daarvoor is naar onze visie vakmanschap nodig en methodisch handelen om die beweging en perspectief te kunnen bieden. Op basis van de analyses op de deelnemers -en procesevaluatie vragenlijsten en vooral ook de focusgroep gesprekken denken we dat de werkzame bestanddelen zijn: een 'mensgerichte' of vraaggerichte bejegening; vertrouwen en aandacht geven maar ook loslaten, bouwen aan een respectvolle relatie, wederkerigheid, maatwerk en vraaggerichte ondersteuning maar zeker ook vakmanschap en methodisch handelen. De sleutels tot succes zijn het opbouwen van een persoonlijke relatie, vertrouwen geven en aandacht, meer tijd voor het 'echte' gesprek, sturen op intrinsieke motivatie, een luisterend oor en loslaten. Vanuit de consultants wordt vooral waardering uitgesproken voor extra begeleiding maar er is in onze visie alle reden om met name ook ontheffing en eigen regie een grotere plaats te geven in de begeleidingsmethodiek. Vanuit de deelnemers worden andere accenten gelegd die ook goed aansluiten bij eigen regie evenals bij intensieve begeleiding waar het gaat om beweging en perspectief, intrinsieke motivatie, handelingsvrijheid en een op vertrouwen gebaseerde bejegening. Op grond van de onderzoeksbevindingen valt de eigen regiegroep op vanwege de gunstige uitstroomcijfers naar voltijd en vast werk. Hoewel nader onderzoek nodig is denken we dat de werkzame bestanddelen daarvoor zijn geweest vertrouwen, handelingsvrijheid en intrinsieke motivatie.

Andere onderzoeksvragen die niet zijn gesteld

Naast deze vragen zijn nog vele andere vragen te stellen die niet zijn beantwoord. Het experiment heeft een rijkdom aan data opgeleverd waarvan in dit voorlopig eindverslag slechts een klein deel is geanalyseerd. In het aanvullend verslag zoomen we nog nader in op uitstroom naar deeltijdwerk en een kwalitatieve vergelijking van de resultaten met andere gemeenten maar niet op andere vraagstukken zoals armoede, duurzame uitstroom naar verschillende typen van werk, de invloed van gezondheid op uitstroom naar werk en op loopbanen en dergelijke.

In Wageningen bleek door de kleine aantallen sprake te zijn van een power probleem waardoor resultaten soms net niet significant werden. Om dit op te lossen kan een kwantitatieve vergelijking met andere gemeenten soelaas bieden omdat dan de aantallen en daarmee de power wordt vergroot. Op dit moment is niet voorzien in vergelijkend onderzoek waarin dit power probleem kan worden opgelost of waarin inhoudelijke vragen vanuit vergelijkend perspectief worden geanalyseerd. We hopen tot slot dat het verslag aanleiding geeft om ook deze vragen te gaan stellen en beantwoorden.

Literatuur en referenties

- Abbring, J.H., G.J. van den Berg, P. Mullenders en J.C. van Ours (1996). Sancties in de ww; een werkend perspectief. In: *Economisch Statistische Berichten*, jg. 81, p. 750-753.
- Abbring, J.H., G.J. van den Berg en J.C. van Ours (2005). The Effect of Unemployment Insurance. *Sanctions on the Transition Rate from Unemployment to Employment*. In: *Economic Journal*, jg. 115, nr. 505, p. 602-630.
- Atkinson, A. B. (1996) 'The case for a participation income', *Political Quarterly*, 67, 1, 67–70.
- Bigotta, M., Bonoli, G., Fossati, F., Lalive, R., & Oesch, D. (2018). Helping social assistance recipients find jobs: Evidence from a controlled trial (ESPANet working paper, University of Lausanne, Switzerland)
- Bohnet, I., Frey, B. S. and Huck, S. (2001) 'More order with less law: on contract enforcement, trust and crowding', *The American Political Science Review*, 95, 1, 131–44.
- Bond, G. R., Drake, R. E., & Luciano, A. (2015). Employment and educational outcomes in early intervention programs for early psychosis: A systematic review. *Epidemiology and Psychiatric Sciences*, 24(5), 446–457. <https://doi.org/10.1017/S2045796014000419>
- Blonk, R.W.B., M.W. van, Twuijver, H.A. van de Ven en A.M. Hazelzet (2015). Quickscan wetenschappelijke literatuur gemeentelijke uitvoeringspraktijk. Leiden: tno.
- Blonk, Roland (2018), We zijn nog maar net begonnen, Openbare rede prof. dr. R. Blonk, Tilburg University op 6 april 2018.
- Card, D., Kluve J., en A. Weber (2010), Active labour market policy evaluations: a meta-analysis, *The Economic Journal*, 120 (November), F452–F477.
- Card, D., Kluve, J., & Weber, A. (2015). What works? A meta analysis of recent active labor market program evaluations (No. w21431). National Bureau of Economic Research.
- CBS (2017), De Geluk meter, CBS Statline, Voorburg/Heerlen.
- Cuelenaere, B., E. van den Brink en J. Leenen (2019). Ervaringen gemeenten met Participatiewet. Rapportage derde meting. Tilburg: Centerdata.
- Deaton, A., Nancy Cartwright (2017). Understanding and misunderstanding randomized controlled trials, *Social Science & Medicine*, Vol. 210(2018):2-21(<https://doi.org/10.1016/j.socscimed.2017.12.005>)
- Deci, E. L. and Ryan, R. M. (1985) *Intrinsic Motivation and Self-Determination in Human Behavior*, New York: Plenum.
- Delos, S. en J. Oude Alink (2016), Experiment regelluwe bijstandsverlening, Brief aan portefeuillehouders, Werk en Inkomen, Gemeente Oss.
- Diener, E., Suh, E. M., Lucas, R. E., & Smith, H. L. (1999). Subjective well-being: Three decade of progress. *Psychological Bulletin*, 125(2), 276–302. <https://doi.org/10.1037/0033-2909.125.2>.
- Dohmen, T., A. Falk, D. Huffman and U. Sunde (2008), Representative trust and reciprocity: Prevalence and determinants, *Economic Inquiry*, 46 (1), 84–90.
- Fehr, E. & Gächter, S. (2000). Fairness and Retaliation: the economics of reciprocity. *The Journal of Economic Perspectives*, 14(3), 159-181.
- Forget, E. L. (2011). The town with no poverty: the health effects of a Canadian guaranteed annual income field experiment. *Canadian Public Policy*, 37(3), 283-305.
- Fehr, E. and Schmidt, K. M. (2003) 'Theories of fairness and reciprocity: evidence and economic

- applications', *Advances in Economics and Econometrics*, Econometric Society, Eighth World Congress, 1, 208–57.
- Fishbein, M., & Ajzen, I. (2010). *Predicting and changing behavior: The reasoned action approach*. New York: Psychology Press.
- Frey, B. S. and Jegen, R. (2001) 'Motivation crowding theory', *Journal of Economic Surveys*, 15, 5, 589–611.
- Goodin, R. E., Headey, B., Muffels, R. and Dirven, H.-J. (1999) *The Real Worlds of Welfare Capitalism*, Cambridge: Cambridge University Press.
- Groot, N. de en B. van der Klaauw (2019). The effects of reducing the entitlement period to unemployment insurance benefits. In: *Labour Economics*, jg. 57, p. 195-208.
- Groot, L. and van der Veen, R. (2000) 'Clues and leads in the policy debate on basic income in the Netherlands', in R. van der Veen and L. Groot (eds.), *Basic Income on the Agenda: Policy Objectives*
- Groot, L., Muffels, R. J., & Verlaat, T. (2019). Welfare states' social investment strategies and the emergence of Dutch experiments on a minimum income guarantee, *Social Policy & Society* (2019) 18:2, 277–287, Cambridge University Press
- Headey, B., Muffels, R.J.A., & Wagner, G. (2010). Long-running German panel survey shows that personal and economic choices, not just genes, matter for happiness. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)*, 107(42), 17922-17926
- Hemerijck, A. (2013) *Changing Welfare States*, Oxford: Oxford University Press.
- Hemerijck, A. (ed.) (2017) *The Uses of Social Investment*, Oxford: Oxford University Press.
- Hoff Stella Hoff Bart van Hulst Jean Marie Wildeboer Schut Benedikt Goderis (2019), *Armoedekaart 2019* (<http://digitaal.scp.nl/armoedeinkaat2016>), SCP-Den Haag.
- Koning, P. (2012a). Beter een stok dan een wortel. In: J. Ultermark, A.J. Gielen en M. Ham (red.), *Wat werkt nu werkelijk? Politiek en praktijk van sociale interventies* (p. 105-120). Amsterdam: Van Genep.
- Kluve, J., (2010). The effectiveness of European active labor market programs. *Labour Econ.* 17, 904–918. doi:10.1016/j.labeco.2010.02.004.
- Kok, L., D. Hollanders en J.P. Hop (2006). *Kosten en baten van reïntegratie*. Amsterdam: SEO Economisch Onderzoek.
- Koning, J. de, A. Gelderblom, K. Zandvliet en L. van den Boom (2005). *Effectiviteit van reïntegratie. De stand van zaken. Literatuuronderzoek*. Den Haag: Ministerie van Sociale Zaken en Werkgelegenheid.
- Koning, J. de, P. de Hek, L. Mallee, F. Rosing en M. Groenewoud (2014). *Uitkomsten en ervaringen experimenten netto-effectiviteit re-integratie*. Rotterdam/Amsterdam: seor/Regioplan Beleidsonderzoek/Epsilon Research.
- Kremer, M., J. van der Meer, en M. Ham (2017). Van werk naar welbevinden – werkt de zachte hand in de bijstand? *Sociale Vraagstukken* (<https://www.socialevraagstukken.nl/van-werk-naarwelbevinden-werkt-de-zachte-hand-in-de-bijstand/>)
- Mani, A., Mullainathan, S., Shafir, E. and Zhao, J. (2013) 'Poverty Impedes Cognitive Function', *Science*, August, 976–80.

- Marshall, T., Goldberg, R. W., Braude, L., Dougherty, R. H., Daniels, A. S., Sushmita Shoma Ghose, E. D., et al. (2014). Supported employment: Assessing the evidence. *Psychiatric Services*, 65(1), 16–27. <https://doi.org/10.1176/appi.ps.201300262>
- Morel, N., Palier, B. and Palme, J. (2012) *Towards a Social Investment Welfare State? Ideas, Policies and Challenges*, Bristol: The Policy Press.
- Muffels, R. and B. Headey (2013). Capabilities and choices: Do they make Sense for understanding objective and subjective well-being? An empirical test of Sen's capability framework on German and British panel data." *Social Indicators Research*, **110**(3): 1159-1185.
- Muffels, R. & J. Van der Klink (2017), *Offerte Vertrouwensexperimenten Wageningen, Oss, Renkum, Epe en Mierlo-Geldrop: Onderzoekdesign en kostenraming onderzoek*, Tilburg University, 10 mei 2017.
- Muffels, R. & S. Vos (2018), *Voortgangsrapport Vertrouwensexperiment Wageningen: No. 1*, 8 april 2018, Tilburg University, p. 1-18.
- Muffels, R. & S. Vos (2018), *Voortgangsrapport Vertrouwensexperiment Wageningen: No. 2*, 8 oktober 2018, Tilburg University, p. 1-29.
- Muffels, R. & S. Vos (2018), *Voortgangsrapport Vertrouwensexperiment Wageningen: No. 3*, 8 april 2019, Tilburg University, p. 1-35.
- Mullainathan, S. and Shafir, E. (2013) *Scarcity – The True Cost of Not Having Enough*. London: Penguin Books.
- Nazar, Amir (2010), *Wat werkt? Wat weten we over effectiviteit?*, Ministerie van Sociale Zaken en Werkgelegenheid (<http://www.onderzoekwerkeninkomen.nl/rapporten/xbdiusf2>).
- Ryan, R. M., Kuhl, J. and Deci, E. L. (1997) 'Nature and autonomy: an organizational view of social and neurobiological aspects of self-regulation in behavior and development', *Development and Psychopathology*, 9, 4, 701–28.
- Sen, A. K. (2004) 'Capabilities, lists, and public reason: continuing the conversation', *Feminist Economics*, 10, 3, 77–80.
- Sen, A. K. (1999). *Development as Freedom*, New York, Knopf.
- Soentken, M., van Hooren, F. and Rice, D. (2017) 'The impact of social investment reforms on income and activation in the Netherlands', in A. Hemerijck (ed.), *The Uses of Social Investment*, OUP Oxford, Kindle Edition, 475.
- Van Echtelt, Patricia e.a. (2019), *Eindevaluatie van de Participatiewet*, SCP.
- Thaler, R. H. and Sunstein, C. R. (2008) *Nudge: Improving Decisions about Health, Wealth, and Happiness*, New Haven, Conn.: Yale University Press.
- Van der Klink JJ, Bultmann U, Burdorf A, Schaufeli WB, Zijlstra FR, Abma FI, Brouwer S, van der Wilt GJ. Sustainable employability--definition, conceptualization, and implications: A perspective based on the capability approach. *Scand J Work Environ Health* 2016; 42:71-9.

Bijlage A. Resultaten regressieanalyses deelnemersvragenlijsten

Tabel A.1. Niet-gestandaardiseerde regressie coëfficiënten (*b*) geschat op basis van OLS of logit regressie model (armoede)

Uitkomstmaten	Eigen regie	Extra begeleiding	Vrijlating	Meting 2	Meting 3
Welbevinden	0	0	0	.35**	.38**
Subjectieve gezondheid	0	.47*	0	0	.33*
Mentale gezondheid	0	0	0	0	.27*
Keuzemogelijkheden	0	0	0	0	-.28*
Perceptie sociale participatie	0	0	0	0	0
Sociaal netwerk	0	0	0	-1.58***	0
Vrijwilligerswerk/mantelzorg	0	-3.3*	-4.0*	-2.1*	-2.77**
Zelfredzaamheid	0	0	0	0	0
Vertrouwen instituties	-.39 ²⁾	0	0	-.74***	-.75***
Sociaal vertrouwen	0	0	0	0	0.29 ¹⁾
Inkomensarmoede	-.69*	0	0	0	0
Deprivatiearmoede (3 uit 5)	0	0	0	0	0

Noot 1: *** $p < 0,01$; ** $p < 0,05$; * $p < 0,10$ ¹⁾ Effecten net niet significant $p < 0,15$.

Noot 2: Voor de testen van de effecten van eigen regie en extra begeleiding was de onderzoekgroep de referentiecategorie waartegen werd getest. Bij de effecten van meting 2 en 3 was de nulmeting de referentiecategorie.

Noot 3: In dit model waren de uitkomstmaten in de eerste kolom de afhankelijke variabelen *Y*. De belangrijkste onafhankelijke variabelen waren treatment en meting. Verder is gecontroleerd voor zittend bestand/nieuwe instroom, nummer klantmanager en een aantal controlevariabelen (leeftijd, geslacht, opleiding, duur uitkering, leefvorm, en huisvesting). Deze uitkomsten zijn de test voor de verschillen van de buitenste rij en kolomtotalen van tabellen 3.1.1 t/m 3.1.12. Bij een 0 in de tabel, kon er geen significant effect gevonden worden.

Bron: Deelnemers vragenlijsten in de drie metingen Vertrouwensexperiment Wageningen, Tilburg University, 2020.

Tabel A.2: Regressieuitkomsten interactiemodel treatment en meting productevaluatie (significante treatment effecten per meting)

Uitkomstmaten	Eigen regie			Extra begeleiding			Vrijlating			Onderzoekgroep		
	M0	M2	M3	M0	M2	M3	M0	M2	M3	M0	M2	M3
Welbevinden	0	.82***	0	0	0	0	0	0	0	0	0	.82***
Subjectieve gezondheid	0	.68 ¹⁾	0	0	.76 ¹⁾	1.02**	0	0	0	0	0	0
Mentale gezondheid	0	.76**	0	0	0	.78**	0	0	.56 ¹⁾	0	0	.68**
Keuzemogelijkheden	0	0	0	0	0	0	0	0	0	0	0	0
Perceptie sociale participatie	0	0	0	0	0	0	0	0	0	0	0	0
Sociale netwerk	0	-1.7***	0	0	-1.0*	0	0	-1.3**	0	0	-1.4***	0
Vrijwilligerswerk/ mantelzorg	0	-4.7 ¹⁾	-5.6*	0	-4.7 ¹⁾	0	0	-5.4*	-7.0**	0	0	0
Zelfredzaamheid	0	0	0	0	0	0	0	0	0	0	0	0
Vertrouwen instituties	0	-1.2***	-1.0***	0	-.88**	-.86**	0	-.97***	-.89***	0	-.67***	-1.0***
Sociaal vertrouwen	0	0	0	0	0	0	0	0	0	0	0	0
Inkomensarmoede	-1.45***	-1.04***	-0.95**	0	-1.10*	0	0	-1.35**	0	0	0	-1.22**
Deprivatiearmoede (3 uit 5)	0	0	0	0	0	0	0	0	0	0	0	0

Noot 1: ***p<0,01; **p<0,05; *p<0,10. ¹⁾ Effecten net niet significant p<0.13.

Noot 2: voor de effecten van de onderzoekgroep in meting 2 en 3 is de onderzoekgroep in de nulmeting de referentiecategorie. Voor de effecten van eigen regie in de nulmeting is de referentie de onderzoekgroep in de nulmeting en voor metingen 2 en 3 geeft de parameter het effect weer van de groep in die meting dat bovenop het effect komt van het effect van de eigen regiegroep in de nulmeting waar de onderzoekgroep de referentie is. Hetzelfde geldt voor het effect van intensieve begeleiding.

Noot 3: In dit model zijn de concepten in de eerste kolom de afhankelijke variabelen Y. De belangrijkste onafhankelijke variabelen zijn treatment en meting en de interactie tussen deze twee. Verder is gecontroleerd voor zittend bestand/nieuwe instroom, consulentnummer en een aantal controlevariabelen (leeftijd, geslacht, opleiding, duur uitkering, leefvorm, en huisvesting). Deze uitkomsten zijn de test voor de verschillen van de tussenliggende cellen van tabellen 3.1.1 t/m 3.1.12. Bij een 0 in de tabel, is er geen significant effect gevonden.

BIJLAGE B. Resultaten regressieanalyses consulent vragenlijsten

Tabel B.1: niet-gestandaardiseerde regressie coëfficiënten (*b*) geschat op basis van OLS-regressiemodel per treatment groep over de tijd en per meting

Uitkomstmaten	Eigen regie	Extra begeleiding	Vrijlating	Meting 2	Meting 3
Motivatie eigen treatment	0,7**	1**	0	-0,6**	0,43*
Motivatie hoogst bereikbare doel	0,7**	1,1***	0	0	1,3***
Realistische eisen	0	1***	0,6**	0	1***
Zelf in actie komen	0	1,2***	0,8**	0	1,5***
Zelf doelen stellen	0,5*	1,5***	1***	0	0,8***
Verwacht effect treatment	0	1,2***	0	-0,7**	0
Sociale participatie	0	0,6**	0,8***	0	0,7***
Participatieladder	0,3**	0	0,3**	1***	1,2***
Leereffect deelnemer	0	1,1***	0,5**	0	0,5**
Leereffect consulent	1,1***	1,2***	1,1***	0	0

Noot 1: *** $p < 0,01$; ** $p < 0,05$; * $p < 0,10$

Noot 2: Voor het testen van de effecten van eigen regie, extra begeleiding en vrijlating is de onderzoekgroep de referentiecategorie. Bij de effecten van meting 2 en 3 is de onderzoekgroep in de nulmeting de referentiecategorie.

Noot 3: In dit model zijn de uitkomstmaten in de eerste kolom de afhankelijke variabelen *Y*. De belangrijkste onafhankelijke variabelen zijn treatment en meting. Verder is gecontroleerd voor zittend bestand/nieuwe instroom, of iemand afgevallen is. Deze uitkomsten geven de test weer voor de verschillen van de buitenste rij en kolomtotalen van tabellen 3.3.1 t/m 3.3.10. Bij 0 (niet significant) in de tabel, is er geen significant effect gevonden.

Bron: Vragenlijsten procesevaluatie Vertrouwensexperiment Wageningen, Tilburg University, 2020.

Tabel B.2: niet-gestandaardiseerde regressie coëfficiënten (b) geschat op basis van OLS-interactie model (treatment effecten per meting)

Uitkomstmaten	Onderzoeksgroep		Eigen regie			Extra begeleiding			Vrijlating		
	M2	M3	M1	M2	M3	M1	M2	M3	M1	M2	M3
Motivatie eigen treatment	0	1*	1,3**	0	-1,1*	1,1**	0	0	0	0	0
Motivatie hoogst bereikbare doel	0	1,5**	0,9**	0	0	1,4**	0	0	0	0	0
Realistische eisen	0	1,4**	0,7*	0	0	1,1**	0	0	0,8**	0	0
Zelf in actie komen	0	1,6**	0	0	0	1,2**	0	0	0	0	0
Zelf doelen stellen	0	1,2**	0,7*	0	0	1,4**	0	0	1**	0	0
Verwacht effect treatment	-1,3**	0	0	0	0	1,1**	0	0	-0,7*	1,4**	0
Sociale participatie	0	0,9**	0	0	0	0,8**	0	0	1,2***	0	-1*
Participatieladder	1,2***	1,4***	0,4*	0	0	0	0	0	0,6**	0	-0,6*
Leereffect deelnemer	0	0	0	0	0	1,1**	0	0	0	1,1**	0
Leereffect consulent	-1,2**	0	0	1,4**	0	0	2,1***	1,9**	0	1,1*	0

Noot 1: *** $p < 0,01$; ** $p < 0,05$; * $p < 0,10$

Noot 2: Voor het testen van de effecten van eigen regie, extra begeleiding en vrijlating is de onderzoeksgroep de referentiecategorie. Voor de effecten van eigen regie in de nulmeting is de referentie de onderzoeksgroep in de nulmeting en voor de vervolgmetingen geeft de parameter het effect weer van de groep in die meting dat bovenop het effect komt van de eigen regiegroep in de nulmeting waar de onderzoeksgroep de referentie is. Hetzelfde geldt voor het effect van extra begeleiding en vrijlating.

Noot 3: In dit model waren de uitkomstmaten in de eerste kolom de afhankelijke variabelen Y. De belangrijkste onafhankelijke variabelen waren treatment en meting en de interactie tussen deze twee. Verder is gecontroleerd voor zittend bestand/nieuwe instroom en of iemand een afmelder is. Deze uitkomsten zijn de test voor de verschillen van de tussenliggende cellen van tabellen 3.2.1 t/m 3.2.10. Bij 0 (niet significant) in de tabel, kon er geen significant effect gevonden worden.

Bron: Vragenlijsten nulmeting en vervolgmetingen procesevaluatie Vertrouwensexperiment Wageningen, Tilburg University, 2020.