

University of Groningen

Perspectieven voor interventies in de bijstand

Edzes, Arjen; Muffels, R. J. A.; Zulkarnain, Alice; Jongen, Egbert; Sanders, Mark W. J. L.; Verlaat, T. L.; Betkó, János; De Boer, Henk-Wim; Bolhaar, Jonneke; Gramberg, Peter

Published in:
Tijdschrift voor Arbeidsvraagstukken

DOI:
[10.5117/TVA2021.3.007.EDZE](https://doi.org/10.5117/TVA2021.3.007.EDZE)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Final author's version (accepted by publisher, after peer review)

Publication date:
2021

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Edzes, A., Muffels, R. J. A., Zulkarnain, A., Jongen, E., Sanders, M. W. J. L., Verlaat, T. L., Betkó, J., De Boer, H.-W., Bolhaar, J., Gramberg, P., Groot, L., Rosenkranz, S., Rijnks, R., Spierings, N., & Venhorst, V. (2021). Perspectieven voor interventies in de bijstand: Experimenten Participatiewet. *Tijdschrift voor Arbeidsvraagstukken*, 37(3), 327-356. <https://doi.org/10.5117/TVA2021.3.007.EDZE>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Perspectieven voor interventies in de bijstand

Experimenten Participatiewet

Arjen Edzes, Ruud Muffels, Alice Zulkarnain, Egbert Jongen, Mark Sanders, Timo Verlaat, János Betkó, Henk-Wim de Boer, Jonneke Bolhaar, Peter Gramberg, Loek Groot, Stephanie Rosenkranz, Richard Rijnks, Niels Spierings, Viktor Venhorst

Arjen Edzes is werkzaam bij de Rijksuniversiteit Groningen. E-mail: a.j.e.edzes@rug.nl. Ruud Muffels is werkzaam bij de Universiteit Tilburg. Alice Zulkarnain is werkzaam bij het Centraal Planbureau. Egbert Jongen is werkzaam bij het Centraal Planbureau en de Universiteit Leiden. Mark Sanders is werkzaam bij de Universiteit Utrecht en Universiteit Maastricht. Timo Verlaat is werkzaam bij de Universiteit Utrecht. János Betkó is werkzaam bij de Radboud Universiteit Nijmegen. Henk-Wim de Boer is werkzaam bij het Centraal Planbureau. Jonneke Bolhaar is werkzaam bij het Centraal Planbureau. Peter Gramberg is werkzaam bij Hogeschool Saxion. Loek Groot is werkzaam bij de Universiteit Utrecht. Stephanie Rosenkranz is werkzaam bij de Universiteit Utrecht. Richard Rijnks is werkzaam bij University College Cork. Niels Spierings is werkzaam bij de Radboud Universiteit Nijmegen. Viktor Venhorst is werkzaam bij de Rijksuniversiteit Groningen.

Samenvatting

Welke handelings- en interventiestrategieën van gemeenten dragen bij aan perspectief op werk en een verbetering van het welbevinden en vertrouwen van mensen in de bijstand? Zes gemeenten – Groningen, Utrecht, Tilburg, Wageningen, Deventer en Nijmegen – hebben de afgelopen jaren (1 oktober 2017 – 31 december 2019) unieke Randomised Controlled Trials uitgevoerd in de bijstand. De volgende (combinatie van) interventies zijn onderzocht: Ontheffing van re-integratieverplichtingen, Intensivering van begeleiding en Vrijlating van bijverdiensten. Er is zowel gekeken naar uitstroom naar werk als naar baanzoekintensiteit, welbevinden, zelfredzaamheid en sociaal vertrouwen. De uitstroom naar werk is bij alle interventies niet lager dan de huidige aanpak, terwijl de uitstroom naar deeltijdwerk in sommige gemeenten hoger is. Vooral *Intensivering* op maat en *Vrijlating* kan uitstroom naar (minimaal deeltijd)werk vergroten. Voor wat betreft de effecten op baanzoekintensiteit, zelfredzaamheid, welbevinden en vertrouwen, is het beeld diffuus. We vinden kleine, en soms ook positieve effecten, vooral voor zelfredzaamheid en vertrouwen, maar het beeld is niet eenduidig. De kleine aantallen deelnemers in de experimenten maken het lastig statistisch significante effecten te vinden.

Key words: Bijstand, Participatiewet, Experimenten, Uitstroom

Summary

Which municipalities' action and intervention strategies contribute to job prospects and an improvement in people's well-being and confidence in social assistance benefits? Six municipalities - Groningen, Utrecht, Tilburg, Wageningen, Deventer and Nijmegen - have conducted unique Randomized Controlled Trials for social assistance benefits in recent years (October 1, 2017 - December 31, 2019). The following (combination of) interventions have been investigated: *Exemption* from reintegration obligations, *Intensification* of counselling and *Earnings release* of extra earnings. Both outflow to work and job search intensity, well-being, self-efficacy and trust were examined. The outflow to work for all interventions is not lower than the current approach, while the outflow to part-time work is higher in some municipalities. In particular, tailor-made intensification and earnings release can increase outflow to (at least part-time) work. The picture is diffuse concerning the effects on job search intensity, well-being, self-efficacy and trust. We find small, and sometimes positive effects, especially for self-reliance and confidence, but the picture is not clear-cut. The small numbers of participants in the experiments make it difficult to find statistically significant effects.

Inleiding¹

Welke handelings- en interventiestrategieën van gemeenten dragen bij aan perspectief op werk en een verbetering van het welbevinden en vertrouwen van mensen in de bijstand? Met deze vraag in het achterhoofd hebben zes gemeenten – Groningen, Utrecht, Tilburg, Wageningen, Deventer en Nijmegen - de afgelopen jaren (1 oktober 2017 – 31 december 2019) unieke Randomised Controlled Trials (RCT) uitgevoerd in de bijstand onder het Tijdelijk Besluit Experimenten Participatiewet (Ministerie van Sociale Zaken en Werkgelegenheid, 2017).² Het doel was om meer te leren over de uitvoering van de Participatiewet door de huidige praktijk te vergelijken met verschillende interventiestrategieën: meer ruimte voor eigen regie, extra begeleiding en maatwerk, en financiële prikkels via een verruiming van de bijverdienregeling.

De interventiestrategieën zijn een uitvloeisel van veranderingen in onderliggende mensbeelden en discussies over handelingstheorieën onder beleidsmakers en wetenschappers (Kampen, Sebrechtsm, Knijn & Tonkens (red.), 2020; Groot, Muffels & Verlaat, 2019; Wetenschappelijke Raad voor het Regeringsbeleid, 2017; Tiemeijer, Thomas & Prast, 2009). We onderzoeken of een verschuiving naar vertrouwen als uitgangspunt en de herwaardering van noties als eigen regie, aandacht en maatwerk in de dienstverlening tot verandering leidt op een reeks van uitkomstmaten.

Zowel in de uitvoering als in de methodologische opzet van de RCT's is nauw samengewerkt tussen de betrokken kennisinstellingen. Om de interventies met elkaar te kunnen vergelijken zijn in alle zes gemeenten vergelijkbare vragenlijsten afgenomen, dezelfde uitkomstmaten gebruikt en is het analysemodel hetzelfde. De effecten van de interventies zijn gemeten op de kans op werk en inkomen en op baanzoekintensiteit, welbevinden, zelfredzaamheid en sociaal vertrouwen. De experimenten zijn wetenschappelijk ondersteund door een commissie van deskundigen van ZonMw die tevens onafhankelijk een eigen procesevaluatie heeft uitgevoerd (De Graaf & Van Gastel, 2020). Naast de lokale studies heeft het Centraal Planbureau (CPB) een onafhankelijk onderzoek uitgevoerd naar de resultaten op de uitkomstmaat uitstroom naar werk (De Boer, Bolhaar, Jongen & Zulkarnain, 2020).

Doordat de experimenten op gemeentelijk niveau zijn vormgegeven, zijn er uiteraard verschillen tussen de contexten en het reguliere beleid, alsook de daar geïmplementeerde experimenten.

¹ Delen van dit artikel zijn ontleend aan Sanders et al. (2020) en De Boer et al. (2020) waarin zowel opzet, uitvoering als uitkomsten uitgebreid zijn beschreven. Details over de losse experimenten zijn te vinden in de lokale rapportages die over elk experiment in mei 2020 zijn uitgekomen. Een overzicht van de officieel verschenen rapportages is te vinden op: <https://www.rijksoverheid.nl/documenten/kamerstukken/2020/05/28/kamerbrief-reactie-op-uitkomsten-experimenten-participatiewet>. Geraadpleegd op 3 maart 2021.

² Naast deze zes erkende of officiële experimenten in de Participatiewet, mogelijk gemaakt door een algemene maatregel van bestuur (AMvB) en vastgelegd in art. 83 van deze wet, hebben vier andere gemeenten (Amsterdam, Apeldoorn-Epe, Oss en Geldrop-Mierlo) soortgelijke experimenten uitgevoerd maar met slechts twee van de drie interventies (onthefving en intensivering).

Deze verschillen zijn dusdanig dat de data niet in één model geanalyseerd kunnen worden. Maar de uitkomsten kunnen wel in samenhang bekeken worden, zoals we in deze bijdrage doen. Hiermee leveren de experimenten relevante informatie op voor (toekomstig) beleid en voor de discussie over de toekomst van de sociale zekerheid in Nederland.

Dit artikel gaat eerst in op de theoretische achtergronden en beleidsveronderstellingen van de interventies. Het beschrijft vervolgens de gevolgde methodologie en de RCT-opzet van de experimenten en gaat achtereenvolgens in op de effecten op werk, baanzoekintensiteit, zelfredzaamheid, welbevinden, en vertrouwen. We ronden af met onze belangrijkste conclusies.

1. Achtergrond³

Het doel van de Participatiewet is om mensen weer op eigen benen te laten staan en (snel) uit te laten stromen naar (betaald) werk of, als dat niet kan, op een andere manier te laten participeren in de samenleving. Het huidige bijstandsregime is gestoeld op het principe van wederzijdse verplichtingen. De overheid verschaft een (aanvullend) inkomen ter hoogte van het sociaal minimum. In ruil daarvoor legt de overheid verplichtingen en beperkingen op die een activerende effect beogen te hebben. De verplichtingen worden voorts met (de dreiging van) strafkortingen afgedwongen. Het onderliggende mensbeeld bij deze benadering is dat van de calculerende burger die een kosten-baten afweging maakt. In veel gemeenten leidt deze aanpak echter tot matige resultaten, wanneer we bijvoorbeeld kijken naar re-integratie (Van Echtelt et al., 2019).

De beleidstheorie achter het loslaten van verplichtingen, mensen vertrouwen te geven en hen te faciliteren (door middel van vraaggestuurde maatwerkbegeleiding) en te belonen voor eigen initiatief, kan worden gevonden in de gedragseconomie, gedragspsychologie en in sociaalpsychologische motivatietheorieën (Groot et al., 2019).

In de eerste plaats laat onderzoek naar de invloed van armoede op de 'mindset' ofwel geestesgesteldheid van mensen zien dat (financiële) schaarste en stress door armoede de cognitieve hulpbronnen van mensen verminderen (Mani, Mullainathan, Shafir & Zhao, 2013; Mullainathan & Shafir, 2013). Mensen hebben alle energie, geestkracht en inventiviteit nodig om de eindjes aan elkaar te knopen, door het ontwikkelen van zogenaamde huishoudbezuinigingsstrategieën om in de behoeften van het huishouden te voorzien. Indien financiële schaarste en het voldoen aan bijstandsverplichtingen een groot deel van de cognitieve hulpbronnen van mensen innemen, is er weinig ruimte voor andere belangrijke en cognitief uitdagende taken, zoals bijscholing, het onderhouden van het sociaal netwerk of het actief zoeken naar passend werk. Wanneer verplichtingen plaats maken voor vertrouwen, ontstaat er meer (cognitieve) ruimte voor andere taken, waaronder werkzoekgedrag, zo is de redenering (Groot et al., 2019).

³ Deze sectie van dit artikel is grotendeels ontleend aan Groot et al., 2019.

Een 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. Onderliggende waarden zijn bijvoorbeeld reciprociteit (wederkerigheid) en vertrouwen. Wederkerigheid betekent dat individuen een goede behandeling of het ontvangen van vertrouwen (een investering in sociale relaties), belonen door zich bijvoorbeeld 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 & Schmidt, 2003). Negatieve prikkels, zoals opgelegde financiële sancties en kortingen, zijn niet noodzakelijkerwijs de beste manier om mensen te prikkelen mee te werken en zich aan de regels te houden. Intussen laat economisch onderzoek zien dat er positieve korte termijn werkgelegenheidseffecten van sancties uitgaan (men accepteert sneller werk (Abbring, Van den Berg & Van Ours, 2005; Hullegie & Van Ours, 2014)). Daarnaast heeft actieve werkzoekbemiddeling vooral effect op werkhervatting bij kortdurende werkloosheid, maar minder bij langdurige werkloosheid (Koning, 2009). Er is nog weinig empirisch onderzoek verricht naar de effecten van werkzoek verplichtingen en sancties op de langere termijn. Theoretisch zou een slechte 'baanmatch' tussen werkzoekende en baan op de iets langere termijn welbevinden, gezondheid en de loopbaan negatief kunnen beïnvloeden (zie o.a. Blonk, 2018). Bevindingen uit de experimentele economie laten ook zien dat mensen in ruil voor het vertrouwen dat zij ontvangen juist extra gemotiveerd zijn en meer hun best doen om daarmee degene die hen vertrouwt te belonen en het vertrouwen niet te beschamen. Zo leidt vertrouwen dus tot gevoelens van positieve reciprociteit en daardoor ook tot blijvende inzet en toegenomen productiviteit (Bohnet, Frey & Huck, 2001).

Een derde inzicht is afkomstig uit de psychologische motivatietheorie en leert ons dat extrinsieke prikkels de intrinsieke motivatie kunnen wegnemen (Frey & Jegen, 2001). De zelfbeschikkingstheorie 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 hoger welbevinden ervaren (Deci & Ryan, 1985). Deze theorie stelt ook dat het geven van vertrouwen aan mensen een gevoel van zelfsturing oplevert, dat weer effect heeft op het baanzoekgedrag 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 (Frey et al., 2001).

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, 1999; 2004). 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 inzichten vormen de basis voor de manier waarop alternatieve interventiebenaderingen (of treatments) zijn vormgegeven binnen de experimenten. De ruimte voor eigen regie of

zelfmanagement was geïnspireerd door ideeën over de rol van zelfbeschikking, autonomie en intrinsieke motivatie voor gedrag, terwijl geïntensiverde ondersteuning op maat werd geïnspireerd door ideeën over de rol van wederkerigheid bij het geven van vertrouwen en keuzevrijheid. Ten slotte werd het ontwerp van de financiële prikkel via de vrijlatingsregeling geïnspireerd door ideeën over de impact van belonen of verstrekken van positieve financiële prikkels op het zoeken naar werk.

2. Overeenkomsten en verschillen tussen de experimenten

De experimenten bestonden grofweg uit drie typen interventies en vonden plaats in Deventer, Groningen, Nijmegen, Tilburg, Utrecht en Wageningen. Tegelijkertijd zijn de experimenten op belangrijke punten anders vormgegeven. De drie typen, of een combinatie daarvan⁴, zijn in tabel 1 weergegeven.

- A. Een tijdelijke ontheffing van de arbeids- en re-integratieverplichtingen (ontheffingsgroep).
- B. Een tijdelijke intensivering van de begeleiding van bijstandsgerechtigden (intensiveringsgroep).
- C. Een vrijlating van de inkomsten uit arbeid tot maximaal 50 procent van deze inkomsten, met een maximum van circa 200 euro per maand (vrijlatingsgroep).

De combinatie van ontheffing en intensivering was niet toegestaan. De gemeenten hanteerden verschillende instroommomenten en sommigen hadden meerdere instroommomenten (tabel 2).⁵ Ook op andere belangrijke punten verschilden de experimenten, enkele lichten we er hieronder uit. Voor uitvoerige beschrijvingen, zie: Sanders et al., 2020; De Boer et al., 2020; De Graaf & Van Gastel, 2020.

[HIER KOMT TABEL 1]

[HIER KOMT TABEL 2]

⁴ Alleen de combinatie van ontheffing en intensivering was niet toegestaan.

⁵ Vanwege de verschillende instroommomenten per gemeente varieert de totale duur van de experimenten over de verschillende gemeenten.

In Deventer en Groningen⁶ is de doelgroep eerst gerandomiseerd over de interventies en daarna uitgenodigd. De controlegroep bestaat in alle gemeenten uit bijstandsgerechtigden die wel voor het experiment zijn uitgenodigd, maar geen interventie toegewezen hebben gekregen. Bovendien bestaat er een aselekt gekozen referentiegroep, die niet voor het onderzoek werd uigenodigd en alleen administratief is gevolgd. Omdat zowel de referentiegroep als de controlegroep aselekt gekozen zijn, zouden ze bij de start van het experiment vergelijkbaar moeten zijn. De referentiegroep kan daarom gebruikt worden om te testen voor zogenaamde John-Henry-effecten,⁷ gedragseffecten in de controlegroep, die veroorzaakt worden door deelname aan een experiment zonder een expliciete interventie.

In Nijmegen, Tilburg, Utrecht en Wageningen is de doelgroep eerst uitgenodigd en vervolgens gerandomiseerd. Hier bestaat de referentiegroep uit bijstandsgerechtigden die wel werden uitgenodigd, maar niet zijn meegenomen in de randomisatie omdat ze niet reageerden op de uitnodiging of actief weigerden om mee te doen. De controlegroep bestaat uit personen die na uitnodiging en toezegging aselekt zijn toegewezen aan de controlegroep. Omdat de referentiegroep niet aselekt gekozen is, kan niet voor John Henry-effecten getest kan worden, wel kan de controlegroep vergeleken worden met de referentiegroep om representativiteit vast te stellen (Batkó, Spierings, Gesthuizen & Scheepers, 2019).

In Box 1 vindt u relevante bijzonderheden in de randomisatie en implementatie. In de bijlage zijn methodische kanttekeningen per gemeente opgenomen.

[HIER KOMT BOX 1]

⁶ In Groningen is het zogenaamde Zelen's design gehanteerd, zie: Adamson, Cockayne, Puffer & Torgerson, 2005.

⁷ Volgens de legende werkte John Henry in de jaren rond 1870 als staalbewerker bij de Amerikaanse spoorwegen. Nadat hij hoorde dat zijn productiviteit in een 'veld experiment' zou worden vergeleken met een door stoom aangedreven drillboor werkte hij extra hard. Hij wist de stoommachine te verslaan, maar moest dat door uitputting met de dood bekopen, zie: https://en.wikipedia.org/wiki/John_Henry_effect.

3. Resultaten per uitkomstmaat

3.1. Uitstroom naar regulier betaald werk

Om de uitstroom naar regulier betaald werk te meten, maken we gebruik van administratieve data over gewerkte uren en het looninkomen. We geven de voorkeur aan gewerkte uren om gedeeltelijke of volledige uitstroom naar werk te meten, omdat verschillen in uurloon de vraag of iemand gedeeltelijk of volledig is uitgestroomd compliceren.⁸ In onze analyse definiëren we gedeeltelijke of volledige uitstroom naar werk als meer dan 12 uur werk per week; en volledige uitstroom naar werk als meer dan 27 uur werk per week.⁹

Om de verschillen in uitkomsten tussen de interventie- en controlegroepen te kunnen interpreteren als het effect van (toewijzing) aan de interventiegroep, is het van belang dat groepen op elkaar lijken en dat ze een vergelijkbaar verloop van arbeidsdeelname hebben voor de start van het experiment (zie box 1). Een vergelijking van de waarneembare demografische karakteristieken wijst uit dat de interventie- en controlegroepen goed overeenkomen (De Boer et al., 2020). Bovendien laat een vergelijking van het verloop in arbeidsdeelname tussen de interventie- en controlegroepen zien, dat deze in Deventer, Groningen, Utrecht en Wageningen vergelijkbaar waren voor de start van de interventie. Maar in Nijmegen en Tilburg zien we dat de controlegroep al voor de start van het experiment vaak (statistisch) significant meer werkt dan de interventiegroepen. Hierdoor kunnen in deze gemeenten verschillen in uitkomsten niet direct geïnterpreteerd worden als het effect van de interventies.

In Deventer en Groningen kunnen we ook testen voor zogenaamde John Henry-effecten, gedragseffecten in de controlegroep (zie box 1 en voetnoot 7), door de controlegroep te vergelijken met de aselechte referentiegroep. Omdat beide groepen aselekt gekozen zijn, zouden ze op elkaar moeten lijken. Bovendien hebben ze beiden geen interventie ontvangen, waardoor het verloop van arbeidsdeelname voor en na de start ook vergelijkbaar zou moeten zijn. Een afwijkend verloop kan wijzen op een John Henry-effect, waardoor de effecten van de interventiegroep ten opzichte van de controlegroep niet volledig toegewezen kunnen worden aan de interventie.

In Deventer en Groningen zijn waarneembare karakteristieken meestal goed vergelijkbaar tussen de controle- en referentiegroep. Ook volgt de controlegroep in Deventer gedurende het experiment een vergelijkbaar traject als de referentiegroep, en lijkt er geen sprake van een John Henry-effect. In Groningen zien we echter dat het verloop van de controlegroep meteen bij de start al flink afwijkt van die van de referentiegroep.¹⁰ Dit zou een indicatie kunnen zijn van een John Henry-effect, waardoor verschillen tussen de interventiegroepen en de controlegroep niet direct

⁸ De uitkomsten, gemeten aan de hand van het inkomen geven een vergelijkbaar beeld (Sanders et al., 2020; De Boer et al., 2020)

⁹ Data over zelfstandig ondernemerschap was nog niet beschikbaar ten tijde van de uitvoering van de analyses.

¹⁰ De verschillen t.o.v. de referentiegroep zijn vaak statistisch significant (De Boer et al., 2020).

geïnterpreteerd kunnen worden als het effect van alleen de beoogde interventie.¹¹ Het is dan een combinatie van een effect van de interventie op de interventiegroep en een neveneffect van deelname aan een experiment.

Tabel 3 geeft de effecten van toewijzing aan een interventie op de gedeeltelijke en/of volledige uitstroom naar werk weer. Omdat de effecten in Nijmegen, Tilburg en Groningen om bovengenoemde redenen niet direct te interpreteren zijn als het effect van de interventies, zijn deze in grijs en cursief weergegeven. In de overige gemeenten zijn de effecten wel te interpreteren als effecten van de interventies. In Wageningen en Deventer vinden we geen (statistisch significante) effecten. De schattingen zijn in beide gemeenten met veel onzekerheid omgeven, mede doordat de aantallen in de interventiegroepen laag zijn.¹²

In Utrecht vinden we geen (significante) effecten op volledige uitstroom naar werk. Wel zien we positieve significante effecten op minimaal 12 uur werk per week in zowel de *intensiveringsgroep* en de *vrijlatingsgroep*, van respectievelijk +7,7 en +6,5%-punt (ten opzichte van 12% in de controlegroep). We vinden geen (significant) effect van het ontheffen van arbeids- en re-integratieverplichtingen. Hierbij is het belangrijk op te merken dat het verschil met de uitgangssituatie voor een deel van de deelnemers vermoedelijk niet groot is, omdat het budget voor de begeleiding van bijstandsgerechtigden in de uitgangssituatie vaak beperkt is en een deel van de mensen *de facto* ontheffen is.¹³ In het algemeen geldt dat de effecten, gegeven de beperkte aantallen deelnemers, niet heel precies te meten zijn.

Tot nu toe hebben we alleen effecten van toewijzing aan de interventies besproken, waarbij daadwerkelijke deelname buiten beschouwing is gelaten (Intention-To-Treat analyse). Voor de gemeenten waarvoor we de effecten kunnen interpreteren als de effecten van de interventies, schatten we ook de effecten voor de deelnemers (Local-Average-Treatment Effect). Voor Utrecht zijn de effecten van deelname vergelijkbaar met de effecten voor toewijzing aan de interventie.¹⁴ In Deventer en Wageningen, waar de effecten van toewijzing niet significant waren, zijn ook de effecten van deelname niet significant.

[HIER KOMT TABEL 3]

¹¹ Bij John Henry-effecten zou de controlegroep beter kunnen presteren, omdat ze zich competitief opstellen tegenover de interventiegroepen. Een andere vorm van gedragseffecten die door het experiment zelf ontstaan, is een Hawthorne effect, waarbij deelnemers zich anders gedragen omdat ze weten dat ze worden geobserveerd.

¹² Deventer heeft meerdere startmomenten, waardoor de later gestarte groepen dus minder lang gevolgd konden worden, wat leidt tot (nog) lagere aantallen voor die schattingen. In de toekomst kan onderzoek gedaan worden naar de langere termijn effecten van deze interventies.

¹³ Zie ook Koning (2017).

¹⁴ Omdat er ook in Utrecht al enkele verschillen voor de start aanwezig waren tussen deze interventiegroepen en de controlegroep voor sommige uitkomstmaten, kunnen we deze effecten niet volledig toeschrijven aan de interventies. Gevoeligheidsanalyses, waarbij gecontroleerd wordt voor deze verschillen voor de start, laten iets kleinere puntschattingen zien die nog wel statistisch significant zijn op 10% niveau (zie De Boer et al., 2020).

3.2. Zoekgedrag, subjectief welbevinden en vertrouwen

Voor het meten van de 'zachte' of subjectieve uitkomstmaten, zoals baanzoekintensiteit, zelfredzaamheid, welbevinden en vertrouwen zijn op drie momenten enquêtes onder de deelnemers gehouden. De eerste enquête is afgenomen voor aanvang of kort na start van deelname in het experiment en diende als nulmeting. Na 1 jaar volgde de tweede meting en na 2 jaar de eindmeting.¹⁵ In de figuren hieronder worden de verschillen gepresenteerd met de controlegroep op vijf van de 10 onderzochte uitkomstmaten. Daarbij is uitgegaan van de derde en laatste meting van het onderzoek.¹⁶ Door in de drie enquêtes aan te sluiten bij gebruikte meetinstrumenten en vraagstellingen in de literatuur en deze in alle gemeenten op dezelfde manier te bevragen, zijn de resultaten vergelijkbaar. Voor een verantwoording van de gebruikte vragen en uitgebreide resultaten verwijzen we naar de lokale rapportages. We bespreken hier de resultaten van vier uitkomstmaten die direct (actief zoekgedrag, zelfredzaamheid in vinden van werk) of indirect (subjectief welbevinden en sociaal vertrouwen) invloed hebben op het werkzoekgedrag en de kans op werkhervatting van de deelnemers (zie ook Sanders et al., 2020).

Werkzoekgedrag

In Figuur 1 en 2 laten we de verandering in uitkomsten zien die te maken hebben met het zoeken naar werk. Dat zijn de baanzoekintensiteit en zelfredzaamheid in het vinden van een baan. Baanzoekintensiteit is gemeten op basis van het aantal uren per week dat men in de laatste vier weken heeft gezocht.

[HIER KOMT FIGUUR 1]

Je zou verwachten dat ontheffing van sollicitatieverplichtingen leidt tot minder zoekintensiteit, terwijl een hogere beloning voor werk zou leiden tot een hogere zoekintensiteit. De gemeten effecten staan afgebeeld in Figuur 1.¹⁷

We zien dat alleen bij *Ontheffing* en in mindere mate voor *Intensivering* in Groningen de zoekintensiteit afneemt (op 10% significantieniveau) terwijl de effecten voor de andere groepen en voor alle groepen in de andere steden insignificant zijn. De grote verschillen in het effect van

¹⁵ Voor de wijze waarop de uitkomstmaten zijn uitgevraagd en berekend, verwijzen we naar een van de lokale onderzoeksrapporten. Zie: <https://www.rijksoverheid.nl/documenten/kamerstukken/2020/05/28/kamerbrief-reactie-op-uitkomsten-experimenten-participatiewet>. Geraadpleegd op 3 maart 2021.

¹⁶ De respons op de vragenlijsten verschilt per gemeente. De uitval kan selectief zijn en de resultaten beïnvloeden. Robuustheidschecks geven op dit punt echter geen reden tot zorg. In de analyses is in alle gemeenten gecontroleerd voor het antwoord op de betreffende uitkomstmaat in de nulmeting en voor achtergrondkenmerken. Op deze manier geven de resultaten het additionele effect weer op de uitkomstmaat ten opzichte van de nulmeting en vergeleken met de controlegroep.

¹⁷ De variable is erg scheef verdeeld, waarbij de meesten niet meer zoeken en een kleine minderheid actief zoekt. Deze maat vertoont wel een vertekening. De mensen die uitstromen vanwege werk, zoeken veelal niet meer en hebben bovendien een lagere respons in de latere metingen. Wat het netto effect van deze tegengestelde vertekeningen is, is niet te bepalen.

Intensivering op deze variabele (zie bijvoorbeeld Groningen (-), Nijmegen (+), Tilburg (+) en Wageningen (+)) is wellicht een aanwijzing dat de invulling van de *Intensivering* interventie tussen de gemeenten verschilde in de nadruk op het zoeken naar betaald werk. Belangrijk om te vermelden is dat de effecten op het baanzoekintensiteit zich niet één op één vertalen naar effecten op werk vind- en uitstroomkansen. De uiteindelijke conclusie is dat we (i) geen eenduidig effect kunnen vaststellen op de intensiteit waarmee men naar werk zoekt, en (ii) dat de effecten op werk daarom waarschijnlijk door andere mechanismes moeten worden verklaard (bv. kwaliteit of richting van zoekgedrag of begeleiding).

Kijken we naar de uitkomsten voor zelfredzaamheid in het vinden van werk, dan vinden we significante en positieve effecten in Utrecht voor zowel de *Ontheffingsgroep* als de *Intensiveringsgroep*. De effecten voor de *Vrijlatingsgroep* zijn in Utrecht, Groningen en Wageningen positief, maar niet significant. In Groningen vinden we ook positieve effecten voor de *Ontheffings-* en *Intensiveringsgroep*, maar opnieuw niet significant. In Deventer en Tilburg vinden we positieve effecten voor de *Intensiveringsgroep* gecombineerd met de werk bonus, waarbij het effect in Deventer niet significant is, terwijl het effect in Tilburg wel significant is. In Wageningen en Tilburg vinden we een klein negatief maar niet significant effect voor de *Intensiveringsgroep*. We concluderen dat we positief significante effecten vinden in Utrecht en in Tilburg, maar dat voor het overige het beeld niet eenduidig is, wat te maken kan hebben met de verschillen in de wijze waarop de *Intensivering* in de gemeenten is vorm gegeven (zie tabel 3).

[HIER KOMT FIGUUR 2]

Bij de scores op welbevinden in Figuur 3 valt op dat geen van de gemeenten significante positieve of negatieve effecten laten zien, met uitzondering van de groep *Ontheffing* in Groningen.¹⁸

[HIER KOMT FIGUUR 3]

Dat we voor de meeste gemeenten en voor alle groepen, uitgezonderd de *ontheffingsgroep* in Groningen, wel positieve maar insignificante effecten vinden mag niet verbazen, aangezien we uit de literatuur weten dat de verschillen tussen mensen op deze indicator redelijk groot zijn, maar

¹⁸ Er zijn in individuele gemeenten wel, ook op deze index, licht positieve en soms significante negatieve effecten gemeten in de tussenmeting. De reden dat deze effecten in de laatste maand weer verdwijnen en/of minder uitgesproken zijn, kan deels liggen aan het vooruitzicht dat het experiment en daarmee de treatments zullen eindigen of het effect dat uitkomstmaten zoals welbevinden zelfs op redelijk korte termijn een terugkeer naar het gemiddelde vertonen. Tevens kan sprake zijn van selectieve non-respons doordat mensen die gestopt zijn, bijvoorbeeld vanwege verhuizing, of werk hebben gevonden, de laatste vragenlijst niet meer invullen. Maar er kan natuurlijk ook sprake zijn van effecten die slechts een tijdelijke aard hebben.

ook vrij stabiel over de tijd en dat gemiddelden over grote groepen dicht bij elkaar liggen (Kunzmann, Little & Smith, 2000; Helliwell & Huang, 2014; Headey, Muffels & Wagner, 2010).

[HIER KOMT FIGUUR 4]

Als we kijken naar het sociale vertrouwen in Figuur 4 dan variëren de scores tussen de gemeenten van 5.1 in Tilburg tot 6.7 in Nijmegen in de controlegroep. De effecten van de verschillende interventies zijn in vier van de zes gemeenten voor alle groepen positief maar klein en alleen in Groningen significant voor de *Intensiveringsgroep* en de *vrijlatingsgroep*. Alleen in Nijmegen zien we voor beide groepen kleine negatieve maar insignificante effecten.

4. Conclusies en beleidsimplicaties

De zes experimenten in de Participatiewet geven een eerste uniek inzicht in interventies gebaseerd op handelingstheorieën die onder beleidsmakers en wetenschappers in discussie zijn: meer ruimte voor eigen regie, extra begeleiding en maatwerk, en financiële prikkels via een verruiming van de bijverdienregeling. Alle metingen hebben vlak na het sluiten van de experimenten plaatsgevonden en zullen nog worden opgevolgd door analyses gericht op de middellange termijn effecten en de effecten op zelfstandig ondernemerschap.

Uitkomsten werk

Wat uitstroom naar werk (>12 uur of meer) betreft vonden we slechts in één stad (Utrecht) significant positieve effecten voor de *intensiveringsgroep* en de *vrijlatingsgroep* (zie De Boer et al. 2020). *Ontheffing* van zoekverplichtingen lijkt bovendien de uittreding naar deeltijdwerk (>12 uur) niet te schaden, waarbij het wel de vraag is in hoeverre de interventie afwijkt van de uitgangssituatie. Op volledige uitstroom naar werk (>27 uur) vinden we in geen van de steden een statistisch significant effect van de interventies.

De impliciete veronderstelling van het huidige beleid dat strikte handhaving en toetsing van verplichtingen en sanctionering een effectievere manier is om mensen weer aan het werk te krijgen, kan worden betwist op basis van onze bevindingen (Vgl. Van Echtelt et al., 2019). De uitstroom naar werk is bij alle interventies niet lager dan de huidige aanpak, terwijl de uitstroom naar deeltijdwerk in sommige gemeenten hoger is. De hogere uitstroom komt met name naar voren wanneer meer begeleiding op maat wordt geboden of wanneer ontvangers meer mogelijkheden hebben om extra geld te verdienen door deeltijdwerk (zie De Boer et al. 2020).

Overige uitkomsten

Op de subjectieve uitkomstmaten, zoals baanzoekintensiteit, zelfredzaamheid, welbevinden, en sociaal vertrouwen, verwachtten we op grond van de theorie eenduidiger positieve resultaten (Groot et al., 2019). Immers, je zou verwachten dat bij de deelnemers meer vrijheid en

autonomie, meer vertrouwen, meer mogelijkheden om bij te verdienen of meer maatwerk in de begeleiding zou leiden tot hogere uitkomsten op deze aspecten. Maar ook hier is het beeld diffuus. In nagenoeg alle gemeenten overheerst het beeld dat er weinig significant positieve resultaten worden gevonden op de secundaire uitkomstmaten. Voor baanzoekintensiteit kunnen we geen eenduidig effect vaststellen op de intensiteit waarmee men naar werk zoekt. Het effect van intensivering verschilt sterk tussen gemeenten.

Voor welbevinden vinden we over het algemeen wel positieve maar insignificante effecten noch voor *ontheffing* noch voor *Intensivering*. De positieve effecten voor *Intensivering* zijn veelal groter dan voor *ontheffing*, maar niet statistisch significant verschillend. *Sociaal vertrouwen* zegt iets over het sociale kapitaal en de sociale integratie van mensen in de samenleving. We vonden veelal positieve resultaten, significant in Groningen voor de intensivering groep.

Het diffuse beeld is deels te herleiden in opzet van de experimenten. Vooral de *Intensivering* is in gemeenten op een verschillende manier ingevuld.

Reflectie op onderzoek

Er zijn wat ons betreft, zowel uit het proces als uit de resultaten, verschillende lessen te trekken.

1. Uit de gepresenteerde resultaten zou de conclusie kunnen worden getrokken dat het er eigenlijk niet zo veel toe doet wat het beleid doet. Dat is naar ons idee een te snelle en eenzijdige interpretatie. Met de aantallen waarmee we in de experimenten hebben gewerkt is het per definitie lastig statistisch significante effecten te vinden. Dan geven indicaties van positieve en negatieve effecten al een beeld van wat we zouden mogen verwachten wanneer de deelname aantallen groter geweest zouden zijn. Bovendien zijn de experimenten beperkt in tijd geweest. Zeker voor Intensivering geldt dat veranderingen een lange aanlooptijd nodig hebben voordat ze zich vertalen in meetbare effecten. Dit heeft alles van doen met het karakter van een veldexperiment ofwel Randomized Controlled Trial. De strikte voorwaarden die RCT's aan de experimenten opleggen schuren met de dominantie van andere factoren van budgetaire, politieke of praktische aard. Vanuit de landelijke en gemeentelijke politiek zijn beperkingen opgelegd die de bruikbaarheid van het onderzoek hebben beperkt. De politieke wens uit de landelijke politiek om de interventies zelf klein te houden (bv. marginale ruimte voor deelnemers om extra inkomsten te mogen houden vegeleken met de grotere ruimte die gemeenten en onderzoekers wilden), maar het aantal typen interventies uit te breiden (bv. de verplichting om ook een intensiveringsgroep mee te nemen), leidde tot kleinere aantallen per interventie en minder statistische power in de resultaten.

2. De tweede les uit het proces is dat, wanneer er behoefte bestaat aan onderling goed vergelijkbare data, het van groot belang is om het onderzoek van meet af aan daarop in te richten. Dat betekent een uniform design wat betreft randomisatie, typen interventies, doelgroepen en implementatie te zorgen, op een manier dat de data ook samengevoegd kunnen worden en analyses op het totaal gedaan kunnen worden.

3. Niet alle resultaten waren bruikbaar. Dat had te maken met enerzijds John Henry-effecten (de controlegroep gedraagt zich anders omdat deelnemers weten dat ze een controlegroep vormen), en anderzijds met een randomisatie die leidde tot groepen die niet vergelijkbaar waren. Wat betreft de

randomizatie is een les om bij een experiment (zeker met kleine groepen) voldoende tijd in te bouwen om te controleren of de randomizatie gelukt is of niet. Als randomizatie bij controle heeft geleid tot evident ongelijke groepen op belangrijke kenmerken, kan het proces worden herhaald.

4. Ten slotte lijkt het ons raadzaam om in vervolgonderzoek de onderzoeksvraag te verleggen van "weten wat werkt" naar "weten wat werkt voor wie". Dat betekent dat we bij een te testen interventie eerst moeten bedenken of het iets is dat we uit zouden willen rollen over het hele bijstandsbestand, of dat het gaat om een interventie die geschikt is voor een beperkt deel van het bestand. In het eerste geval is de huidige opzet een goede, maar in het tweede geval zou de opzet anders van aard moeten zijn. Uiteraard zullen ook hier politieke wensen, praktische uitvoering en wetenschappelijke integriteit met elkaar verenigd moeten worden.

Lessen voor beleid .

De lokale totstandkoming van deze experimenten illustreren de wens van lokale uitvoerders en politici voor maatwerk om te experimenteren met en te leren van de uitvoering. Maatwerk is ook van belang om na te gaan wat werkt voor wie, al was het alleen maar om te voorkomen dat te grofmazig aan de knoppen van controleren en sanctioneren wordt gedraaid. Door goed, bij voorkeur centraal gecoördineerd onderzoek op lokaal niveau te koppelen met ruimte in de wettelijke kaders om op basis van die onderzoeken het beleid op de regionale en lokale noden te kunnen aanpassen, kan meer inzicht gekregen worden in effectievere en efficiëntere uitvoering (Edzes & Van Dijk, 2015; Broersma, Edzes & Van Dijk 2013; Edzes, 2010).

De gevonden effecten op zowel werk als de subjectieve uitkomstmaten, zoals welbevinden en vertrouwen, zijn bescheiden, maar zijn soms significant positief voor uitstroom naar deeltijdwerk. Vooral *Intensivering* op maat en *Vrijlating* kan uitstroom naar (minimaal deeltijd)werk vergroten. Wat betreft de effecten op baanzoekintensiteit, zelfredzaamheid, welbevinden en vertrouwen, is het beeld diffuus. We vinden wel degelijk kleine, en soms ook positieve effecten, vooral voor zelfredzaamheid en vertrouwen, maar het beeld is niet eenduidig. Gezien de lage deelnemersaantallen is dit niet verwonderlijk en juist reden te meer om in vervolganalyses na te gaan welke interventies voor welke klanten werken.

Het Sociaal en Cultureel Planbureau concludeerde in de eindevaluatie van de Participatiewet dat het tijd is voor een (terugkeer naar een) menselijkere maat met meer aandacht en meer vertrouwen (Van Echtelt et al., 2019). Onze resultaten ondersteunen deze conclusie. Positieve aandacht is goed ontvangen in de uitvoeringspraktijk en door de deelnemers, zo blijkt uit kwalitatief onderzoek bij de experimenten (Vgl. Betkó et al., 2020; Edzes et al., 2020, Gramberg et al, 2020; Muffels et al, 2020a, 2020b, 2020c, 2020d; Verlaat et al., 2020). De meeste betrokkenen bij de experimenten — deelnemers, dienstverleners en beleidsmedewerkers — schetsen in gesprekken een positief beeld over de experimenten. Zij spreken dan over wat het doet met de deelnemers en dienstverleners en de effectiviteit in de uitvoering. Hoewel niet duidelijk zichtbaar in de cijfers, hebben betrokkenen de indruk dat mensen sneller in beweging komen, zelfredzamer worden, minder trajecten nodig hebben, de angst voor de gemeente minder is en daardoor ook weer andere perspectieven mogelijk worden. In verschillende gemeenten gaan de dienstverleners met meer plezier naar het werk, ze ervaren meer voldoening en ontdekken dat

de relatie met de deelnemer beter wordt, waardoor meer mogelijk is en de betere verstandhouding ook leidt tot meer beweging en initiatief.

Referenties

- Abbring, J. H., Van den Berg, G.J. & Van Ours, J.C. (2005). The Effect of Unemployment Insurance Sanctions on the Transition Rate from Unemployment to Employment. *The Economic Journal*, 115, 505, 602–630.
- Adamson, J., Cockayne, C., Puffer, S. & Torgerson, D.J. (2005). Review of randomised trials using the post-randomised consent (Zelen's) design. *Contemporary Clinical Trials*, 27, 4, 305-319.
- Betkó J., Spierings N., Gesthuizen M., Scheepers P. (2019). The Who and the Why? Selection Bias in an Unconditional Basic Income Inspired Social Assistance Experiment. In: Delsen L. (eds). *Empirical Research on an Unconditional Basic Income in Europe*. Contributions to Economics. Springer, Cham.
- Blonk, R.W.B. (2018). We zijn nog maar net begonnen. Oratie. Tilburg University
- Broersma, L., Edzes, A.J.E. & Van Dijk, J. (2013). Have Dutch municipalities become more efficient in managing the costs of social assistance dependency. *Journal of Regional Science*, 53, 2, 274-291.
- 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.
- De Boer, H-W, Bolhaar, J., Jongen, E., & Zulkarnain, A. (2020). *Evaluatie experimenten Participatiewet: Effecten op de uitstroom naar werk*. CPB: Den Haag.
- De Graaf, R. & Van Gastel, W. (2020). *Experimenten Participatiewet: Procesevaluatie*. ZonMw: Den Haag.
- Deci, E. L. & Ryan, R. M. (1985). *Intrinsic Motivation and Self-Determination in Human Behavior*, New York: Plenum
- Edzes, A.J.E. (2010). *Werk en bijstand: Arbeidsmarktstrategieën van gemeenten*. Maastricht: Shaker Publishing.
- Edzes, A.J.E. & Van Dijk, J. (2015). Lokale onmacht rond de Participatiewet. *Tijdschrift voor Arbeidsvraagstukken*, 31, 1, 88-95.
- Fehr, E. & 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.
- Frey, B. S. & Jegen, R. (2001). Motivation crowding theory. *Journal of Economic Surveys*, 15, 5, 589–611.
- 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*, 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.
- Helliwell, J. F., & Huang, H. (2014). New measures of the costs of unemployment: Evidence from the subjective well-being of 3.3 million Americans. *Economic Inquiry*, 52(4), 1485-1502.
- Hullegie, P, & Van Ours, J. (2014). Seek and ye shall find: How search requirements affect job finding rates of older workers, *De Economist*, Vol. 162, 377-395.
- Kampen, T., Sebrechtsm, M., Knijn, T. & Tonkens, E. (red.) (2020). *Streng maar onrechtvaardig: De bijstand gewogen*. Amsterdam: Uitgeverij van Genneep
- Kunzmann, U., Little, T. D., & Smith, J. (2000). Is age-related stability of subjective well-being a paradox? Cross-sectional and longitudinal evidence from the Berlin Aging Study. *Psychology and aging*, 15(3), 511.
- Koning, P. (2009). The effectiveness of public employment service workers in the Netherlands, *Empirical Economics*, Vol. 37, 393-409.
- Mani, A., Mullainathan, S., Shafir, E. and Zhao, J. (2013). Poverty Impedes Cognitive Function, *Science*, 341, 976–80.

- Ministerie van Sociale Zaken en Werkgelegenheid (2017). Tijdelijke regeling experimenten Participatiewet. *Staatsblad van het Koninkrijk der Nederlanden*, 69, 1-18. Beschikbaar op: <https://wetten.overheid.nl/BWBR0039258/2019-10-01>
- Mullainathan, S., & Shafir, E. (2013). *Scarcity: Why having too little means so much*. Times Books/Henry Holt and Co;
- Sanders, M. W. J. L., Betkó J. G., Edzes, A.J.E., Gramberg, P. J., Groot, L. F. M., Muffels, R. J. A., Rosenkranz, S., Rijnks, R. H., Spierings, C. H. B. M., Venhorst, V. A., & Verlaat, T. L. L. (2020). De bijstand kan beter. *Me Judice*, 2020 (28 mei 2020).
- Sen, A.K. (1999). *Development as freedom*, New York: Knopf.
- Sen, A. K. (2004). Capabilities, lists, and public reason: continuing the conversation. *Feminist Economics*, 10, 3, 77-80.
- Tiemeijer, W.L., Thomas, C.A. & Prast, H.M. (2009). *De menselijke beslisser: over de psychologie van keuze en gedrag*. Den Haag: WRR
- Wetenschappelijke Raad voor het Regeringsbeleid (2017). Weten is nog geen doen: een realistisch perspectief op zelfredzaamheid*. Den Haag
- Van der Klink, J., Bültmann, U., Burdorf, A., Schaufeli, W., Zijlstra, F., Abma, F., . . . Van der Wilt, G. (2016). Sustainable employability — definition, conceptualization, and implications: A perspective based on the capability approach. *Scandinavian Journal of Work, Environment & Health*, 42(1), 71-79.
- Van Echtelt, P., Sadiraj, K., Hoff, S., Muns, S., Karpinska, K., Das, D. & Versantvoort, M., m.m.v. L. Putman (2019). *Eindevaluatie van de Participatiewet*. Den Haag: Sociaal en Cultureel Planbureau.

Lokale rapporten

- Betkó, J., Spierings, N., Gesthuizen, M. & Scheepers, P.(2020). *Rapportage experiment Participatiewet gemeente Nijmegen*, Nijmegen.
- Edzes, A.J.E., Rijnks, R., Kloosterman, K., & Venhorst, V.A. (2020). *Bijstand op Maat: Beleidsrapport*. URSI-onderzoeksrapport 366. Groningen: Rijksuniversiteit Groningen.
- Gramberg, P. & De Swart, J. (2020). *Wat werkt op weg naar werk? Eindrapport Experiment Participatiewet gemeente Deventer*. Enschede: Saxion Hogeschool.
- Muffels, R., Blom-Stam, K. & Van Wanrooij, S. (2020a). *Vertrouwensexperiment Tilburg: Werkt het en waarom wel of niet?* Maart 2020, Tilburg University/Tranzo-ReflecT, p.1-117.
- Muffels, R., Blom-Stam, K. & Van Wanrooij, S. (2020b). *Vertrouwensexperiment Wageningen: Werkt het en waarom wel of niet?* Voorlopig eindverslag, April 2020, Tilburg University/Tranzo-ReflecT, p.1-113.
- Muffels, R., Blom-Stam, K. & Van Wanrooij, S. (2020c). *Vertrouwensexperiment Oss: Werkt het en waarom wel of niet?* February 2020, Tilburg University/Tranzo-ReflecT, p.1-110.
- Muffels, R., Blom-Stam, K. & Van Wanrooij, S. (2020d). *Vertrouwensexperiment Apeldoorn-Epe: Werkt het en waarom wel of niet?* April 2020, Tilburg University/Tranzo-ReflecT, p.1-105.
- Verlaat, T., de Kruijk, M., Rosenkranz, S., Groot, L., & Sanders, M. (2020). *Onderzoek Weten wat werkt: samen werken aan een betere bijstand*, Eindrapport. Utrecht: Universiteit Utrecht

[HIER KOMT BOX 2]

Tabel 1: Overeenkomsten en verschillen tussen de experimenten in de zes AMvB gemeenten

		Deventer	Groningen	Nijmegen	Tilburg	Utrecht	Wageningen
Aantal deelnemers		695	890 (8.724) ¹	339	780	752	410
Interventies ²	T1	A + C	A	A + C	A + C	A	A
	T2	B + C	B	B + C	B	B	B
	T3	D + C	C		B + C	C	C
	T4		E (A, B of C)				
Kenmerken <i>Ontheffingsgroep</i>		Geen arbeids- en re-integratieplicht, wel informatieplicht. Maken van plan van aanpak. 1x per 6 maanden gesprek met vaste werkmakelaar en ondersteuning op eigen initiatief.	Geen arbeids- en re-integratieplicht, wel informatieplicht. Geen verplichte bijeenkomsten. Geen contact met deelnemers, tenzij zelf geïnitieerd.	Geen arbeids- en re-integratieplicht, wel informatieplicht. Geen contact met deelnemers, tenzij zelf geïnitieerd. Alleen in combinatie met vrijlating inkomsten.	Geen arbeids- en re-integratieplicht, wel informatieplicht. Gebaseerd op 'Leren zelfredzaamheid' dus contact, zeker aan begin (2x per jr).	Geen arbeids- en re-integratieplicht, wel informatieplicht. Geen verplichte bijeenkomsten. Geen contact met deelnemers, tenzij zelf geïnitieerd.	Geen arbeids- en re-integratieplicht, wel informatieplicht. Ondersteuningsmethodiek en 'Leren zelfredzaamheid'.
Kenmerken <i>Intensiveringsgroep</i>		Maken van plan van aanpak. Iedere 14 dagen contact met vaste werkmakelaar.	Op maat dienstverlening door vijf extern aangetrokken ervaringsdeskundige begeleiders.	Voor dit experiment ontwikkelde begeleiding, met oa intensieve groepscoaching en veel autonomie deelnemers. Alleen in combinatie met vrijlating inkomsten.	Gesprekken (6x per jr in groep 2 en 5x in groep 3) met vaste coach, aanvullende instrumenten sociale domein en intensief contact met coach.	Take-in gesprek met vaste werkcoach, aanvullende instrumenten en intensiever contact, ook op initiatief van de afdeling Werk en Inkomen.	Treatment door vaste reguliere consulenten, intensief contact en ondersteuning met instrumenten PW.
Kenmerken <i>Vrijlatingsgroep</i>		25% tot €202.	50% tot €202 i.p.v. 25% tot €202.	50% tot €202 i.p.v. 25% tot €202.	50% tot €202 i.p.v. 25% tot €202.	50% tot €202 i.p.v. 25% tot €202.	50% tot €202 i.p.v. 25% tot €202.
Werkbonus					€2.392 (max. 2 keer uit te keren per half jr of jr)		

¹ Groningen heeft eerst de hele onderzoeksgroep (n=8.724) gerandomiseerd en daarna aan cliënten gevraagd of ze wilden deelnemen. Uiteindelijk hebben 890 deelnemers meegedaan.

² Toelichting: A=Ontheffing, B=Intensivering, C=Inkomstenvrijlating; D=e-dienstverlening met een APP; E=keuzevrijheid

(A, B of C). Bron: De Boer et al., 2020.

Tabel 2: Instroommomenten en duur experimenten.

Gemeente	Deventer	Groningen	Nijmegen	Tilburg	Utrecht	Wageningen
Instroom						
Experiment	1 oktober '17	1 november '17	1 december '18	1 december '17	1 juni '18	1 oktober '17
	1 februari '18		1 april '18	1 december '17 tot		2 oktober '17 tot
	1 juli '18			1 juli '18		1 september '18
Einde						
experiment	1 oktober '19	31 oktober '19	31 december '19	1 oktober '19	31 december '19	31 december '19
Duur	14-24 maanden	24 maanden	21-25 maanden	22 maanden	19 maanden	27 maanden

Bron: De Boer et al., 2020

Box 1. Toelichting op de methode van onderzoek

De experimenten zijn opgezet als een RCT dat moet voldoen aan strikte onderzoekstechnische voorwaarden. Mensen worden daarbij (voor of na aanmelding) aselekt toegewezen aan de experimentgroepen. Bij de duiding van resultaten uit een dergelijk experiment zijn twee vragen belangrijk: (1) Kunnen de geobserveerde verschillen in uitkomsten tussen experimentgroepen uitsluitend worden toegeschreven aan een andere behandeling ofwel interventie (interne validiteit); (2) Gelden de geobserveerde effecten ook voor de doelpopulatie of andere populaties (externe validiteit)?

Hiervoor zijn drie aspecten belangrijk:

1. *Aselecte toewijzing.* Door randomisatie zijn de experimentgroepen in beginsel vergelijkbaar zowel op waargenomen (bijv. opleiding, duur bijstand) als niet-waargenomen kenmerken (bijv. gezondheid, motivatie) en startpositie op de uitkomstmaten. In de praktijk kunnen er toch toevallige verschillen zijn tussen de groepen die invloed kunnen hebben op de uitkomsten. Door statistische analysemethoden kan hiervoor deels worden gecorrigeerd.
2. *Geen post-randomisatie verschillen.* De wetenschap deel te nemen aan een experiment kan gedragseffecten hebben bij de deelnemers en consumenten (experimenteffecten). Daardoor kan de dienstverlening in de interventie- en controlegroepen veranderen en kunnen effecten niet meer volledig aan de interventie worden toegeschreven. Vergelijking met een gerandomiseerde referentiegroep kan dit aan het licht brengen. Bij toewijzing ná aanmelding kan selectiviteit in de uitval ertoe leiden dat de groepen niet meer met elkaar te vergelijken zijn (randomisatie bias). Uitval gedurende het experiment kan ook selectief zijn (drop-out bias) Als voor uitvallers data beschikbaar is kan door statistische analysemethoden hiervoor deels gecorrigeerd worden.
3. *Selectiviteit van de steekproef.* Zijn de deelnemers (gerandomiseerde controlegroep en interventiegroepen) bij aanvang niet vergelijkbaar met de doelgroep, dan is er sprake van selectie bij deelnemers en/of niet-deelnemers (selectie bias). Bij toewijzing vóór aanmelding zijn de resultaten voor de experimentgroepen in hoge mate geldig voor de doelpopulatie van bijstandsgerechtigden, bij toewijzing na aanmelding is dit alleen het geval als de uitkomsten in de controlegroep vergelijkbaar zijn met een gerandomiseerde referentiegroep.

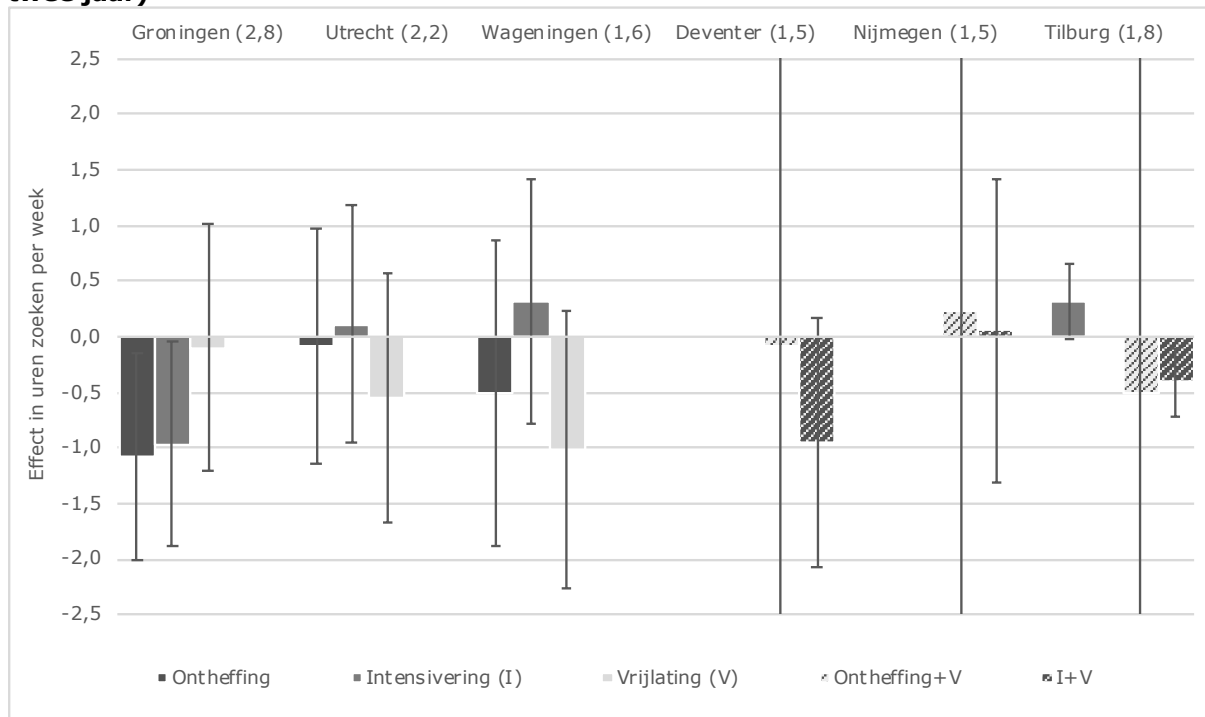
Onze experimenten zijn gebaseerd op twee soorten bronnen: administratieve data, en survey data. Wat betreft de uitkomstmaten gebaseerd op administratieve bronnen geldt dat de we data hebben van alle deelnemers, ook degenen die het experiment voortijdig verlaten. Drop-out bias wordt hiermee voorkomen. Bij de uitkomsten op basis van survey-data is het niet uit te sluiten dat uitval selectief is, en dat dit de resultaten beïnvloedt. Robuustheidschecks geven op dit punt echter geen reden tot zorg. We beschouwen onze resultaten overwegend als intern valide. Externe validiteit is altijd lastiger bij sociale experimenten (o.a. Greenberg & Schröder, 2004, Shadish, Cook & Campbell, 2002). Dit wordt versterkt doordat deelname vrijwillig is, waardoor we er niet van uit kunnen gaan dat de deelnemers aan de experimenten niet afwijken van de niet-deelnemers. Representativiteitsanalyses en een uitgebreide studie in Nijmegen (Betkó et al., 2019) laten zien dat er inderdaad verschillen zijn. Dat betekent dat we voorzichtig moeten zijn met generalisaties van de effecten naar het totale bijstandsbestand.

Tabel 4 Overzichtstabel effecten interventies, de effecten in het grijs en cursief zijn niet direct als effect van de interventie te interpreteren

	Utrecht		Wageningen		Nijmegen ^f		Tilburg ^f		Deventer		Groningen ^f	
A: Aandeel dat meer dan 27 uur werk per week werkt												
	Effect in %-punt / 100	Effect in %	Effect in %-punt / 100	Effect in %	Effect in %-punt / 100	Effect in %	Effect in %-punt / 100	Effect in %	Effect in %-punt / 100	Effect in %	Effect in %-punt / 100	Effect in %
Ontheffing ^a	+0,018	+31	+0,061	+59							-0,023⁺⁺	-29
Intensivering ^b	+0,043	+73	+0,046	+45			-0,042	-41			-0,020⁺⁺	-25
Vrijlating ^c	+0,025	+42	-0,003	-3							-0,010	-13
Keuze interventie											-0,009	-11
Ontheffing +vrijlating					-0,080⁺⁺	-48	-0,053⁺⁺	-52	+0,013	+14		
Intensivering + vrijlating					-0,041	-25	-0,055⁺⁺	-54	+0,045	+49		
Ontheffing + app + vrijlating									+0,020	+22		
Aandeel meer dan 27 uur werk per week in de controlegroep ^d	6		10		17		10		9		8	
B: Aandeel dat meer dan 12 uur werk per week werkt												
Ontheffing	+0,040	+34	-0,008	-3							-0,034⁺⁺	-24
Intensivering	+0,077^{**}	+66	-0,023	-8			-0,079⁺⁺	-32			-0,028⁺⁺	-20
Vrijlating	+0,065^{**}	+56	-0,081	-28							-0,013	-9
Keuze interventie											-0,026⁺⁺	-19
Ontheffing +vrijlating					-0,188⁺⁺	-48	-0,063	-25	-0,052	-25		
Intensivering + vrijlating					-0,129⁺⁺	-33	-0,068 [†]	-27	-0,003	-1		
Ontheffing + app + vrijlating									-0,001	0		
Aandeel meer dan 12 uur werk per week in de controle groep ^e	12		29		39		25		21		14	

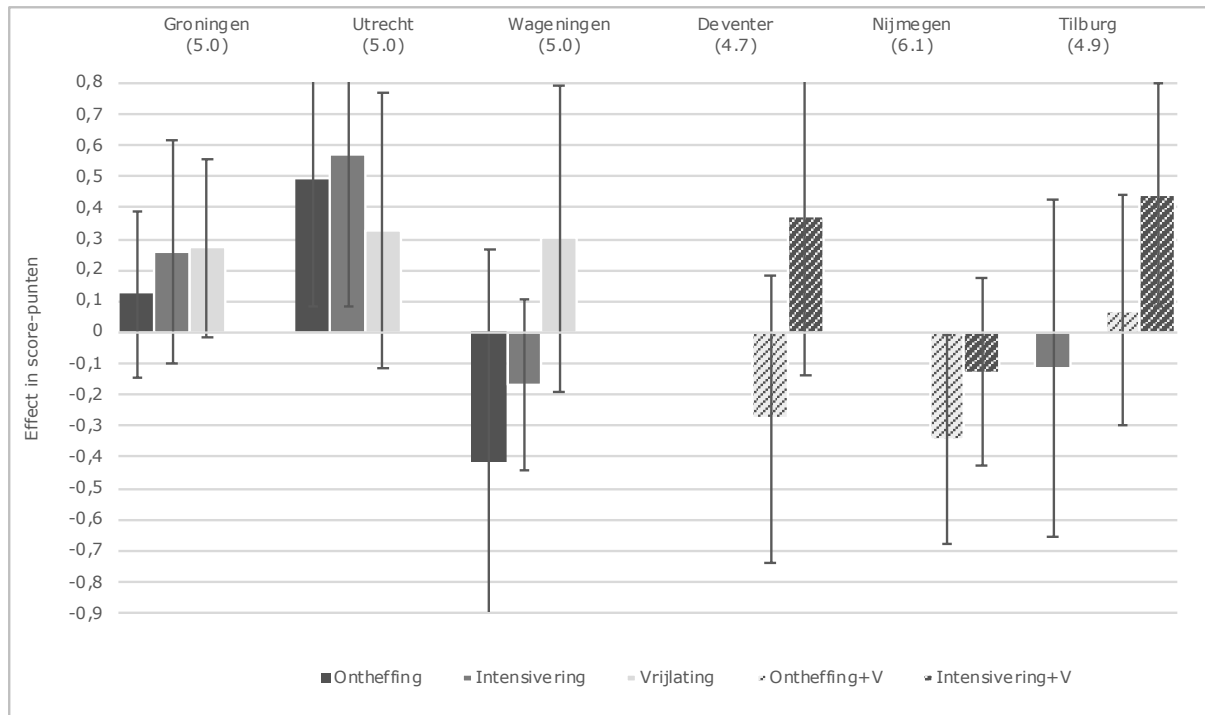
Noten: ** is significant op 5% niveau, * is significant op 10% niveau. ^a Ontheffing = ontheffing van de verplichtingen in de bijstand. ^b Intensivering = een intensievere begeleiding naar werk voor bijstandsgerechtigden. ^c Vrijlating = hogere vrijlating voor bijverdiensten in de bijstand. ^d Aandeel personen in de controlegroep dat meer dan 27 uur per week werkt aan het eind van het experiment (voor Tilburg 16 maanden na de start). ^e Aandeel personen in de controlegroep dat meer dan 12 uur per week werkt aan het eind van het experiment. ^f Voor Nijmegen, Tilburg, en Groningen is ++ significant op 5% niveau, + significant op 10% niveau. Maar deze effecten zijn niet direct te interpreteren als alleen het effect van de interventies.

Figuur 1: Baanzoek-intensiteit in uren per week in de laatste vier weken, eindmeting (na twee jaar)



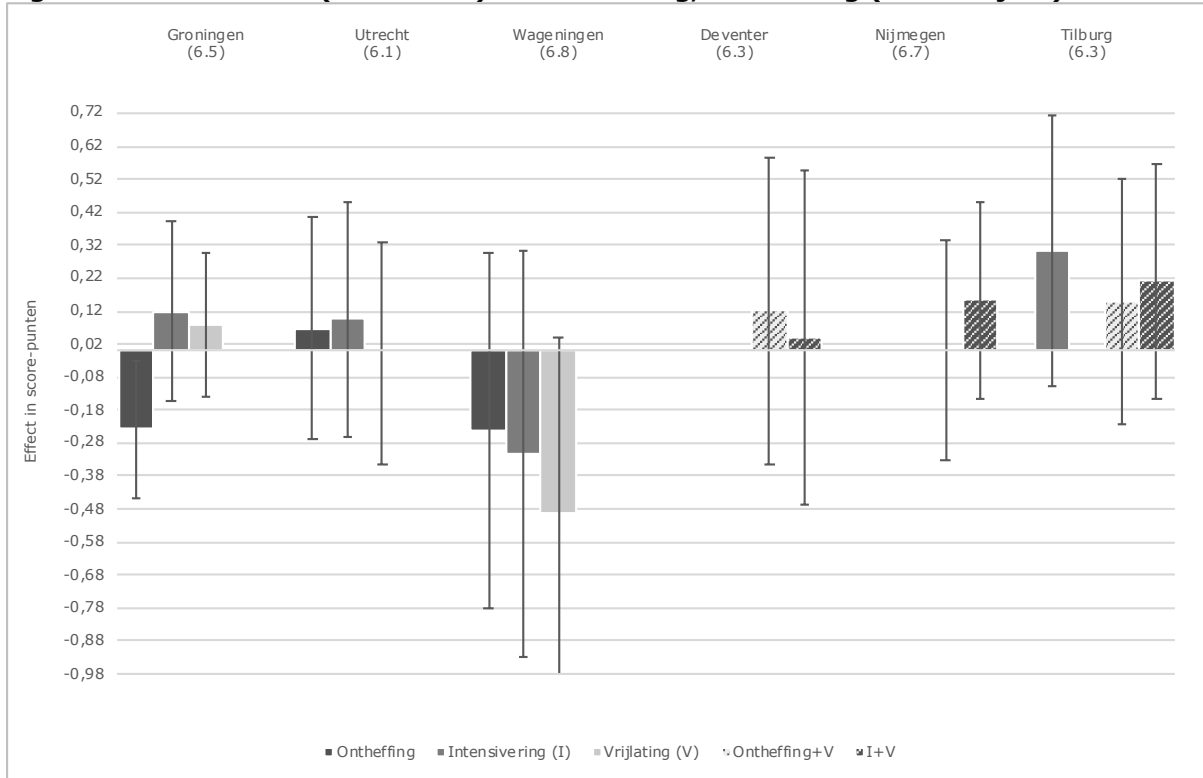
Toelichting: Treatmenteffecten t.o.v. de vergelijkingsgroep in punten op een schaal van 0-10 en 90% betrouwbaarheidsinterval. Uitkomsten voor de vergelijkingsgroep in procent staan boven de figuur. Ontheffing + V in Tilburg is Ontheffing + Vrijlating + Werkbonus. Voor Groningen is de groep waarmee vergeleken wordt deze keer de controlegroep. Zie: Edzes et al, 2020.

Figuur 2: Zelfredzaamheid in het vinden van een baan in de laatste vier weken, eindmeting (na twee jaar)



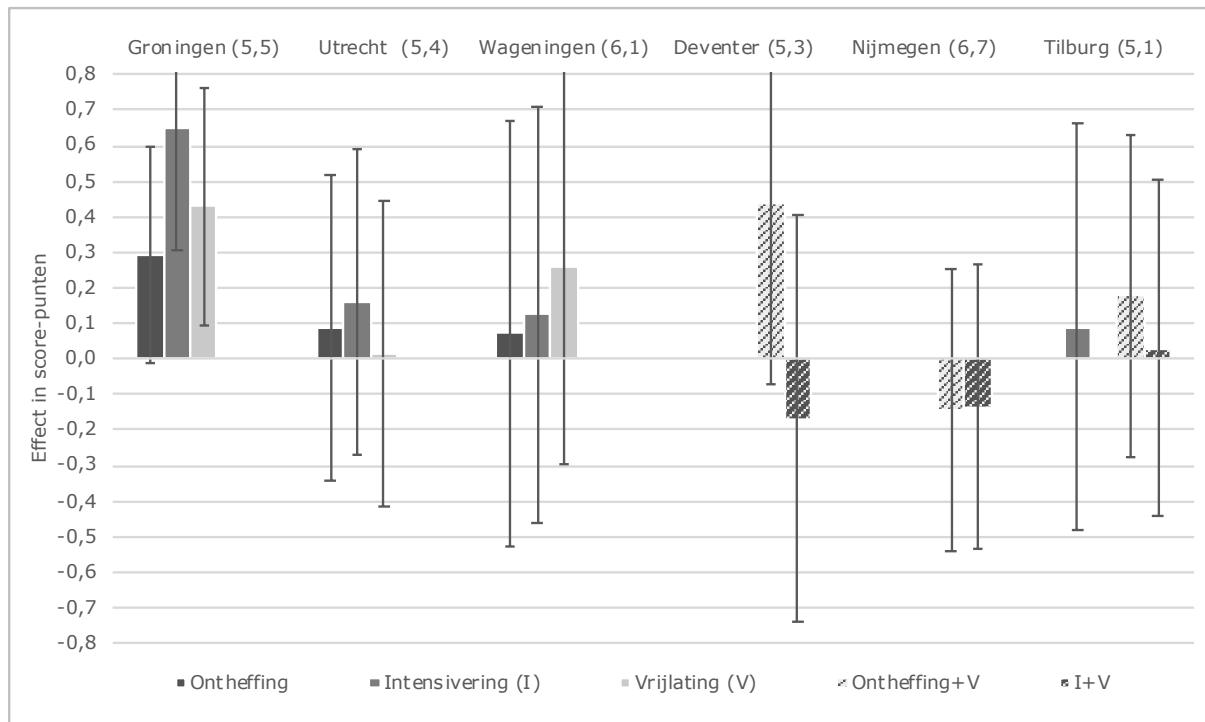
Toelichting: Treatmenteffecten t.o.v. de vergelijkingsgroep in punten op een schaal van 0-10 en 90% betrouwbaarheidsinterval. Uitkomsten voor de vergelijkingsgroep in procent staan boven de figuur. Ontheffing + V in Tilburg is Ontheffing + Vrijlating + Werkbonus. Voor Groningen is de groep waarmee vergeleken wordt deze keer de controlegroep. Zie: Edzes et al, 2020.

Figuur 3: Welbevinden (score 0-10) laatste meting, eindmeting (na twee jaar)



Toelichting: Treatmenteffecten t.o.v. de vergelijkingsgroep in punten op een schaal van 0-10 en 90% betrouwbaarheidsinterval. Uitkomsten voor de vergelijkingsgroep in procent staan boven de figuur. Ontheffing + V in Tilburg is Ontheffing + Vrijlating + Werkbonus. Voor Groningen is de groep waarmee vergeleken wordt deze keer de controlegroep. Zie: Edzes et al, 2020.

Figuur 4: Sociaal vertrouwen (score 0-10) laatste maand, eindmeting (na twee jaar)



Toelichting: Treatmenteffecten t.o.v. de vergelijkingsgroep in punten op een schaal van 0-10 en 90% betrouwbaarheidsinterval. Uitkomsten voor de vergelijkingsgroep in procent staan boven de figuur. Ontheffing + V in Tilburg is Ontheffing + Vrijlating + Werkbonus. Voor Groningen is de groep waarmee vergeleken wordt deze keer de controlegroep. Zie: Edzes et al, 2020.

Box 2: Methodische kanttekeningen per gemeente

Deventer

- *Randomisatie:* Vond plaats vóór aanmelding na opzij zetten van een willekeurig getrokken referentiegroep. Vanwege randomisatiemethode is vergelijking met de willekeurige referentiegroep zuiver. Doorgaans geen grote verschillen tussen experimentgroepen m.b.t. waarneembare achtergrondkenmerken.
- *Implementatie:* In de resultaten voor Intensivering is het effect van begeleiding door een groep vaste klantmanagers verweven.
- *Steekproef:* Niet iedereen mocht meedoen aan het experiment. Randomisatie vond plaats voor aanmelding. De deelnemers hebben een iets gunstiger arbeidsmarktpositie dan de doelgroep. Rond een kwart van de doelgroep deed mee aan het onderzoek.
- *Overig:* Door verschillende instroommomenten hebben deelnemers die vanaf het begin zijn ingestroomd vaker een vragenlijst ingevuld dan latere deelnemers.

Tilburg

- *Randomisatie:* Willekeurige groepsindeling vond plaats na aanmelding en is geslaagd.
- *Implementatie:* De controlegroep werd begeleid door consultants van de Afdeling Werk en Inkomen terwijl de treatmentgroepen werden begeleid door extern aangeworven coaches en projectleider. Ontheffing (met werkbonus) was 'het leren van zelfredzaamheid' van de deelnemers door daarvoor getrainde coaches en dus niet loslaten. De begeleiding was daardoor niet veel minder intensief dan bij de twee Intensivering groepen. De controlegroep lijkt daarentegen iets intensiever te zijn begeleid dan gangbaar. Er wordt melding gemaakt van mogelijke experiment effecten doordat controlegroep bij aanvang en direct daarna op uitstroom naar werk niet vergelijkbaar is met interventiegroepen. Effecten t.o.v controlegroep zijn daardoor niet toe te schrijven aan interventies.
- *Overig:* In het Tilburgse eindverslag is gebruik gemaakt van lokale BUS-data en zijn andere modellen gebruikt dan in dit verslag waardoor de resultaten iets kunnen afwijken. Er is middels propensity score matching gecontroleerd voor selectie in referentiegroep van alle niet-deelnemers in bijstand om uitstroom van interventiegroepen te kunnen vergelijken.

Nijmegen

- *Randomisatie:* Vond plaats ná aanmelding. Experimentgroepen zijn vergelijkbaar m.b.t. achtergrondkenmerken. In de maanden vóór start zijn er verschillen in de uitkomsten tussen de interventiegroepen en de controlegroep op het onderwerp 'deeltijdwerk', waarbij controlegroep kansrijker is. Bij aanvang zijn er (op basis van de vragenlijst) verschillen in oa (mentale) gezondheid, verslaving, en persoonlijke problemen, waarbij de controlegroep kansrijker is.
- *Implementatie:* In de resultaten voor Intensivering is het effect van begeleiding door een groepje vaste klantmanagers verweven.
- *Steekproef:* Niet iedereen mocht meedoen aan het experiment en randomisatie vond plaats na aanmelding. Als gevolg heeft de groep deelnemers een gunstigere arbeidsmarktpositie dan de doelgroep. Onderzoek wijst uit dat er binnen de potentiële deelnemers ook een verschil tussen mensen die zich opgaven en zij die dat niet deden, waarbij degenen die zich opgaven kansrijker waren.

Utrecht

- *Randomisatie:* Vond plaats ná aanmelding. Doorgaans geen grote verschillen tussen experimentgroepen m.b.t. waarneembare achtergrondkenmerken. In enkele maanden vóór start kleine verschillen in uitkomsten tussen de interventiegroepen en de controlegroep, waarbij controlegroep minder kansrijk.
- *Implementatie:* In resultaten voor Intensivering is het effect van begeleiding door een groep vaste klantmanagers verweven.
- *Steekproef:* Sommige groepen waren op voorhand uitgesloten van deelname; randomisatie vond plaats na aanmelding. Als gevolg heeft de groep deelnemers een gunstigere arbeidsmarktpositie dan de doelgroep.
- *Overig:* Uitkomsten uit de vragenlijsten voor Intensivering moeten i.v.m. een licht lagere respons voorzichtig worden geïnterpreteerd.

Groningen

- *Randomisatie:* Vond plaats vóór aanmelding na opzij zetten van een willekeurig getrokken referentiegroep. Vanwege randomisatiemethode is vergelijking met de willekeurige referentiegroep zuiver. Doorgaans geen grote verschillen tussen experimentgroepen m.b.t. waarneembare achtergrondkenmerken.
- *Implementatie:* Mogelijk gedragseffecten in de controlegroep waardoor vergelijking met deze groep niet zuiver is; waar mogelijk daarom vergelijking met willekeurige referentiegroep. Voor uitkomsten uit de vragenlijsten wel vergelijking met controlegroep omdat referentiegroep geen vragenlijsten heeft ingevuld.
- *Steekproef:* Door randomisatie van hele doelgroep hebben resultaten op basis van CBS data in hoge mate betrekking tot de doelgroep. Kortdurende bijstandsafhankelijken (< ½ jaar) horen niet tot de doelgroep.

Wageningen

- *Randomisatie:* Vond plaats ná aanmelding. Experimentgroepen zijn vergelijkbaar m.b.t. achtergrondkenmerken. In sommige maanden voor start kleine verschillen in uitkomsten tussen de intensieve begeleidingsgroep en de controlegroep, waarbij controlegroep voor start kansrijker is voor kleine banen en meer dan 50% minimumloon.
- *Implementatie:* Voor alle groepen is een werkmethode ontwikkeld waardoor de begeleiding in de controlegroep intensiever was dan gangbaar. Er wordt melding gemaakt van mogelijke experimenteffecten, doordat de interventie in de controlegroep onder andere door de nieuwe methode na de start is veranderd.
- *Steekproef:* Niet iedereen mocht meedoen aan het experiment en randomisatie vond plaats na aanmelding. Rond de helft van de bijstandspopulatie deed mee aan het onderzoek.
- *Overig:* In het Wageningse eindverslag is gebruik gemaakt van lokale BUS-data en zijn andere modellen gebruikt dan in dit verslag waardoor de resultaten iets afwijken. Er is middels propensity score matching gecontroleerd voor selectie in referentiegroep van alle niet-deelnemers in bijstand om uitstroom te kunnen vergelijken.

