



Evaluatie experimenten Participatiewet

Effecten op de uitstroom naar werk

Het ontheffen van bijstandsgerechtigden van hun verplichtingen heeft geen negatief effect op de uitstroom naar werk. Er stromen soms wel meer bijstandsgerechtigden naar deeltijdbanen als ze meer begeleid worden en als ze meer mogelijkheden krijgen om bij te verdienen. In veel gemeenten zijn de uitkomsten echter met veel onzekerheid omgeven.

Het CPB is door het ministerie van Sociale Zaken en Werkgelegenheid gevraagd de experimenten met de bijstand te evalueren. Deze experimenten hebben plaatsgevonden in Deventer, Groningen, Nijmegen, Tilburg, Utrecht en Wageningen.

CPB Notitie

Henk-Wim de Boer, Jonneke Bolhaar,
Egbert Jongen, Alice Zulkarnain

Lessen uit bijstandsexperimenten

Zes gemeenten hebben experimenten uitgevoerd om te bepalen wat werkt om meer mensen vanuit de bijstand aan het werk te krijgen. Wat hebben we daarvan geleerd?



Deelnemers

Deventer

1.584 uitgenodigd
364 deelnemers

Utrecht

752 uitgenodigd
710 deelnemers

Tilburg

780 uitgenodigd
780 deelnemers



Groningen

6.551 uitgenodigd
890 deelnemers

Wageningen

410 uitgenodigd
410 deelnemers

Nijmegen

366 uitgenodigd
304 deelnemers



Type experimenten + effecten



Ontheffing

Arbeids- en re-integratieverplichtingen vervallen

Gevonden effect

Geen. Verschil met uitgangssituatie vermoedelijk niet groot



Extra begeleiding

Deelnemer krijgt intensievere begeleiding

Gevonden effect

Soms positief. Utrecht: 20% deelnemers werkt in deeltijd, tegen 12% in de controlegroep



Hogere bijverdiensten

Deelnemer mag extra bijverdiensten met behoud van uitkering

Gevonden effect

Soms positief. Utrecht: 19% deelnemers werkt in deeltijd, tegen 12% in de controlegroep



Hoe verder?

Al zijn de uitkomsten niet eenduidig, ze bieden wel hoop. Een aantal aanbevelingen voor toekomstig onderzoek



Minder tegelijk

Minder maatregelen tegelijk maakt dat effecten preciezer meetbaar zijn



Eerst uitnodigen

Eerst uitnodigen en daarna maatregel toewijzen leidt tot meer precisie



Langere termijn

Effecten treden mogelijk pas op langere termijn op



Zelfstandigheid

Vervolgonderzoek nuttig wanneer data over uitstroom naar zelfstandig ondernemer beschikbaar zijn

Samenvatting

Het ministerie van Sociale Zaken en Werkgelegenheid heeft het CPB gevraagd de bijstandsexperimenten onder de Participatiewet te evalueren. Meer begeleiding en meer mogelijkheden om bij te verdienen in de bijstand leiden soms tot een hogere uitstroom naar deeltijdbanen. Het ontheffen van bijstandsgerechtigden van hun verplichtingen heeft geen negatief effect op de uitstroom uit de bijstand, ten opzichte van de uitgangssituatie. In veel gemeenten zijn de uitkomsten echter met veel onzekerheid omgeven. Bij toekomstige experimenten is het zaak om het aantal verschillende maatregelen te beperken, om meer precieze uitspraken te kunnen doen.

Grofweg drie typen experimenten in zes gemeenten. De experimenten bestaan uit het verlenen van een ontheffing van de arbeids- en re-integratieverplichtingen in de bijstand en/of een intensievere begeleiding naar werk en/of een hogere vrijlating voor bijverdiensten. De experimenten hebben plaatsgevonden in zes gemeenten: Deventer, Groningen, Nijmegen, Tilburg, Utrecht en Wageningen.

Ontheffing van verplichtingen in de bijstand heeft geen effect ten opzichte van de uitgangssituatie. Voor alle gemeenten vinden we geen (statistisch significant) effect van de ontheffing van de arbeids- en re-integratieverplichtingen in de bijstand op de uitstroom naar werk. Daarbij is het verschil met de uitgangssituatie voor een deel van de deelnemers vermoedelijk echter niet groot, en zijn de effecten niet heel precies te meten.

Intensievere begeleiding naar werk heeft geen effect op de volledige uitstroom naar werk, soms wel een positief effect op de gedeeltelijke of volledige uitstroom naar werk. Voor alle gemeenten vinden we geen effect van een intensievere begeleiding naar werk op de volledige uitstroom naar werk. Alleen in Utrecht is er sprake van een effect op de gedeeltelijk of volledige uitstroom naar werk (op 5% significantieniveau). 16 maanden na de start van het experiment werkt in Utrecht 20% van de intensiveringsgroep meer dan 12 uur per week, ten opzichte van 12% van de controlegroep.

Een hogere vrijlating van bijverdiensten heeft geen effect op de volledige uitstroom naar werk, soms wel een positief effect op de gedeeltelijke of volledige uitstroom naar werk. Voor alle gemeenten vinden we geen effect van een hogere vrijlating van bijverdiensten in de bijstand op de volledige uitstroom naar werk. Dit was wellicht te verwachten, omdat deelnemers niet meer dan 200 euro per maand mochten bijverdienen naast hun uitkering. Alleen in Utrecht is sprake van een positief effect op de gedeeltelijke of volledige uitstroom naar werk (op 5% significantieniveau). 16 maanden na de start van het experiment werkt in Utrecht 19% van de vrijlatingsgroep meer dan 12 uur per week, ten opzichte van 12% van de controlegroep.

De effecten zijn niet voor alle gemeenten direct te interpreteren. Voor Groningen, Nijmegen en Tilburg kunnen de uitkomsten niet direct worden geïnterpreteerd als effect van de verschillende interventies. In Groningen laat de controlegroep (uitgenodigd voor deelname aan het experiment, maar geen verandering van behandeling), andere uitkomsten zien na de start van het experiment dan de zogenaamde referentiegroep (niet uitgenodigd voor deelname aan het experiment). In Nijmegen zien we voor de start van het experiment al aanzienlijke verschillen in de uitkomsten tussen de interventiegroepen en de controlegroep. Ook in Tilburg zien we al verschillen tussen de interventiegroepen en de controlegroep voor de start van het experiment. Hierdoor zijn de waargenomen effecten in deze gemeenten niet direct toe te schrijven aan (alleen) de interventies.

Nader onderzoek gewenst naar effecten op zelfstandigheid en de langere termijn. Vanwege databeperkingen was het niet mogelijk om te bepalen of de experimenten effect hebben op de kans om als

zelfstandige aan het werk te gaan. Daarnaast konden de effecten alleen voor de eerste twee jaar worden bepaald.

Lessen voor de toekomst: minder verschillende interventies leiden tot meer precisie. Experimenten dragen bij aan *evidence-based* beleid. De inzet bij deze experimenten van projectleiders, onderzoekers, ZonMw en het ministerie van SZW is dan ook zeer te waarderen. Naar de toekomst toe kunnen de lessen uit experimenten doeltreffender zijn als het aantal verschillende interventies beperkt wordt gehouden, zodat het effect per interventie preciezer te bepalen is. Daarbij is het ook aan te raden om de interventies tussen gemeenten zoveel mogelijk gelijk te houden, wederom om de effecten van een bepaalde interventie meer nauwkeurig te kunnen bepalen. Verder is het raadzaam om deelnemers eerst uit te nodigen en daarna pas te randomiseren. Hierdoor worden effecten anders dan de interventie zelf zoveel mogelijk gelijk gehouden tussen de controlegroep en de interventiegroep.

1 Inleiding

Gemeenten hebben gedurende twee jaar bijstandsexperimenten mogen doen, om te onderzoeken of er mogelijkheden zijn de Participatiewet doeltreffender uit te voeren. Op 1 april 2017 trad het Tijdelijk besluit experimenten Participatiewet in werking. Deze wet maakt het mogelijk dat gemeenten bevoegd zijn om bij wijze van experiment, gedurende twee jaar op onderdelen van de Participatiewet af te wijken met het oog op het onderzoeken van mogelijkheden om de Participatiewet op onderdelen doeltreffender uit te voeren. Het ministerie van Sociale Zaken en Werkgelegenheid (hierna: SZW) heeft de volgende gemeenten aangewezen als experimenteergemeenten: Deventer, Groningen, Nijmegen, Tilburg, Utrecht en Wageningen.¹ Elke gemeente levert ook een eigen rapportage van het in de betreffende gemeente uitgevoerde experiment aan het ministerie van SZW.

SZW heeft het Centraal Planbureau gevraagd om een onafhankelijke kwantitatieve analyse van de effecten van de experimenten op de gedeeltelijke en volledige uitstroom naar werk. In deze kwantitatieve analyse evalueert het Centraal Planbureau (hierna: CPB) wat de effecten van de experimenten zijn geweest op uitstroom uit de uitkering in het kader van de Participatiewet naar regulier, betaald werk. Het betreft daarbij zowel de gedeeltelijke als volledige uitstroom naar regulier, betaald werk. Tevens is het CPB gevraagd om te bepalen in hoeverre de effecten direct als het effect van de beoogde interventies kunnen worden geïnterpreteerd.

De opbouw van de notitie is als volgt. In hoofdstuk 2 geven we een overzicht van de experimenten in de verschillende gemeenten. Hoofdstuk 3 behandelt de empirische methode. Hoofdstuk 4 geeft een overzicht van de gebruikte datasets, en beschrijvende statistieken voor de persoons- en huishoudenkenmerken en de uitkomsten voor en na de interventie. Dit geeft mede² inzicht in de mate waarin de effecten van de experimenten direct kunnen worden geïnterpreteerd als gevolg van de beoogde interventie. Hoofdstuk 5 behandelt vervolgens de effecten van de experimenten op de uitstroom naar werk per gemeente, en geeft het beeld over de gemeenten heen. Hoofdstuk 6, ten slotte, behandelt enkele kanttekeningen bij de analyse en mogelijkheden voor toekomstig onderzoek.

2 Beschrijving experimenten Participatiewet

Er zijn grofweg drie typen experimenten gedaan in zes gemeenten. De experimenten bestaan uit het verlenen van een ontheffing van de arbeids- en re-integratieverplichtingen in de bijstand en/of een intensievere begeleiding naar werk en/of een hogere vrijlating voor bijverdiensten. De experimenten hebben plaatsgevonden in zes gemeenten: Deventer, Groningen, Nijmegen, Tilburg, Utrecht en Wageningen. In

¹ We bedanken het ministerie van SZW, de projectleiders in de zes gemeenten, de gemeentelijke onderzoekers, de wetenschappelijke begeleidingscommissie en ZonMw voor het beschikbaar stellen van de gemeentelijke gegevens over de experimenten en voor hun opmerkingen en suggesties.

² Om de effecten van de experimenten direct te kunnen interpreteren als gevolg van de beoogde interventie is het noodzakelijk dat de interventie- en controlegroepen voor de interventie voldoende vergelijkbaar zijn op zowel kenmerken als op de uitgangssituatie voor de interventie. Dit is echter geen voldoende voorwaarde om de effecten van de interventie te kunnen interpreteren als direct gevolg van (alleen) de interventie, het is daarnaast ook noodzakelijk dat de controlegroep dezelfde behandeling krijgt na de interventie als voor de interventie en zijn of haar gedrag niet aanpast.

Deventer en Groningen is er eerst gerandomiseerd en daarna uitgenodigd, in de andere gemeenten is er eerst uitgenodigd en daarna gerandomiseerd. Na de start van het experiment valt een deel van de deelnemers uit, maar we kunnen iedereen administratief blijven volgen in de CBS data.

Op grond van het Tijdelijk besluit experimenten Participatiewet heeft de minister van SZW de volgende gemeenten aangewezen als experimenteergemeenten: Deventer, Groningen, Nijmegen, Tilburg, Utrecht en Wageningen. Door de aanwijzing zijn deze gemeenten bevoegd om, bij wijze van experiment, gedurende maximaal twee jaar, op onderdelen van de Participatiewet af te wijken, met het oog op het onderzoeken van mogelijkheden om de Participatiewet op onderdelen doeltreffender uit te voeren.

De instroommomenten verschilden tussen de gemeenten en sommige gemeenten gebruikten meerdere instroommomenten. Tabel 2.1 toont de verschillende instroommomenten, einddatum en totale duur van de experimenten per gemeente. Zo hadden Deventer en Nijmegen respectievelijk drie en twee instroommomenten. Vanwege de verschillende instroommomenten per gemeente varieert de totale duur van de experimenten over de verschillende gemeenten.

Tabel 2.1 Overzichtstabel instroommomenten en duur experiment, per gemeente

	Deventer	Groningen	Nijmegen	Tilburg	Utrecht	Wageningen
Instroom Experiment	1 oktober '17 1 februari '18 1 juli '18	1 november '17	1 december '17 1 april '18	1 oktober '17	1 juni '18	1 oktober '17
Einde Experiment	1 oktober '19	31 oktober '19	1 januari '20	1 oktober '19	1 oktober '19	31 december '19 ³
Totale duur	14-24 maanden	24 maanden	21-25 maanden	Maximaal 24 maanden ⁴	16 maanden	Maximaal 27 maanden ⁵

De interventies bestonden uit een ontheffing van de arbeids- en re-integratieverplichtingen, een intensievere begeleiding en een hogere vrijlating voor bijverdiensten, dan wel een combinatie daarvan. Meer specifiek konden de gemeenten de volgende interventies inzetten bij de experimenten:

- Een tijdelijke ontheffing van de arbeids- en re-integratieverplichtingen (ontheffingsgroep).
- Een tijdelijke intensivering van de begeleiding van bijstandsgerechtigden (intensiveringsgroep).
- Een vrijlating van de inkomsten uit arbeid tot maximaal 50 procent van deze inkomsten, met een maximum van circa 200 euro per maand (vrijlatingsgroep).⁶
- Gemeenten konden er ook voor kiezen om bovenstaande interventies te combineren. Alleen de combinatie van ontheffing en intensivering was niet toegestaan.

De precieze vormgeving van de interventies verschilde tussen gemeenten en wordt hieronder in meer detail besproken (paragraaf 2.2).

³ SZW heeft toestemming gegeven aan de verschillende gemeenten om de experimenten te verlengen tot 31 december 2019.

⁴ Tilburg kende verschillende instroommomenten tot aan 1 juli 2018, waarbij elke twee weken is gerandomiseerd op de nieuwe aanmeldingen.

⁵ Wageningen kende verschillende instroommomenten tot aan 1 juli 2018, waarbij elke twee weken is gerandomiseerd op de nieuwe aanmeldingen.

⁶ De Participatiewet kent een vrijlating van maximaal 25 procent van de inkomsten uit arbeid, met een maximum van circa 200 euro per maand. Het exacte bedrag kan per gemeente verschillen.

Via randomisatie zijn personen toegewezen aan verschillende groepen. Via randomisatie zijn personen toegewezen aan een zogenaamde referentiegroep (geen uitnodiging voor het experiment), een controlegroep (een uitnodiging voor het experiment, maar in principe geen speciale behandeling) en verschillende behandelgroepen (een uitnodiging voor het experiment, en een speciale behandeling). Deelname aan het experiment is vrijwillig, een deel van de personen kiest er voor om niet (meer) aan het experiment deel te nemen. Met behulp van de administratieve data kunnen we echter zowel de personen die deelnemen aan het experiment als personen die niet (meer) deelnemen volgen over de tijd. De randomisatie verschilt tussen gemeenten, dit bespreken we in paragraaf 2.2.

2.1 Genodigden, deelnemers en uitval

In alle gemeenten zijn de interventiegroepen gerandomiseerd toegewezen, daarbij verschilt het moment van randomisatie wel tussen gemeenten. De randomisatie is uitgevoerd onder begeleiding van een begeleidingscommissie en ZonMw. Zoals reeds opgemerkt verschilt het moment van randomiseren per gemeente. In Deventer en Groningen is ervoor gekozen om de doelpopulatie van de bijstandsgerechtigden te randomiseren en vervolgens uit te nodigen voor de controlegroep of één van de verschillende interventiegroepen.⁷ In Nijmegen, Utrecht, Tilburg en Wageningen is er eerst een uitnodiging voor deelname aan het experiment gestuurd aan potentiële deelnemers. Vervolgens zijn de personen die wilden deelnemen *random* toegewezen aan de controlegroep of één van de verschillende interventiegroepen.

Een deel van de personen stopt met het experiment, maar we kunnen alle personen volgen in de administratieve data. In de databestanden is bekend welke personen zijn uitgenodigd voor het experiment, welke personen zich hebben aangemeld voor het experiment, welke personen daadwerkelijk zijn gestart met het experiment en welke personen tussentijds zijn gestopt. Tabel B1.1 in de bijlage toont een overzicht van de aantallen personen in de verschillende groepen die zijn uitgenodigd, aangemeld, gestart en eventueel gestopt. Voor alle gemeenten geldt dat de meeste aangemelde personen ook daadwerkelijk gestart zijn. Wel zien we in de meeste gemeenten dat ook een aantal van de deelnemers voor het einde stopt met het experiment (van 2% tot 39%, gemiddeld 16%). In Deventer, Groningen en Utrecht is er ook een *random referentiegroep* samengesteld die niet is uitgenodigd voor het experiment. We kunnen deze personen wel volgen in de administratieve data. In Nijmegen, Tilburg en Wageningen zijn er referentiegroepen samengesteld uit de personen die uitgenodigd waren voor een interventie, maar die zich niet hebben aangemeld. Deze laten wij in deze notitie buiten beschouwing omdat ze wegens selectie minder goed vergelijkbaar zijn met de interventiegroepen.

2.2 Interventies per gemeente

Tabel 2.2 geeft een overzicht van de interventies per gemeente. Hieronder lichten we de interventies per gemeente toe.

⁷ De doelpopulatie voor de experimenten bestaat uit alle bijstandsgerechtigden minus de bijstandsgerechtigden die zijn uitgesloten door exclusiecriteria. Zo nemen de meeste gemeenten bijstandsgerechtigden jonger dan 27 jaar niet mee. Hetzelfde geldt voor bijstandsgerechtigden die ook arbeidsongeschikt zijn of in de schuldsanering zitten. De exclusiecriteria verschillen per gemeenten, zie ZonMw (2020) voor een volledig overzicht.

Tabel 2.2 Overzichtstabel verschillende interventies, per gemeente

Groep	Deventer	Groningen	Nijmegen	Tilburg	Utrecht	Wageningen
1	A ₁ + C	A	A + C	A + C	A	A
2	A ₂ + C	B	B + C	B	B	B
3	B + C	C		B + C	C	C
4		A, B of C				

Noot: A = Ontheffing, B = Intensivering, C = Vrijlating inkomensgrens.

Deventer. In Deventer is eerst gerandomiseerd en daarna uitgenodigd. De controlegroep bestaat uit bijstandsgerechtigden die wel voor het experiment zijn uitgenodigd, maar in principe geen interventie hebben gehad.⁸ De referentiegroep is ook een aselect gekozen groep, maar is alleen administratief gevolgd en niet voor het onderzoek uitgenodigd. Omdat zowel de referentiegroep als de controlegroep aselect gekozen zijn, zouden ze bij de start van het experiment vergelijkbaar moeten zijn.

Er zijn drie interventiegroepen in Deventer:

- Voor interventiegroep 1 geldt een ontheffing (A₁). Deze groep krijgt geen begeleiding en er is alleen een gesprek met een klantmanager (1x per 6 maanden). De ontheffing wordt gecombineerd met een vrijlating om bij te verdienen (C). De vrijlating van de inkomsten uit arbeid wordt uitgebreid tot 25% van het verdiende inkomen (in plaats van 25% voor het experiment), maximaal circa 200 euro per maand⁹ gedurende de gehele looptijd van het experiment (in plaats van de maximaal 6 maanden voor het experiment).
- Voor interventiegroep 2 geldt ook een ontheffing (A₂). Deze groep krijgt daarnaast digitale begeleiding en een gesprek met een klantmanager (1x per 6 maanden). De ontheffing wordt gecombineerd met een vrijlating om bij te verdienen (C). De vrijlating van de inkomsten uit arbeid wordt uitgebreid tot 25% van het verdiende inkomen, maximaal circa 200 euro per maand, gedurende de gehele looptijd van het experiment (in plaats van de maximaal 6 maanden voor het experiment)..
- Interventiegroep 3 is de intensiveringsgroep, deze groep krijgt intensieve begeleiding en actieve ondersteuning. Het gaat dan om een tweewekelijks gesprek met een klantmanager en ondersteuning bij het maken van een plan van aanpak bij het vinden van werk. De intensivering wordt gecombineerd met een vrijlating om bij te verdienen (C). De vrijlating van de inkomsten uit arbeid wordt uitgebreid tot 25% van het verdiende inkomen, maximaal circa 200 euro per maand, gedurende de gehele looptijd van het experiment (in plaats van de maximaal 6 maanden voor het experiment)..

Groningen. In Groningen is ook eerst gerandomiseerd en daarna uitgenodigd. De controlegroep bestaat uit bijstandsgerechtigden die wel voor het experiment zijn uitgenodigd, maar in principe geen interventie hebben gehad. De referentiegroep is ook random gekozen, maar is alleen administratief gevolgd en niet voor het onderzoek uitgenodigd. De referentiegroep en de controlegroep zouden bij de start van het experiment vergelijkbaar moeten zijn, omdat beide via randomisatie bepaald zijn.

⁸ Voor de controlegroep en de referentiegroep in alle gemeenten geldt een vrijstelling van 25% van de inkomsten, tot maximaal 200 euro per maand, gedurende maximaal 6 maanden. Voor het experiment kregen bijstandsgerechtigden geen uitstroompremie bij het vinden van werk.

⁹ De exacte bedragen verschillen per gemeenten, maar bedraagt in alle gevallen circa 200 euro per maand.

Er zijn vier interventiegroepen in Groningen:

- Interventiegroep 1 is een ontheffingsgroep (A), waarbij meedoen aan het bestaande begeleidingsaanbod geheel vrijwillig is.
- Interventiegroep 2 is een intensiveringsgroep (B). Deze groep wordt begeleid door ingehuurde ervaringsdeskundigen.
- Interventiegroep 3 is een vrijlatingsgroep (C), waarbij de vrijlating om bij te verdienen wordt uitgebreid tot 50% van het verdiende inkomen, maximaal circa 200 euro per maand, gedurende de gehele looptijd van het experiment (in plaats van de maximaal 6 maanden voor het experiment)..
- Interventiegroep 4 is een groep waarbij deelnemers een vrije keuze hebben tussen ontheffing (A), intensivering (B) of vrijlating (C).

Nijmegen. In Nijmegen is eerst uitgenodigd en vervolgens gerandomiseerd. De referentiegroep is anders tot stand gekomen dan in Deventer en Groningen. 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 aselect zijn toegewezen aan de controlegroep. Omdat de referentiegroep niet aselect gekozen is en de controlegroep bestaat uit een aselecte toewijzing onder personen die zich hebben aangemeld voor het experiment zijn deze twee groepen niet vergelijkbaar.

Er zijn twee interventiegroepen in Nijmegen:

- Interventiegroep 1 is een combinatie van een ontheffing (A) en een uitbreiding van de vrijlating (C). Er is geen sollicitatie- en re-integratieverplichting, en de vrijlating om bij te verdienen is uitgebreid tot 50% van het verdiende inkomen, maximaal circa 200 euro per maand, gedurende de gehele looptijd van het experiment (in plaats van de maximaal 6 maanden voor het experiment).
- Interventiegroep 2 is een combinatie van intensivering (B) en vrijlating (C). De intensieve begeleiding bevat een nieuwe re-integratiemethode, inclusief verplichte bijeenkomsten, en de vrijlating om bij te verdienen wordt uitgebreid tot 50%, maximaal circa 200 euro per maand, gedurende de gehele looptijd van het experiment (in plaats van de maximaal 6 maanden voor het experiment).

Tilburg. In Tilburg is eerst uitgenodigd en daarna gerandomiseerd. Net als in Nijmegen bestaat de referentiegroep in Tilburg 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 mensen die zich na uitnodiging hebben aangemeld, en vervolgens gerandomiseerd zijn in de controlegroep. Wederom zijn deze twee groepen niet vergelijkbaar.

Er zijn drie interventiegroepen in Tilburg:

- Interventiegroep 1 betreft een combinatie van een ontheffing (A) en een vrijlating (C). Deelname aan het bestaande begeleidingsaanbod is vrijwillig. De vrijlating van de inkomsten uit arbeid wordt uitgebreid tot 50% van het verdiende inkomen, maximaal circa 200 euro per maand, gedurende de gehele looptijd van het experiment (in plaats van de maximaal 6 maanden voor het experiment). Daarnaast is er een uitstroompremie van 2400 euro per jaar bij uitstroom naar werk. De uitstroompremie wordt alleen toegekend bij volledige uitstroom uit de bijstand.
- Interventiegroep 2 is een intensiveringgroep (B), met 6 afspraken per jaar en op maat gemaakte supervisie.
- Interventiegroep 3 bevat een combinatie van intensivering (B) en vrijlating (C). De intensivering is gelijk aan de intensivering bij interventiegroep 2, 5 afspraken per jaar en op maat gemaakte supervisie. De vrijlating van de inkomsten uit arbeid wordt uitgebreid tot 50% van het verdiende inkomen, maximaal circa 200 euro per maand, gedurende de gehele looptijd van het experiment (in plaats van de maximaal 6 maanden voor het experiment).

maanden voor het experiment). Daarnaast is er een uitstroompremie van 2400 euro per jaar bij uitstroom naar werk.

Utrecht. In Utrecht zijn er twee referentiegroepen. De eerste referentiegroep is een voor de start van het experiment aselekt gekozen groep bijstandsgerechtigden die alleen administratief werd gevolgd, en niet voor het onderzoek was uitgenodigd (cf. de referentiegroep in Deventer en Groningen). De tweede referentiegroep bestaat 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 (cf. de referentiegroep in Nijmegen en Tilburg). De aselekt gekozen referentiegroep is minder vergelijkbaar met de controlegroep, omdat die laatste bestaat uit mensen die zich na uitnodiging hebben aangemeld voor het experiment. De tweede referentiegroep is ook niet vergelijkbaar met de controlegroep, omdat de laatste bestaat uit een (aselecte) groep personen die zich heeft aangemeld.

Er zijn drie interventiegroepen in Utrecht:

- Interventiegroep 1 is een ontheffingsgroep (A) waarvoor geen sollicitatie- en re-integratieverplichting geldt.
- Interventiegroep 2 is een intensiveringsgroep (B), met intensieve begeleiding door een speciale groep klantmanagers.
- Interventiegroep 3 betreft de vrijlating (C) om bij te verdienen die wordt uitgebreid tot 50% van het verdiende inkomen, maximaal circa 200 euro per maand, gedurende de gehele looptijd van het experiment (in plaats van de maximaal 6 maanden voor het experiment).

Wageningen. In Wageningen is eerst uitgenodigd en daarna gerandomiseerd. Net als in Nijmegen en Tilburg bestaat de referentiegroep in Wageningen 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 mensen die wel zijn uitgenodigd voor het experiment, maar in principe geen andere behandeling hebben ontvangen. De referentiegroep en controlegroep zijn daarom niet vergelijkbaar.

Er zijn drie interventiegroepen in Wageningen:

- Interventiegroep 1 is een ontheffingsgroep (A) waarvoor geen sollicitatie- en re-integratieverplichting geldt. Wel is er een halfjaarlijks voortgangsgesprek.
- Interventiegroep 2 is een intensiveringsgroep (B), met meer contactmomenten (naast het halfjaarlijkse voortgangsgesprek) en is er supervisie op maat.
- Interventiegroep 3 betreft de vrijlating (C) om bij te verdienen die wordt uitgebreid tot 50% van het verdiende inkomen, maximaal circa 200 euro per maand, gedurende de gehele looptijd van het experiment (in plaats van de maximaal 6 maanden voor het experiment).

3 Methode

Gedeeltelijk of volledige uitstroom naar werk is gedefinieerd als meer dan 12 uur werk per week; volledige uitstroom naar werk is gedefinieerd als meer dan 27 uur werk per week. Met behulp van een *intention-to-treat* (ITT) schatting bepalen we het effect van toewijzing aan een bepaalde interventie. Met behulp van een *local average treatment effect* (LATE) schatting bepalen we het effect van deelname aan een bepaalde interventie. Door randomisatie zijn de controle- en interventiegroepen (op basis van toewijzing) in principe vergelijkbaar op zowel waargenomen als niet-waargenomen kenmerken, maar dan nog zijn er valkuilen die een zuivere

effectmeting van de interventie in de weg kunnen staan. Om te bepalen in hoeverre deze valkuilen van toepassing zijn doen we een aantal tests.

Dit hoofdstuk bespreekt de gehanteerde empirische methodologie om het effect te bepalen op de gedeeltelijke of volledige uitstroom naar werk. We bespreken eerst de gehanteerde definitie voor gedeeltelijke en volledige uitstroom naar werk. Vervolgens bespreken we de voordelen van randomisatie, met enkele kanttekeningen. Tot slot bespreken we de gehanteerde empirische methodologie.

3.1 Definitie gedeeltelijke en volledige uitstroom

Gedeeltelijke of volledige uitstroom naar werk is gedefinieerd als meer dan 12 uur werk per week; volledige uitstroom naar werk is gedefinieerd als meer dan 27 uur werk per week. Om uitstroom naar regulier betaald werk te meten zouden we idealiter gebruikmaken van administratieve data over de bijstandsuitkeringen. Deze data zijn echter nog niet beschikbaar voor de volledige periode van de uitgevoerde experimenten. Administratieve data over lonen en banen zijn wel beschikbaar tot en met het derde kwartaal van 2019 en deze data bevatten informatie over gewerkte uren en het looninkomen. We geven de voorkeur aan de 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. De uitkomsten, gemeten aan de hand van het inkomen geven overigens een vergelijkbaar beeld, deze presenteren we als gevoeligheidsanalyse. We definiëren volledige uitstroom naar werk als 27 uur werk per week¹⁰, en 12 uur per week of meer als gedeeltelijke of volledige uitstroom naar werk. In een gevoeligheidsanalyse kijken we ook naar volledige uitstroom gedefinieerd als meer dan 32 uur werk per week en gedeeltelijke of volledige uitstroom gedefinieerd als meer dan 8 uur werk per week.

Vooralsnog kan niet onderzocht worden of mensen uitstromen naar zelfstandigheid (zelfstandig ondernemer). De administratieve data over het inkomen als zelfstandige is op het moment van schrijven alleen nog beschikbaar tot en met eind 2018.¹¹

3.2 Gerandomiseerde toewijzing

Door randomisatie zijn de controle- en interventiegroepen in principe vergelijkbaar op zowel waargenomen en niet-waargenomen kenmerken, maar dan nog zijn er valkuilen die een zuivere effectmeting van de interventie in de weg kunnen staan. Buiten de setting van een experiment is het doorgaans niet voldoende om de uitkomsten van een interventiegroep te vergelijken met die van een willekeurige andere groep die deze interventie niet heeft gehad. Vaak zijn die groepen niet vergelijkbaar, omdat personen die ervoor kiezen om een interventie te volgen vaak andere waargenomen en niet-waargenomen kenmerken hebben dan personen die daar niet voor kiezen. Om de interventie- en controlegroepen vergelijkbaar te maken, wordt daarom vaak gebruikgemaakt van randomisatie. Randomisatie leidt tot aselechte toewijzing aan interventie- en controlegroepen, dit zorgt er, in het algemeen, voor dat deze groepen vergelijkbare waargenomen en niet-waargenomen kenmerken hebben. Het hebben van vergelijkbare waargenomen en niet-waargenomen kenmerken van interventie- en controlegroepen is een noodzakelijke maar nog niet voldoende voorwaarde om de verschillen in de uitkomsten tussen de groepen als het directe effect van (alleen) de interventie te interpreteren:

¹⁰ 27 uur werk op minimumloonniveau levert een inkomen op vergelijkbaar met het bijstandsniveau van een alleenstaande.

¹¹ Het CBS verwacht dat data over zelfstandigen, voor het jaar 2019, in januari 2021 beschikbaar komen. Het CPB raadt aan om een vervolganalyse uit te voeren nadat deze data beschikbaar komen.

- Ten eerste, het kan voorkomen dat de randomisatie door toeval niet tot vergelijkbare groepen heeft geleid. Door een goede implementatie van de experimenten en bij voldoende aantallen wordt dit risico beperkt.
- Ten tweede, personen hebben de keuze om niet deel te nemen aan de interventie (of meer in het algemeen om de voorgestelde interventie niet op te volgen), dit zijn de zogenaamde *non-compliers* (non-conformisten). Tevens kunnen deelnemers ervoor kiezen om voor het einde van de interventie te stoppen.¹² Een lage *compliance rate* (nalevingsgraad) maakt het moeilijker een statistisch significant effect vast te stellen, omdat het effect 'verwatert'.¹³ In Deventer en Groningen is de randomisatie anders uitgevoerd dan in de overige gemeenten. Beide gemeenten hebben ervoor gekozen om het hele bestand te randomiseren en vervolgens uit te nodigen. Een voordeel hiervan is dat de interventiegroepen bij toewijzing goed vergelijkbaar zijn met de gehele doelgroep. Een nadeel is dat het kan leiden tot lage *compliance rates*.
- Ten derde, er kunnen alleen al door de participatie aan een experiment gedragseffecten ontstaan, waardoor verschillen tussen de controle- en interventiegroepen niet volledig aan het effect van de interventie kunnen worden toegeschreven (Duflo et al., 2007).¹⁴ Zo bestaan er mogelijk:
 - Hawthorne-effecten¹⁵ – gedragseffecten in de interventiegroep: participanten die zijn toegewezen aan de interventiegroep kunnen zich anders gaan gedragen, bijvoorbeeld uit dankbaarheid voor de interventie of doordat ze weten dat ze in de gaten worden gehouden.¹⁶ Deze gedragseffecten, die veroorzaakt worden door de experimentele setting los van de interventie zelf, kunnen een effect hebben op de uitkomst.
 - John Henry effecten¹⁷ – gedragseffecten in de controlegroep: de toewijzing aan de controlegroep kan gedragseffecten veroorzaken. Zo kunnen zij gedemotiveerd raken doordat ze niet aan een interventie mee mogen doen, of zich juist competitief opstellen. Dit zijn wederom gedragseffecten die de uitkomst kunnen beïnvloeden.

Om te bepalen of we de effecten per gemeente direct als gevolg van de interventie kunnen interpreteren doen we een aantal tests. Om de eerste mogelijke complicatie te onderzoeken presenteren we in hoofdstuk 4 zogenaamde *balancing tests* van de verschillende groepen. Zo bepalen we in hoeverre de verschillende groepen statistisch significante verschillen laten zien in de waargenomen kenmerken. Tevens presenteren we een grafische vergelijking van de uitkomsten voor de start van het experiment voor de interventie- en controlegroepen, om te zien of er voor de start niet al verschillen waren in gewerkte uren. Bovendien voeren we ook placebo-regressies uit om te zien of er, na correctie voor waarneembare kenmerken, al statistisch significante verschillen zijn voor (of vlak na) de start van de interventies. De tweede complicatie speelt een rol bij de effecten voor enkele gemeenten, waarbij eerst is gerandomiseerd en daarna is uitgenodigd in plaats van andersom (eerst uitgenodigd, daarna gerandomiseerd). Hierdoor is de *compliance rate* relatief laag, en verliezen de schattingen aan statistische kracht. Het bepalen van Hawthorne- en/of John Henry-effecten is niet

¹² Het is ook mogelijk dat mensen die oorspronkelijk toegewezen waren aan de controlegroep toch de interventie ontvangen. Dit was, voor zover wij weten, niet het geval bij de experimenten onder de Participatiewet.

¹³ Een bijkomend probleem is dat de waargenomen en niet-waargenomen kenmerken van deelnemers door zelfselectie mogelijk verschillen van personen die ervoor kiezen om niet deel te nemen. Dit is een potentieel probleem voor LATE-analyses (zie beneden). Voor de door ons geprefereerde ITT-analyses (zie beneden) is het voornaamste probleem van *non-compliance* dat de behandeling 'verwatert'.

¹⁴ Indien de random toewijzing een direct effect heeft op de uitkomst is de *exclusion restriction* geschonden (zie o.a. Angrist en Pischke, 2009; Athey en Imbens, 2017).

¹⁵ Uit een onderzoek naar het effect van wisselende lichtintensiteit op de productiviteit van werknemers in het elektriciteitsbedrijf Hawthorne Works in Chicago in de jaren 1924-1932 kwam naar voren dat de productiviteit vooral veranderde doordat medewerkers meer aandacht kregen, zie: https://en.wikipedia.org/wiki/Hawthorne_effect.

¹⁶ Ook onder klantmanagers zouden gedragseffecten door het experiment zelf kunnen optreden die tot Hawthorne (of John Henry) effecten kunnen leiden.

¹⁷ Volgens de legende werkte John Henry in de jaren 1870 als staalbewerker. Nadat hij hoorde dat zijn productiviteit werd vergeleken met een stoommachine 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.

eenvoudig. Binnen de experimenten zijn ook referentiegroepen van bijstandsgerechtigden bepaald die niet meededen aan het experiment, maar wel tot de doelgroep behoorden. De referentiegroep heeft net als de controlegroep geen behandeling ontvangen. In de gemeenten Deventer en Groningen is een referentiegroep via randomisatie bepaald die vergelijkbaar is met de controlegroep, en die niet is geïnformeerd over het experiment. Door de controlegroep te vergelijken met de aselect bepaalde referentiegroep krijgen we zicht op mogelijke John Henry-effecten.¹⁸ Voor deze gemeenten kijken we in een grafische analyse of de ontwikkeling in de uitkomstmaten voor de controlegroep vergelijkbaar is met die van de referentiegroep.¹⁹

Als aan bovenstaande voorwaarden is voldaan meten we het effect van de beoogde interventie op de groep deelnemers, maar de effecten kunnen anders zijn voor de bredere doelgroep. Als aan bovenstaande voorwaarden is voldaan is er sprake van interne validiteit: de mate waarin binnen de effecten van de interventie causaal kunnen worden verbonden aan die interventie voor de deelnemers. Interne validiteit betekent echter niet automatisch externe validiteit: de mate waarin de effecten voor de experimentele groep van toepassing zijn op de bredere doelgroep.²⁰ De effecten voor de bredere doelgroep kunnen afwijken van de effecten voor de deelnemers, bijvoorbeeld omdat deelnemers een kleinere afstand hebben tot de arbeidsmarkt en/of meer gemotiveerd zijn.

3.3 Intention-to-Treat regressie

Bij een *Intention-to-Treat* (ITT) regressie wordt de daadwerkelijke deelname aan de interventie buiten beschouwing gelaten en worden de uitkomsten vergeleken op basis van de toewijzing aan een interventiegroep of de controlegroep. Het geschatte effect kan geïnterpreteerd worden als het causale effect van de toewijzing aan een interventie, mits de interventie- en controlegroepen vergelijkbaar zijn (door toeval kan dat niet zo zijn) en er geen Hawthorne- of John Henry-effecten optreden. Om het ITT-effect te achterhalen schatten we het volgende lineaire kansmodel (Angrist en Pischke, 2009):

$$y_i^t = \alpha_1 + \sum_{m=1}^M \beta_m D_m + X' \gamma + q + \varepsilon, \quad (1)$$

daarbij is uitkomst y_i^t een binaire variabele gelijk aan 1 bij meer dan 27 uur werk per week (of meer dan 12 uur werk per week voor gedeeltelijke of volledige uitstroom naar werk) op meetmoment t , en D_m is een binaire variabele gelijk aan 1 als persoon i aan interventiegroep m is toegewezen. Vector X is een vector met karakteristieken gemeten 3 maanden voor de start van de interventie, waaronder geslacht, hoogste voltooide opleiding, migratieachtergrond, leefvorm, leeftijd, duur van de lopende uitkering en cumulatief looninkomen in de periode 9 tot 3 maanden voor start van de interventie. Omdat deze karakteristieken gemeten worden voor de interventie variëren ze niet met de tijd. Verder zijn q *fixed effects* voor het startkwartaal van de interventie.

¹⁸ De aselect gekozen referentiegroep in Utrecht is minder vergelijkbaar met de interventie- en controlegroepen, omdat die laatste bestaan uit mensen die zich na uitnodiging hebben aangemeld voor het experiment. In Nijmegen, Tilburg en Wageningen bestaat de referentiegroep uit de deelnemers die zich niet hebben aangemeld voor de interventies. Deze groep is daardoor mogelijk niet meer goed vergelijkbaar met de interventie- en controlegroepen.

¹⁹ De zelfgekozen referentiegroepen zijn niet direct vergelijkbaar met de interventie- en controlegroepen, waardoor mogelijke Hawthorne- of John Henry- effecten niet zijn te onderscheiden van selectie-effecten. In deze notitie presenteren we daarom alleen de vergelijkingen met de aselect toegewezen referentiegroepen die vergelijkbaar zouden moeten zijn met de controlegroep in waargenomen en niet-waargenomen kenmerken, dus voor de gemeenten Deventer en Groningen.

²⁰ De randomisatie in Deventer en Groningen, waarbij bijna de hele bijstandspopulatie is gerandomiseerd, scoort hoog op externe validiteit: de aselect gekozen interventie- en controlegroepen zijn representatief voor de bredere doelgroep. Het nadeel daarvan is dat een lage *compliance rate* het moeilijk maakt statistisch significante effecten te meten voor de interventies. In de andere gemeenten, waar eerst is uitgenodigd en daarna is gerandomiseerd, is de *compliance rate* relatief hoog. Hierdoor neemt de statistische kracht toe bij het bepalen van de causale effecten die intern valide zijn (Athey en Imbens, 2017), maar mogelijk niet zijn te extrapoleren naar de bredere doelgroep.

De coëfficiënten β_m geven, als aan de voorwaarden van de vorige paragraaf is voldaan, de ITT-effecten. De ITT-effecten zijn dan de causale effecten van de toewijzing aan de interventie. Als iedereen de toewijzing opvolgt, is dit bovendien gelijk aan het causale effect van deelname aan de interventie. Maar ook wanneer niet iedereen de toewijzing opvolgt, is het isoleren van het causale effect van toewijzing aan de interventie informatief. Meer problematisch zijn de Hawthorne- of John Henry- effecten, gedragseffecten in de treatment of controlegroep die optreden door de experimentele setting anders dan door de interventie zelf. In dat geval is het causale effect van de interventie niet meer te onderscheiden van andere effecten. Het is dan een combinatie van een causaal effect van de interventie op de interventiegroep en een neveneffect van deelname aan een experimentele setting.

3.4 Local Average Treatment Effect regressie

Bij een *local average treatment effect* (LATE) regressie worden de effecten voor de deelnemers aan de interventie geschat.²¹ Het is mogelijk om, onder bepaalde aannames, het LATE-effect te schatten met een *Two-Stage Least Squares* model (2SLS).^{22, 23} We interpreteren de LATE- schatting als het causale effect van de interventie voor zogenaamde *compliers*. *Compliers* zijn de mensen die na de gerandomiseerde toewijzing besluiten om daadwerkelijk aan de interventie deel te (blijven) nemen.²⁴

De ITT- en LATE-effecten zijn aan elkaar gerelateerd. Het LATE-effect is het ITT-effect gedeeld door de nalevingsgraad, het aandeel mensen dat de aselechte toewijzing opvolgt. Indien iedereen de toewijzing opvolgt is de nalevingsgraad gelijk aan 1, en zijn de ITT- en LATE-effecten aan elkaar gelijk. Een nalevingsgraad kleiner dan 1 leidt tot een LATE-effect dat, in absolute termen, groter is dan het ITT-effect (het totale ITT-effect is het gewogen gemiddelde van een LATE-effect voor de deelnemers en geen effect voor de niet-deelnemers). Bij hoge *compliance rates* is het ITT-effect, het effect van toewijzing aan een interventie, een goede indicatie voor het effect van deelname aan de interventie zelf.

Om de LATE-effecten als causaal te kunnen interpreteren, gelden allereerst dezelfde voorwaarden als bij de ITT, o.a. geen Hawthorne- en John Henry-effecten. Bovendien kunnen de LATE-effecten alleen geschat worden onder bepaalde aannames. Ook als een van deze aannames geschonden is, kan de LATE-schatting niet meer (direct) als causaal effect van de interventie worden geïnterpreteerd (zie voetnoot 22). Een belangrijke eis is dat de relatie tussen de toewijzing en deelname aan de interventie niet zwak is.²⁵

²¹Iedereen die is gestart met een interventie wordt als deelnemer beschouwd. Resultaten waarbij mensen die gestopt zijn als niet-deelnemer worden beschouwd verschillen nauwelijks.

²²LATE is causaal te interpreteren als het effect op de *compliers* onder de volgende aannames (o.a. Angrist en Pischke, 2009): 1) Monotonicity. Dat er geen *defiers* zijn, namelijk dat er geen mensen zijn die altijd ingaan tegen het experiment; 2) Dat het *instrument* (zo goed als) aselechte is. In ons geval is het instrument een aselechte toewijzing; 3) Dat de *exclusion restriction* niet geschonden is, namelijk dat het instrument (de toewijzing) alleen een effect heeft op de uitkomst via deelname aan de interventie; 4) Dat het instrument (toewijzing) een effect heeft op deelname aan de interventie. Een zwakke relatie tussen toewijzing en deelname zou leiden tot een bias in de LATE-schatting.

²³Zie bijlage II voor het LATE-model.

²⁴Dit in tegenstelling tot de *always takers* en, in dit geval niet relevant, de *never takers*. *Always takers* zijn personen die de interventie sowieso zouden kiezen en *never takers* zijn personen die de interventie altijd zouden afwijzen. *Never takers* zijn personen die wel worden uitgenodigd, maar niet deelnemen (o.a. Angrist en Pischke, 2009)

²⁵Een zwakke relatie tussen toewijzing en deelname zou leiden tot bias in de LATE-schatting (o.a. Bound et al., 1995). De vuistregel hiervoor is dat F-statistieken uit de *first-stage* regressie groter zijn dan 10 (o.a. Angrist en Pischke, 2009).

4 Data en beschrijvende statistieken

Voor de empirische analyse koppelen we de data over de toewijzing aan de verschillende groepen van de gemeenten aan de administratieve data van het CBS. De controle- en interventiegroepen zijn meestal goed vergelijkbaar in termen van waargenomen karakteristieken. Uit een vergelijking van de uitkomsten voor de start van het experiment blijkt dat de controlegroep in Nijmegen en in Tilburg voor de start al betere uitkomsten had dan de interventiegroepen. Verder blijkt dat de controlegroep in Groningen het na de start van het experiment significant beter doet dan de aselechte referentiegroep. De effecten gemeten in Groningen, Nijmegen en Tilburg zijn daarom niet direct te interpreteren als (alleen) het effect van de beoogde interventies.

Dit hoofdstuk bespreekt de gemeentelijke data en de CBS-microdata die gebruikt zijn in de empirische analyse. Tevens presenteren we beschrijvende statistieken voor de persoons- en huishoudkenmerken en voor de uitkomsten voor en na de start van het experiment. Deze beschrijvende statistieken geven inzicht in de vergelijkbaarheid van de interventie- en controlegroepen, en of de resultaten van de regressies in het volgende hoofdstuk direct als effect van de beoogde interventies kunnen worden geïnterpreteerd.

4.1 Gemeentelijke data en CBS microdata

Alle gemeenten hebben de data uit hun experimenten aangeleverd aan het CBS. Deze data bevatten informatie over hun lopende bestand met bijstandsgerechtigden, met daarin informatie over de interventie- en controlegroepen. Het gaat hier om de toewijzing aan de groepen, wanneer deze is uitgevoerd, of de bijstandsgerechtigde zich heeft aangemeld voor de interventie, en wanneer de bijstandsgerechtigde is gestopt met de interventie ook de datum waarop de persoon is gestopt. Het CBS heeft deze data geanonimiseerd en gekoppeld aan de CBS microdata met behulp van een rinpersoon-nummer.

De geanonimiseerde gemeentedata hebben wij vervolgens gekoppeld aan de CBS-data. Allereerst aan de gemeentelijke basisadministratie (*GBAPERSOONTAB*). Dit bestand bevat informatie over leeftijd, geslacht en migratieachtergrond (4 categorieën: geen, westers, niet-westers, onbekend). Vervolgens zijn de gegevens gekoppeld aan informatie over het hoogst behaalde opleidingsniveau (4 categorieën: laag, middelbaar, hoog, onbekend) uit het CBS bestand *HOOGSTEOPLTAB*. De duur van de lopende uitkering in maanden voor de start van het experiment is afkomstig van het CBS bestand *BIJSTANDUITKERINGTAB*. Informatie over het type huishouden (5 categorieën: alleenstaande, alleenstaande ouder, gehuwden en samenwonenden zonder kinderen, gehuwden en samenwonenden met kinderen, overig) uit het CBS bestanden *GBAHUISHOUDENSBUS*. Tot slot, gegevens over het aantal gewerkte uren per maand en inkomen uit werk zijn via koppeling verkregen uit de polisadministratie *SPOLISBUS*. Voor onze analyse waren de CBS microdata beschikbaar tot en met het derde kwartaal van 2019.

4.2 Beschrijvende statistieken: vergelijkbaarheid interventie- en controlegroepen

De waarneembare demografische karakteristieken van de 'ITT' interventie- en controlegroepen komen meestal goed overeen. Tabel B3.1a t/m f in de bijlage geven de karakteristieken van de groepen per gemeente weer. Slechts enkele karakteristieken verschillen statistisch significant van elkaar op een significantieniveau van 5%, wat als 'toeval' bestempeld kan worden bij de uitvoering van zoveel statistische testen. Dat geeft aan

dat de aselecte toewijzing van de 'ITT'-groepen, in elk geval op basis van de waarneembare demografische karakteristieken, genoeg op elkaar lijken om straks de verschillen in uitkomsten te mogen interpreteren als het effect van toewijzing aan de verschillende interventies. Voor deze verschillen wordt gecorrigeerd in de regressiemodellen door deze als controlevariabelen op te nemen.²⁶ We analyseren ook de verschillen in inkomen voor de start van het experiment, omdat die kunnen wijzen op verschillen in de uitgangssituatie. In Nijmegen had de ontheffingsgroep in de periode voor de start een 36% lager cumulatief inkomen dan de controlegroep (significant op 10% niveau). In Deventer verdiende de ontheffingsgroep met digitale ondersteuning gemiddeld 44% meer dan de controlegroep (significant op 10% niveau). In Tilburg hadden de twee intensiveringsgroepen (met en zonder inkomensvrijlating) een lager cumulatief inkomen, van respectievelijk 36 en 55% in de periode voor de start van het experiment (significant op 5% niveau). Deze verschillen in cumulatief inkomen doen vermoeden dat er voor de start al enige verschillen waren in de afstand tot de arbeidsmarkt tussen deze controle- en interventiegroepen.

In Deventer en Groningen kunnen we testen voor John Henry-effecten, door de uitkomsten van de controlegroep te vergelijken met de aselecte referentiegroep. Voor Deventer en Groningen vergelijken we daarom ook de karakteristieken van de controle- en referentiegroepen. Beiden zijn aselect gekozen en zouden daarom voor de start vergelijkbaar moeten zijn. Tabellen B3.1a en b laten zien dat deze voor zowel Deventer als Groningen meestal goed vergelijkbaar zijn.²⁷

4.3 Grafische vergelijking uitkomsten over de tijd

Een grafische analyse van het aandeel personen dat gedeeltelijk of volledig werkt laat zien dat de controlegroep in Nijmegen en Tilburg al meer werkte dan de interventiegroepen voor de start van het experiment. Figuur 4.1a tot en met 4.6b laat de uitkomsten zien voor het aandeel personen dat gedeeltelijk of volledig werkt voor de verschillende groepen voor en na de start van het experiment. In Deventer, Groningen, Utrecht, Wageningen en zien we geen duidelijke verschillen tussen de verschillende groepen voor de start van het experiment. In Nijmegen zien we echter dat de controlegroep al snel beter presteert dan de overige groepen wat betreft volledige uitstroom. En voor de uitkomstmaat gedeeltelijke of volledige uitstroom naar werk lijkt de controlegroep het voor de start al veel te doen dan de interventiegroepen. Dit maakt het problematisch om de effecten in Nijmegen direct te interpreteren als alleen het effect van de beoogde interventies. In Tilburg zijn er voor de start van de interventies ook al verschillen tussen de controle- en interventiegroepen. Met name voor gedeeltelijke uitstroom (12 uur werk) is duidelijk te zien dat de controlegroep al voor de start van het experiment boven de andere groepen uitsteekt en erboven blijft. En wat betreft volledige uitstroom lijkt de controlegroep ook al snel beter te presteren dan interventiegroepen. Daarom zijn ook in Tilburg de effecten niet direct te interpreteren als alleen het effect van de beoogde interventies.

In Groningen duiden de verschillen in de ontwikkeling tussen de controlegroep en de aselect gekozen referentiegroep op een John Henry-effect in de controlegroep, in Deventer vinden we daar geen

²⁶ Er zijn enkele verschillen die we, vanwege *multiple hypothesis testing*, kunnen bestempelen als toeval. De volgende karakteristieken verschillen statistisch significant, op een significantieniveau van 5%, van de controlegroep. In Wageningen is de ontheffingsgroep gemiddeld 4 jaar jonger. In Nijmegen is de intensiveringsgroep gemiddeld 2,5 jaar jonger. In Deventer is een groter aandeel van de intensiveringsgroep alleenstaand dan de controlegroep. In Groningen zijn er meer lager opgeleiden in de ontheffings- en de vrijlatinggroep. Ook zijn er minder hoger opgeleiden in de ontheffingsgroep. In Tilburg zijn er relatief veel niet-westerse immigranten in de intensiveringsgroep (zonder extra inkomensvrijlating): 31% ten opzichte van 24% in de controlegroep. Daarnaast is de intensiveringsgroep (met extra inkomensvrijlating) 2,5 jaar ouder dan de controlegroep en is de gemiddelde duur van de bijstandsuitkering circa 16 maanden hoger dan bij de controlegroep.

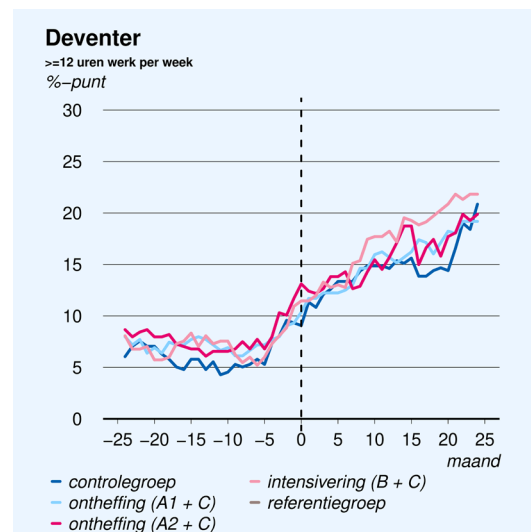
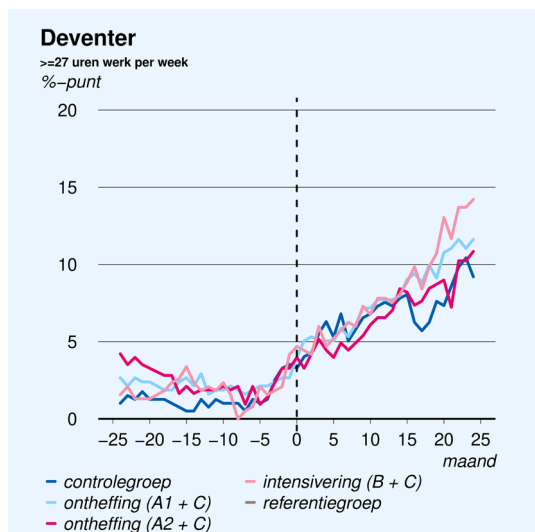
²⁷ In Groningen zijn de volgende verschillen statistisch significant op een significantieniveau van 5%: er zijn minder alleenstaanden in de referentiegroep en minder mensen met een niet-westerse migratieachtergrond dan in de vier interventiegroepen. Tevens is het aandeel mannen kleiner in de referentiegroep dan in de ontheffingsgroep.

aanwijzingen voor.²⁸ In Deventer volgt de controlegroep gedurende het experiment een vergelijkbaar traject als de referentiegroep, en lijkt er geen sprake van een John Henry-effect. In Groningen zien we dat het verloop van de controlegroep meteen bij de start al flink afwijkt van die van de referentiegroep. Dit zou een indicatie kunnen zijn van een John Henry-effect, een gedragseffect in de controlegroep door deelname aan een experiment (zonder een expliciete interventie). De controlegroep zou beter kunnen presteren, omdat ze zich competitief opstellen tegenover de interventiegroepen.²⁹ Verschillen tussen de interventiegroepen en de controlegroep kunnen daardoor niet direct geïnterpreteerd worden als het effect van alleen de beoogde interventie. Het is dan een combinatie van een effect van de interventie op de interventiegroep en een neveneffect van deelname aan een experimentele setting.

In Deventer, Utrecht en Wageningen zien we dat sommige interventiegroepen aan het eind van het experiment betere uitkomsten laten zien. Regressieanalyses moeten uitwijzen of deze verschillen ook statistisch significant verschillen van nul. Voor de gemeenten die voor de start een vergelijkbare ontwikkeling laten zien tussen de controle- en interventiegroepen, en waar we geen John Henry-effecten waarnemen, zien we dat de interventiegroepen zich na de start van het experiment zich wat anders gaan ontwikkelen. Na de start van de interventie zien we in Deventer dat de intensiveringsgroep voor beide uitkomstmaten beter lijkt te presteren dan de controlegroep. Na twee jaar hebben ze een 5 tot 7%-punt hogere kans op volledige uitstroom dan de controlegroep. In Wageningen zien we dat de intensiverings- en ontheffingsgroepen na de start van de interventie de controlegroep achter zich beginnen te laten. Na twee jaar lijken ze wat betreft volledige uitstroom zo'n 5 tot 7 procentpunt boven de controlegroep te liggen. In Utrecht zien we dat alle drie de interventiegroepen 24 maanden na start beter presteren dan de controlegroep, met een 3 tot 4%-punt hogere kans op volledige uitstroom, en 5 tot 8%-punt hogere kans op gedeeltelijke of volledige uitstroom. Regressieanalyses zullen moeten uitwijzen of deze verschillen ook statistisch significant verschillen van nul.

Figuur 4.1a Deventer, meer dan 27 uur werk – over de tijd

Figuur 4.1b Deventer, meer dan 12 uur werk – over de tijd

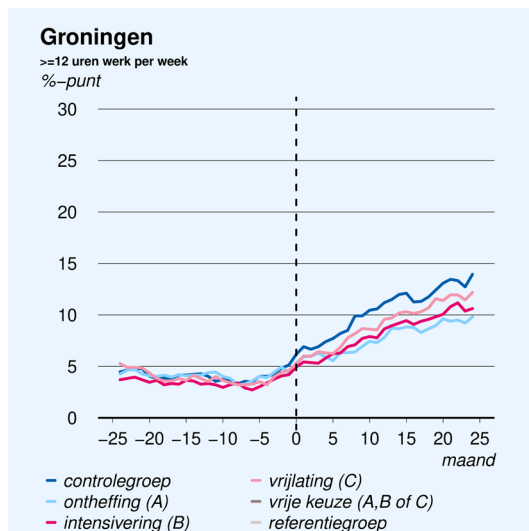
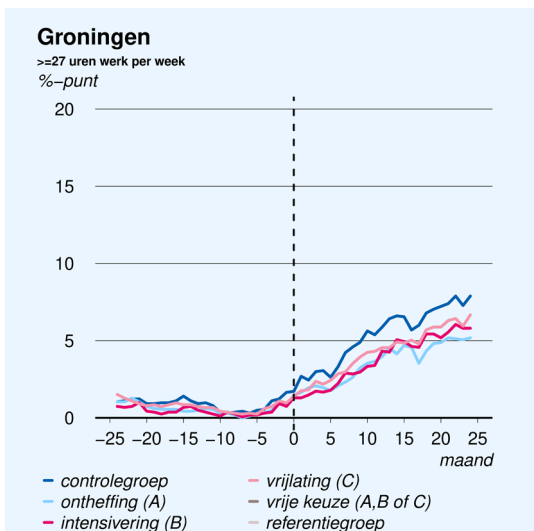


Noot: A1 = Ontheffing, A2 = Ontheffing met app, B = Intensivering, C = Vrijlating inkomensgrens.

²⁸ De aselect gekozen referentiegroep in Utrecht is minder vergelijkbaar met de interventie- en controlegroepen, omdat die laatste bestaan uit mensen die zich na uitnodiging hebben aangemeld voor het experiment.

²⁹ Een mogelijke hypothese is dat de bewoording in de brief is in Groningen een rol heeft gespeeld. Omdat personen in Groningen zijn uitgenodigd na de aselecte toewijzing, is er aan elke groep een andere brief gestuurd. In de brief aan de controlegroep in Groningen stond dat ze ingeloot waren voor de vergelijkingsgroep en nodig waren om het effect van 'Bijstand op maat' te vergelijken. Een mogelijkheid is dat personen in de controlegroep dit hebben ervaren als dat ze minder belangrijk waren dan de interventiegroep. Het is echter onwaarschijnlijk dat de brief de verschillen tussen de controle- en referentiegroep volledig kan verklaren. Omdat de controlegroep en de verschillende interventiegroepen allen een andere brief hebben ontvangen, is het effect van de verschillende brieven niet te onderscheiden van het effect van de verschillende interventies.

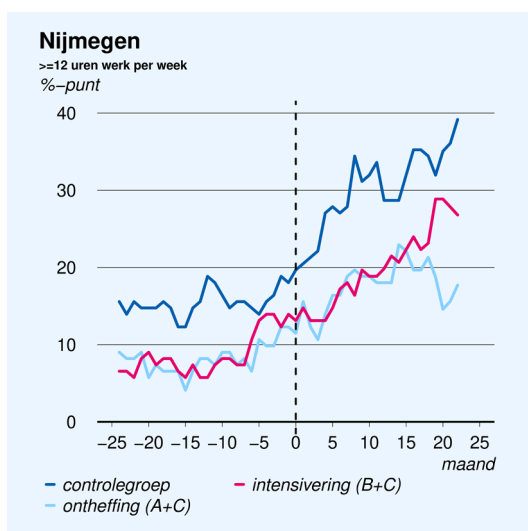
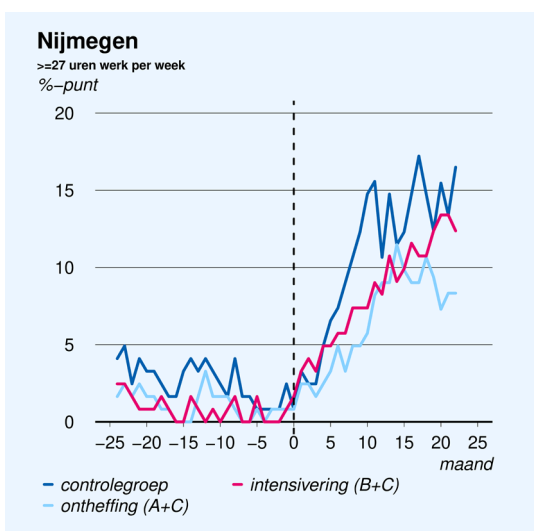
Figuur 4.2a Groningen, meer dan 27 uur werk – over de tijd Figuur 4.2b Groningen, meer dan 12 uur werk – over de tijd



Noot: A = Ontheffing, B = Intensivering, C = Vrijlating inkomensgrens.

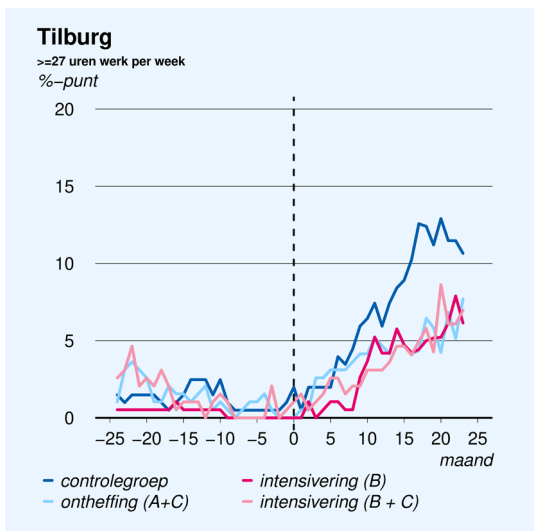
Figuur 4.3a Nijmegen, meer dan 27 uur werk – over de tijd

Figuur 4.3b Nijmegen, meer dan 12 uur werk -over de tijd

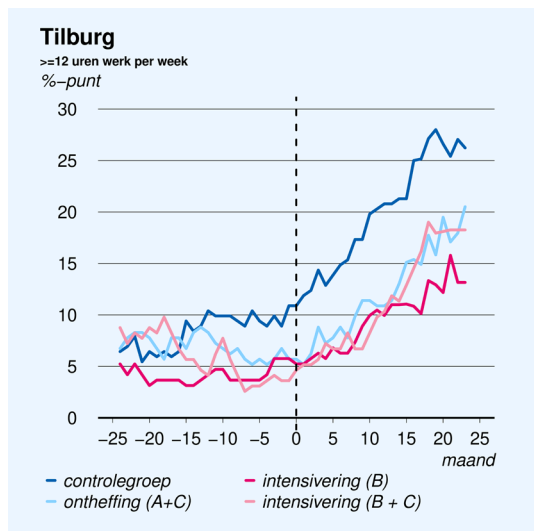


Noot: A = Ontheffing, B = Intensivering, C = Vrijlating inkomensgrens.

Figuur 4.4a Tilburg, meer dan 27 uur werk – over de tijd

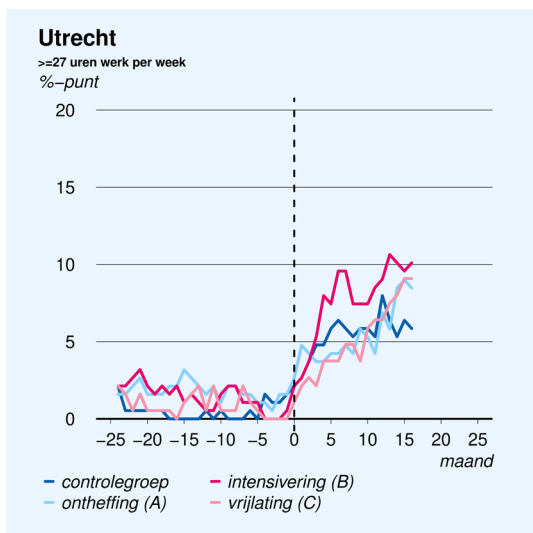


Figuur 4.4b Tilburg, meer dan 12 uur werk – over de tijd

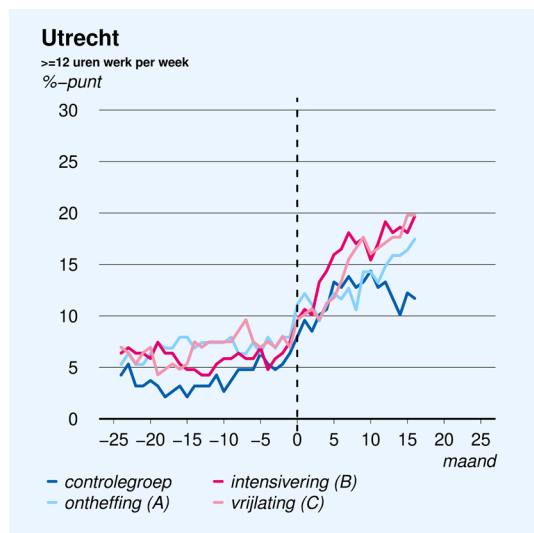


Noot: A = Ontheffing, B = Intensivering, C = Vrijlating inkomensgrens.

Figuur 4.5a Utrecht, meer dan 27 uur werk – over de tijd

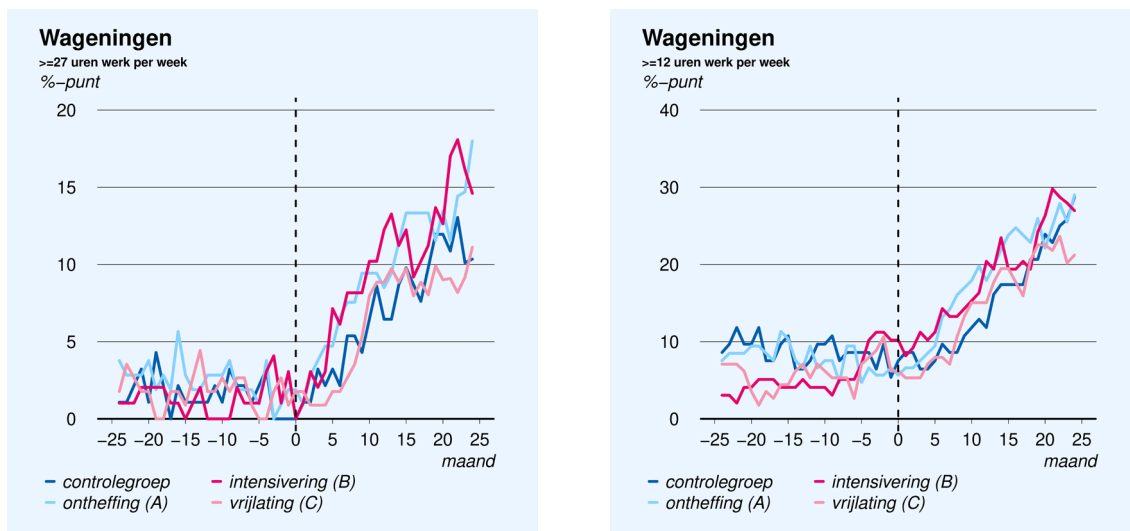


Figuur 4.5b Utrecht, meer dan 12 uur werk – over de tijd



Noot: A = Ontheffing, B = Intensivering, C = Vrijlating inkomensgrens.

Figuur 4.6a Wageningen, meer dan 27 uur werk – over de tijd Figuur 4.6b Wageningen, meer dan 12 uur werk – over de tijd



Noot: A = Ontheffing, B = Intensivering, C = Vrijlating inkomensgrens.

5 Resultaten

Voor de drie gemeenten waarvoor we de effecten kunnen interpreteren als direct effect van de interventie, vinden we het volgende. We vinden geen (statistisch significant) effect van de ontheffing van de arbeids- en re-integratieverplichtingen in de bijstand op de uitstroom naar werk. Daarbij is het verschil met de uitgangssituatie voor een deel van de deelnemers vermoedelijk echter niet groot. We vinden geen effect van een intensievere begeleiding op de volledige uitstroom naar werk. Alleen in Utrecht is er sprake van een effect op de gedeeltelijk uitstroom naar werk (op 5% significantieniveau). 16 maanden na de start van het experiment werkt in Utrecht 20% van de intensiveringsgroep meer dan 12 uur per week, ten opzichte van 12% van de controlegroep. We vinden geen effect van een hogere vrijlating van bijverdiensten in de bijstand op de volledige uitstroom naar werk. Dit was wellicht te verwachten, omdat deelnemers niet meer dan 200 euro per maand mochten bijverdienen naast hun uitkering. Alleen in Utrecht is sprake van een effect op de gedeeltelijke uitstroom naar werk (op 5% significantieniveau). 16 maanden na de start van het experiment werkt in Utrecht 19% van de intensiveringsgroep meer dan 12 uur per week, ten opzichte van 12% van de controlegroep. Voor alle gemeenten geldt verder dat de effecten niet heel precies zijn te schatten door de beperkte aantallen per interventie.

Dit hoofdstuk beschrijft eerst een additionele toets op de vergelijkbaarheid van de interventie- en controlegroepen (placebotoetsen), en vervolgens de ITT- en LATE-regressieresultaten voor de gedeeltelijke en/of volledige uitstroom naar werk. Dit hoofdstuk bespreekt eerst de placebotoetsen, waarbij we de ITT-schattingen in elke maand vanaf 24 maanden voor de start van het experiment grafisch laten zien. Statistisch significante effecten voor of bij de start wijzen op initiële verschillen tussen de interventie- en controlegroepen, waardoor de effecten na de interventie lastig als het directe effect van de beoogde interventie kunnen worden geïnterpreteerd. Vervolgens bespreken we de effecten na de interventie. De hoofdresultaten hebben betrekking op het laatst mogelijke meetmoment. Voor Deventer, Groningen, Tilburg en Wageningen is dit 24 maanden na de start van de interventie. Voor Utrecht is dit 16 maanden na de start van de interventie en voor Nijmegen is dit 22 maanden (zie hoofdstuk 2). Daarnaast bespreken we ook de ontwikkeling van de geschatte effecten na de start van het experiment over de tijd. Ten slotte, voor de

gemeenten waar een aselect gekozen referentiegroep bestaat, Deventer en Groningen, bespreken we ook de resultaten bij een vergelijking met deze groep.

5.1 Placebotoetsen: vergelijking uitkomsten voor de start van het experiment na correctie voor kenmerken

Allereerst toetsen we of er voor de start van het experiment al statistisch en economisch significante verschillen zijn tussen de interventiegroepen en de controlegroep, na correctie voor kenmerken. Voordat we de ITT-resultaten bespreken, kijken we als additionele toets op de vergelijkbaarheid van de interventie- en controlegroepen of er, na correctie voor waarneembare kenmerken, statistisch significante verschillen zijn in de uitkomstmaten voor of op de start van deelname. Bij een aselecte toewijzing zou je verwachten dat er voor en vlak na de start geen verschillen zouden zijn. Als er wel al waarneembare verschillen in de uitkomstmaten zijn, dan kunnen we de eventuele verschillen tussen interventie- en controlegroepen na de start van het experiment niet (alleen) toeschrijven aan de interventie. De figuren in bijlage IV (figuren B4.1 t/m 6) geven de ITT-effecten weer voor elke maand vanaf 24 maanden voor de start van de interventie tot en met 24 maanden, of het laatst beschikbare meetmoment.³⁰

In Nijmegen, Tilburg en Groningen zijn er voor of kort na de start van het experiment al significante verschillen in de uitkomsten, waardoor de effecten aan het eind van het experiment niet direct zijn toe te schrijven aan de beoogde interventies. In Nijmegen zien we voor beide interventiegroepen voor de start, na controle voor waarneembare kenmerken, al regelmatig statistisch significante verschillen. En wanneer de effecten niet statistisch significant zijn, valt de nullijn nog maar net binnen het betrouwbaarheidsinterval. Dit hangt samen met de eerdere constatering in de beschrijvende statistiek dat de controlegroep beter lijkt te presteren voor de start dan de interventiegroepen en daarom minder goed vergelijkbaar is met de interventiegroepen. Wat ook opvallend is, is dat de intensiveringsgroep het tijdens het experiment minder goed gaat doen dan de controlegroep, dit lijkt na een mogelijk initieel *lock-in* effect onwaarschijnlijk naar het eind van het experiment toe. Dit doet vermoeden dat er iets anders aan de hand is in/met de controlegroep in Nijmegen, waardoor deze beter presteert. Voor Tilburg zijn de verschillen voor de start meestal niet significant. Vlak na de start zien we wel snel significante effecten voor de intensiveringsgroep en de vrijlatingsgroep. Voor beide typen interventies is het wederom onwaarschijnlijk dat er negatieve effecten zouden zijn. Met name zou men verwachten dat een hogere vrijlating van bijverdiensten positief zou werken op gedeeltelijke uitstroom (men krijgt dit extra inkomen immers pas bij gedeeltelijke uitstroom). Dit wijst er op dat er mogelijk ook iets aan de hand is in/met de controlegroep in Tilburg (mogelijk een John Henry-effect). Ook voor Groningen zijn er voor de start van de interventie al significante verschillen in volledige uitstroom voor de intensivering, ontheffing, en de keuze interventiegroep. Tevens zien we ook hier na de start onverwacht negatieve effecten voor de intensiveringsgroep (en soms ook voor de vrijlatingsgroep). Wederom wijst dit ook op een afwijkende trend in de controlegroep (zie beneden voor een analyse van John Henry effecten in Groningen). Voor Groningen, Nijmegen en Tilburg zijn de effecten dus niet direct te interpreteren als het effect van (alleen) de interventies.

Ook in Deventer, Utrecht en Wageningen zien we op enkele meetmomenten kleine significante verschillen voor de start van de interventie, waarmee rekening gehouden moet worden bij de interpretatie van de effecten aan het eind. In Deventer, Utrecht en Wageningen zien we op enkele momenten voor de start van de interventie ook kleine significante verschillen. Bij de interpretatie van de regressieresultaten moeten we deze verschillen in acht nemen. In paragraaf 5.4 voeren we een

³⁰ Zie voor vergelijkbare figuren Bolhaar et al. (2019).

gevoeligheidsanalyse uit waarbij wordt gecontroleerd voor verschillen voor de start van het experiment, om te testen hoe gevoelig de resultaten hiervoor zijn.

Er zijn geen verschillen in de uitkomsten tussen de controlegroep en de referentiegroep in Deventer gedurende het experiment, maar wel in Groningen. Voor Deventer en Groningen vergelijken we ook de verschillen in uitstroom tussen de controlegroep en de aselect gekozen referentiegroep, na correctie voor waarneembare kenmerken. In de grafische analyse van paragraaf 4.3 zagen we dat de controlegroep en referentiegroep in Deventer een gelijk verloop volgden. Dit wordt bevestigd in de figuren B4.7a en 7b. Voor Groningen zagen we in paragraaf 4.3 meteen bij de start al verschillen tussen de controle- en de aselecte referentiegroep. Dit wijst mogelijk op een John Henry effect. Ook na correctie voor waarneembare verschillen, zien we in figuren B4.8a en 8b dat de controlegroep bij de start een statistisch significant hogere kans heeft op gedeeltelijke of volledige uitstroom. Met name voor volledige uitstroom wordt dit verschil met de tijd groter. Dit bevestigt dat de geschatte effecten van de interventies in Groningen niet direct te interpreteren zijn als het effect van (alleen) die interventies.

5.2 Effecten van toewijzing aan interventie (ITT)

Tabel 5.1 toont de ITT-resultaten die direct geïnterpreteerd kunnen worden als het effect van de toewijzing aan een interventie op de gedeeltelijke en/of volledige uitstroom naar werk. Niet significante effecten zijn cursief en grijs. Statistisch significante effecten zijn vetgedrukt. Effecten die niet direct zijn te interpreteren als het gevolg van (alleen) de interventie zijn niet weergegeven (zie de bijlage en beneden). De eerste kolom geeft de puntschatting voor elk effect uit het model weer. In de kolom daarnaast staat het effect uitgedrukt als het procentuele verschil met het gemiddelde van de controlegroep. We geven eerst de uitkomsten voor de gemeenten waarbij eerst is uitgenodigd en daarna gerandomiseerd, en daarbinnen voor de gemeenten waarvoor de uitkomsten direct als effect van de beoogde interventie kunnen worden geïnterpreteerd (Utrecht en Wageningen en daarna pas Nijmegen en Tilburg), en daarna voor de gemeenten waar eerst is gerandomiseerd en daarna is uitgenodigd (Deventer en Groningen).

Tabel 5.1 Overzichtstabel effecten interventies, alleen de vetgedrukte coëfficiënten zijn statistisch significant

	Utrecht	Wageningen	Nijmegen	Tilburg	Deventer	Groningen
A: Aandeel dat meer dan 27 uur werk per week werkt						
	Effect in %-punt / 100	Effect in %	Effect in %-punt / 100	Effect in %	Effect	Effect
Ontheffing ^a	+0,018	+31%	+0,061	+59%		Niet direct te interpreteren
Intensivering ^b	+0,043	+73%	+0,046	+45%		Niet direct te interpreteren
Vrijlating ^c	+0,025	+42%	-0,003	-3%		Niet direct te interpreteren
Keuze interventie						Niet direct te interpreteren
Ontheffing +vrijlating					Niet direct te interpreteren	Niet direct te interpreteren
Intensivering + vrijlating					Niet direct te interpreteren	Niet direct te interpreteren
Ontheffing + app + vrijlating					+0,013	+14%
					+0,045	+49%
					+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%		Niet direct te interpreteren
Intensivering	+0,077**	+66%	-0,023	-8%		Niet direct te interpreteren
Vrijlating	+0,065**	+56%	-0,081	-28%		Niet direct te interpreteren
Keuze interventie						Niet direct te interpreteren
Ontheffing +vrijlating					Niet direct te interpreteren	-0,052
Intensivering + vrijlating					Niet direct te interpreteren	-0,003
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. ^e Aandeel personen in de controlegroep dat meer dan 12 uur per week werkt aan het eind van het experiment.						

In Utrecht vinden we geen effect van het ontheffen van arbeids- en re-integratieverplichtingen, en een significant positief effect van intensievere begeleiding en van een hogere vrijlating op de gedeeltelijke of volledige uitstroom naar werk. Er is geen (significant) effect op volledige uitstroom naar werk. In Utrecht is het effect op de volledige uitstroom voor alle drie de interventiegroepen positief, maar niet statistisch significant. Kijkend naar gedeeltelijke of volledige uitstroom, dan hebben zowel de intensiveringsgroep als de vrijlatingsgroep een statistisch significant positief effect, met coëfficiënten van respectievelijk 7,7%- en 6,5%-punt.³¹ ³² Vergeleken met de controlegroep, met een kans van 12%-punt, hebben ze respectievelijk een 66% en 56% hogere kans op meer dan 12 uur werk per week. Het effect van ontheffing op meer dan 12 uur werk per week is niet statistisch significant in Utrecht.³³ Daarbij is wel belangrijk op te merken dat het verschil met de uitgangssituatie voor een deel van de deelnemers vermoedelijk niet groot is, het budget voor de begeleiding van bijstandsgerechtigden is in de uitgangssituatie vaak beperkt.³⁴ Verder zijn de effecten, gegeven de beperkte aantallen deelnemers, niet heel precies te meten.

De effecten in Utrecht lijken toe te nemen over de tijd. Voor de intensiveringsgroep lijken de effecten vanaf de start van het experiment meteen langzaam toe te nemen. Pas in de laatste maanden voor het einde van het experiment is het effect ook statistisch significant voor gedeeltelijke of volledige uitstroom naar werk. Voor de andere twee interventies zijn de puntschattingen in de eerste maanden na de start eerst negatief, na verloop van tijd worden ze steeds positiever, en voor vrijlating uiteindelijk ook significant (op 5%-niveau) voor gedeeltelijke of volledige uitstroom (zie figuur B4.5a-5f en tabel B5.1a en b).

Voorzichtigheid is echter geboden bij het vertalen van de uitkomsten voor de deelnemers in Utrecht naar het effect op de gehele bijstandspopulatie. Ook zijn de effecten mogelijk afhankelijk van de stand van de conjunctuur. De effecten kunnen voor de gehele bijstandspopulatie in de betrokken of andere gemeenten verschillen van personen die deelnemen aan het experiment, voor alle interventies. Deelnemers aan het experiment hebben wellicht een kleinere afstand tot de arbeidsmarkt dan de gemiddelde persoon in de bijstand en/of zijn meer gemotiveerd. Daarnaast vond het experiment plaats in een conjunctureel gunstige periode, ook dit kan invloed hebben op de uitkomsten. De effecten van de interventie kunnen anders zijn indien die in een conjunctureel ongunstigere periode worden geïmplementeerd. Verder blijft relevant dat de effecten niet heel precies waren te meten, gegeven de beperkte aantallen per interventie.

In Wageningen zijn geen van de effecten voor de interventiegroepen statistisch significant verschillend van nul. De effecten zijn aan het eind van het experiment meestal positief voor volledige uitstroom en negatief voor gedeeltelijke of volledige uitstroom, maar niet statistisch significant verschillend van nul (zie tabel B5.1a en b). Figuren B4.6a-6f laten zien dat de schattingen in Wageningen door veel onzekerheid zijn omgeven, waardoor zowel positieve als negatieve effecten niet uit te sluiten zijn. De aantallen in de interventiegroepen blijven te laag om de effecten met meer precisie te kunnen bepalen.

In Nijmegen kunnen we de effecten niet interpreteren als het directe effect van de interventie. Hier waren er in de uitgangssituatie al aanzienlijke verschillen tussen de controle- en interventiegroepen, en ook in het inkomen voor de start van het experiment. Figuren B4.3a-3d laten zien dat er al snel na de start significant negatieve effecten (statistisch significant op 10%-niveau) zijn op gedeeltelijke of volledige uitstroom. Aan het eind van het experiment vinden we op de volledige uitstroom negatieve effecten in de ontheffingsgroep

³¹ Heyma en Van der Werff (2014) vinden een positief effect van verschillende re-integratieactiviteiten voor WW'ers. Deze re-integratieactiviteiten hebben gemiddeld genomen geleid tot een verhoogde kans dat WW'ers weer gaan werken. Koning (2009) vindt een licht positief effect van arbeidsbemiddeling op de korte termijn, maar geen effect op de langere termijn.

³² Knoef en Van Ours (2016) vinden een positief effect van een inkomensvrijlating voor alleenstaande moeders in de bijstand op de kans op een deeltijdbaai. Zij gebruikten hiervoor beleidsexperimenten in 14 gemeenten.

³³ Lammers et al. (2013) en Hulleger en Van Ours (2014) vinden een negatief effect op het vinden van werk van ontheffing onder oudere WW gerechtigden.

³⁴ Zie ook Koning (2017).

(statistisch significant op 10%-niveau) en intensiveringsgroep (niet statistisch significant). Kijkend naar de gedeeltelijke of volledige uitstroom, dan zijn beide negatieve effecten significant op 5%-niveau (zie tabel B5.1a en b). Maar gezien de verschillen die er reeds voor de start van het experiment waren kunnen deze effecten niet geïnterpreteerd worden als het directe effect van de interventies.

In Tilburg zijn de effecten ook niet direct te interpreteren als het effect van de interventies. In Tilburg zagen we voor de controlegroep ook al een aanzienlijk betere uitgangssituatie dan voor de interventiegroepen bij gedeeltelijke of volledige uitstroom naar werk, waardoor de effecten ook niet als het directe effect van de interventies te interpreteren zijn. We vinden negatieve effecten voor alle drie de interventiegroepen voor zowel gedeeltelijke en volledige uitstroom naar werk (zie tabel B5.1a en b). Voor volledige uitstroom zijn er significant negatieve effecten voor de intensiverings- en ontheffingsgroepen met een hogere vrijlating voor bijverdiensten. Voor gedeeltelijke of volledige uitstroom zijn er significant negatieve effecten voor de intensiveringsgroep (exclusief extra inkomensvrijlating) en de ontheffingsgroep. Gezien de verschillen voor de start van het experiment zijn deze effecten echter wederom niet direct toe te schrijven aan de interventies.

In Deventer zijn de effecten van de interventies niet statistisch significant verschillend van nul. In Deventer vinden we geen statistisch significant effect voor de interventiegroepen op de kans van werk voor meer dan 27 uur (zie tabel B5.2a). Ook bij de kans op een gedeeltelijke of volledige uitstroom naar werk is het effect voor de verschillende interventies niet statistisch significant (zie tabel B5.2b). Figuren B4.1a-1f laten zien dat er soms een positieve trend voor de twee intensiveringsgroepen (met en zonder digitale ondersteuning) lijkt te zijn, deze is echter omgeven met veel onzekerheid. Onze conclusie is dat er weinig significante effecten zijn van de interventies in Deventer. Wel is het zo dat Deventer meerdere startmomenten heeft. De later gestarte groepen kunnen dus minder lang gevolgd 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. Ten slotte zijn er ook geen significante effecten wanneer de uitkomsten van de interventiegroepen vergeleken worden met de aselechte referentiegroep (zie tabel B5.3a en b).

In Groningen zijn geen van de effecten ten opzichte van de controlegroep direct te interpreteren, omdat de controlegroep zich beter ontwikkelt dan de aselechte referentiegroep na de start van de interventies. Zowel de ontheffingsgroep als de intensiveringsgroep hebben een significant negatief effect op de gedeeltelijke of volledige uitstroom naar werk. Het effect van de vrije keuze interventie is significant negatief bij de gedeeltelijke uitstroom. Figuren B4.2a-h laten zien dat deze negatieve effecten al vlak na de start van het experiment ontstaan. Zoals we eerder constateerden uit figuur 4.2a en b ontwikkelde de controlegroep zich echter beter dan de aselekt gekozen referentiegroep, wat de aanwezigheid van een John Henry effect suggereert. De gemeten effecten in tabel B5.2a en f kunnen daarom niet direct aan de interventie toegeschreven worden omdat ze een combinatie zijn van het effect van de interventie voor de interventiegroepen en een mogelijk John Henry effect in de controlegroep. We zien geen significante effecten wanneer we de uitkomsten van de interventiegroepen vergelijken met de aselechte referentiegroep (zie tabel B5.3a en b).

5.3 Effecten deelname aan interventie (LATE)

Alleen voor Utrecht vinden we significante LATE-effecten van intensivering en vrijlating van de inkomensgrens op gedeeltelijke of volledige uitstroom naar werk. In Utrecht zijn geen van de LATE-effecten op volledige uitstroom aan het eind van het experiment statistisch significant (zie tabel B6.1a). Voor gedeeltelijke of volledige uitstroom zien we dat intensivering de uitstroom verhoogt met 8,2%-punt en een hogere vrijlating van bijverdiensten leidt tot een 7,1%-punt hogere kans op gedeeltelijke uitstroom (zie tabel

B6.1a).³⁵ Vergeleken met de gemiddelde kans op uitstroom in de controlegroep neemt deze kans toe met respectievelijk 70% en 60%. De effecten zijn vanwege de hoge *compliance rates* sterk vergelijkbaar met de geschatte ITT-effecten. Wegens de verschillen die voor de start al aanwezig zijn, tussen deze interventiegroepen en de controlegroep voor sommige uitkomstmaten, kunnen we deze effecten echter niet volledig toeschrijven aan de interventies (zie beneden voor een gevoeligheidsanalyse waar we corrigeren voor verschillen in de uitkomst voor de start van het experiment).

In Deventer en Wageningen zijn de LATE-effecten net als de ITT-effecten niet significant. In Deventer zijn geen van de LATE-effecten aan het eind van het experiment significant, zoals ook het geval was bij de ITT-effecten. Dit geldt zowel voor de volledige uitstroom naar werk als de gedeeltelijke of volledige uitstroom naar werk. De LATE-effecten, gemeten aan het eind van het experiment, zijn net als de ITT-effecten ook niet significant in Wageningen (zie tabel B6.1a en b en B6.2a en b).

5.4 Gevoeligheidsanalyses

De resultaten zijn robuust voor verschillende specificaties met alternatieve controlevariabelen. Als gevoeligheidsanalyse hebben we ook andere specificaties van de ITT- en LATE-modellen geschat. Zo onderzoeken we of de geschatte effecten gevoelig zijn voor de opname van andere controlevariabelen. Schattingen met leefvorm uit *BIJSTANDUITKERINGTAB* in plaats van huishoudtype volgens *GBAHUISHOUDENSBUS*, extra termen als leeftijd kwadraat, of een dummy voor looninkomen voor de start van het experiment, en de logaritme van de duur van de bijstand (in plaats van het lineaire niveau van de duur) leveren vergelijkbare resultaten op. Schattingen waarbij het cumulatieve looninkomen 12 tot 24 maanden voor 3 maanden voor de start van het experiment (in plaats van de 6 maanden in de basisspecificatie) opnemen leiden ook tot vergelijkbare resultaten. Wel is het effect van een hogere vrijlating van bijverdiensten in Utrecht dan net niet meer significant op een significantieniveau van 10%.³⁶ Verder leidt ook een logistische regressie tot vergelijkbare resultaten als het lineaire kansmodel.³⁷

Alternatieve uitkomstmaten als meer dan 32 of meer dan 8 uur werk per week schetsen een zelfde beeld in de gemeenten waar de effecten als het directe effect van de interventies geïnterpreteerd kan worden. In Deventer en Wageningen zijn geen van de effecten statistisch significant. In Utrecht zijn er geen significante effecten op meer dan 32 uur werk per week. Wat betreft meer dan 8 uur werk per week heeft ontheffing nog steeds geen significant effect. Intensivering verhoogt de kans op meer dan 8 uur werk in Utrecht met 7,7%-punt (+63%, significant op 5%-niveau) ten opzichte van een gemiddelde van 12% in de controlegroep. En een hogere vrijlating van bijverdiensten verhoogt de kans in Utrecht op meer dan 8 uur werk per week met 8,6%-punt (+70%, significant op 5%-niveau) ten opzichte van de controlegroep (zie tabel B7.1a en b en B7.2a en b).

Ook bij de alternatieve uitkomstmaten als het verdienen van meer dan 50% of meer dan 75% van het wettelijk minimumloon vinden we vergelijkbare resultaten voor Deventer, Utrecht en Wageningen. Voor de uitkomstmaten voor het verdienen van meer dan 50% of 75% van het wettelijke minimumloon (wml) vinden we in Deventer en Wageningen ook geen significante effecten. In Utrecht zijn er geen significante effecten op het verdienen van meer dan 75% van wml. Wel vinden we een statistisch significant verhoging van 6,1%-punt (+60%) van de kans op meer dan 50% wml in de intensiveringgroep (significant op 10%-niveau) ten opzichte van de 10% in de controlegroep. Het effect in de vrijlating van inkomensgrensgroep is niet significant

³⁵ De F statistieken uit de *first-stage* regressie zijn groter dan 10. Er is niet sprake van een zwakke relatie tussen toewijzing en deelname.

³⁶ Opname van een controlevariabele voor cumulatief inkomen in maand 15 tot 3 maanden voor de start gaat de p-waarde naar 0,104, en bij een controlevariabele voor cumulatief inkomen in maand 25 tot 3 maanden voor de start gaat de p-waarde naar 0,111. De coëfficiënten van de interventies blijven hetzelfde.

³⁷ De resultaten van de gevoeligheidsanalyses zijn op verzoek beschikbaar.

op 10%-niveau, maar suggereert een verhoging van 4,4 %-punt (44%) ten opzichte van de 10% in de controlegroep (zie tabel B7.4 en b en B7.5a en b).

Regressie-modellen waarbij we het aantal gewerkte uren opnemen als afhankelijke variabele leiden doorgaans tot vergelijkbare conclusies voor de gemeenten waar de effecten direct toegeschreven kunnen worden aan de interventies. Het effect van intensievere begeleiding is dan wel niet meer statistisch significant in Utrecht. Als we in plaats van binaire variabelen voor meer dan 27 of meer dan 12 uur werk per week een (continu) model schatten van gewerkte uren (inclusief de nullen) vinden we vergelijkbare effecten. In Deventer en Wageningen zijn de effecten niet significant. In Utrecht is er wederom geen effect van ontheffing. Het effect van intensievere begeleiding in Utrecht is net niet significant op een significantieniveau van 10% ($p=0,11$), maar de coëfficiënt suggereert een verhoging van 1,6 uur werk per week (+43%) ten opzichte van de gemiddeld 3,6 gewerkte uren per week in de controlegroep. In Utrecht vinden we dat een hogere vrijlating van bijverdiensten het aantal gewerkte uren met 1,8 uur per week (+50%) verhoogd, ten opzichte van de gemiddeld 3,6 gewerkte uren per week in de controlegroep (statistisch significant op 10% niveau) (zie tabel B7.3a en b³⁸).

Wanneer we controleren voor de verschillen voor de start van de interventies vinden we wat kleinere maar nog steeds significante puntschattingen (op 10%-niveau) van het effect van intensievere begeleiding en een hogere vrijlating van bijverdiensten in Utrecht. Zoals reeds opgemerkt hadden de interventiegroepen voor de start op enkele meetmomenten in Utrecht al een hogere kans op werk. Zo is in figuren B4.5a-f te zien dat er soms een significant verschil was (2%-punt) in de kans op volledige uitstroom voor alle drie de interventiegroepen. Ook bij gedeeltelijke of volledige uitstroom waren de verschillen op enkele meetmomenten voor de start significant (4%-punt). Als gevoeligheidsanalyse doen we ook schattingen waarbij we als controlevariabelen de uitkomsten 24, 16 en 8 maanden voor de start meenemen. Dit leidt tot iets kleinere puntschattingen met coëfficiënten van 6,8% voor intensivering (significant op 5%-niveau) en 4,4% voor de hogere vrijlating van bijverdiensten (significant op 10%-niveau). Verder is het effect van ontheffing nog steeds niet significant. Na controle voor de verschillen in uitkomst voor de start van het experiment blijven de effecten in Deventer en Wageningen niet significant (zie tabel B7.6a en b en B7.7a en b).

6 Discussie

De effecten zijn niet direct toepasbaar voor de gehele bijstandspopulatie en kunnen afhankelijk zijn van de stand van de conjunctuur. Voor zover er effecten zijn, komen deze voort uit een experiment uit een deel van de bijstandspopulatie. Zoals reeds opgemerkt kan het effect anders zijn wanneer de hele bijstandspopulatie deel zou nemen aan de interventie. Bovendien zou het effect, wegens verschillen tussen de gemeenten of wegens algemeen-evenwicht effecten (Duflo et al., 2007), ook anders kunnen zijn indien de interventie in alle gemeenten wordt uitgerold. Ook de stand van de conjunctuur kan van invloed zijn op de resultaten, de experimenten werden uitgevoerd in een periode met een sterk dalende werkloosheid. De resultaten kunnen verschillen in een economische neergang met oplopende werkloosheid (zie ook CPB, 2020).

Nader onderzoek gewenst met data over zelfstandigen. Het huidige onderzoek maakt gebruik van de CBS-polisdata, een data set die alleen gegevens bevat over werknemers en niet over zelfstandigen. Data waarmee zelfstandigen voldoende lang gevolgd kunnen worden zijn bij het CBS nog niet beschikbaar voor de gehele periode van het experiment. De huidige resultaten kunnen daarom het effect van interventies onder-

³⁸De effecten in tabel B7.3a en b zijn geschat op een OLS model van het aantal gewerkte uren per maand.

overschatten. Dit effect is op voorhand niet eenduidig omdat uitstroom via werk als zelfstandige niet wordt gemeten in zowel de interventie- als de controlegroepen. Vervolgonderzoek met data over zelfstandig ondernemerschap is nodig om dit verder te onderzoeken.³⁹

Nader onderzoek gewenst naar de onderliggende mechanismen en mogelijke John Henry-effecten. Om de effecten per gemeente en de verschillen tussen de gemeenten beter te begrijpen is het gewenst om nader onderzoek te doen naar de onderliggende mechanismen die bij de verschillende interventies mogelijk een rol hebben gespeeld. Ook is het gewenst om nader onderzoek te doen naar de al (Groningen) dan niet (Deventer) afwijkende patronen tussen de controlegroep en de aselechte referentiegroep. Dit duidt mogelijk op een John Henry-effect, wat ook bij de interventies in de andere gemeenten een rol kan spelen, maar niet gemeten kon worden.

De langeretermijneffecten kunnen verschillen van de onderzoeksperiode. Met de beschikbare data kunnen we de effecten meten tot (maximaal) 24 maanden na de start van het experiment. De langeretermijneffecten kunnen afwijken van de effecten gemeten over deze periode. Dit geldt met name voor volledige uitstroom, omdat men wellicht eerst met een kleinere baan begint en deze daarna uitbreidt. De analyses laten zien dat de effecten over de tijd toe lijken te nemen. In vervolgonderzoek is het mogelijk om te kijken naar de langeretermijneffecten van de interventies.

Minder verschillende interventies tegelijk kunnen bij toekomstige experimenten leiden tot meer precisie. Het initiatief om door middel van experimenten te onderzoeken welke interventies het beste werken voor de bijstandpopulatie, is een goede insteek voor *evidence-based* beleid. Het experiment is, zoals ook door ZonMw beoordeeld, doorgaans goed opgezet en uitgevoerd. Desalniettemin hebben we nog een aantal aanbevelingen voor toekomstige experimenten. Ten eerste, het is belangrijk om veel waarnemingen in de interventie- en controlegroepen te hebben. De ambitie en de noodzaak om in elke gemeente te experimenteren met meerdere interventies heeft geleid tot kleinere groepen per interventie. In kleinere groepen is het risico groter dat er verschillen ontstaan tussen de groepen, waardoor de effecten van een interventie minder precies zijn te schatten. Dit is ook gebleken in het onderzoek. Ten tweede is het aan te raden om de interventies tussen gemeenten zoveel mogelijk gelijk te houden, wederom om het effect per interventie meer precies te kunnen bepalen. Ten derde, de volgorde van uitnodiging en randomisatie is belangrijk voor de *compliance rate*. Bij een lage *compliance rate* verwatert het effect, waardoor effecten lastig vast te stellen zijn. Verder helpt deze volgorde bij het zoveel mogelijk gelijk behandelen van de controle- en interventiegroepen voor de start van de interventie.

³⁹ Daarbij is het ook interessant om de uitstroom volgens de CBS data te vergelijken met de uitstroom volgens de gemeentedata.

Literatuur

Angrist J., en J. Pischke, 2009, *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press, Princeton.

Athey, S en G. Imbens, 2017, The econometrics of randomized experiments, *Handbook of Economic Field Experiments*, vol. 1: 73-140.

Bolhaar, J., N. Ketel en D. van Vuuren, 2019, Job search periods for welfare applicants: Evidence from a randomized experiment, *American Economic Journal: Applied Economics*, vol. 11(1): 92-215.

Bound, J., D. Jaeger en R. Baker, 1995, Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak, *Journal of the American Statistical Association*, vol. 90(43): 443-450.

CPB, 2020, Update Kansrijk arbeidsmarktbeleid – Actief arbeidsmarktbeleid, CPB, Den Haag.

Duflo, E., R. Glennerster en M. Kremer, 2008, Using randomization in development economics research: A toolkit, in: Schultz, T. en J. Strauss (eds), *Handbook of Development Economics*, vol. 4, North Holland, Amsterdam.

Hullegie, P, en J. van Ours, 2014, Seek and ye shall find: How search requirements affect job finding rates of older workers, *De Economist*, vol. 162: 377-395.

Knoef, M. en J. van Ours, How to stimulate single mothers on welfare to find a job: evidence from a policy experiment, *Journal of Population Economics*, vol. 29(4): 1025-1061.

Koning, P., 2009, The effectiveness of public employment service workers in the Netherlands, *Empirical Economics*, vol. 37: 393-409.

Koning, P., 2017, Basisinkomen licht, *Economisch Statistische Berichten*, 11 juli 2017.

Lammers, M., H. Bloemen en S. Hochguertel, 2013, Job search requirements for older unemployed: Transitions to employment, early retirement and disability benefits, *European Economic Review*, vol. 58: 31-57.

Heyma, A. en S. van der Werff, 2014, Een goed gesprek werkt: (Kosten)effectiviteit van re-integratiedienstverlening UWV voor de WW-instroompopulatie 2008-2010, SEO, Amsterdam.

ZonMw, 2020, Rapportage procesevaluatie experimenten Participatiewet, Den Haag.

Bijlage I – Aanvullende informatie gemeenten

Tabel B1.1 Overzichtstabel aantallen verschillende groepen, per gemeente

Deventer	uitgenodigd	aangemeld	gestart	gestopt
Groep 1 (A1 + C)	376	117	113	15
Groep 2 (A2 + C)	427	84	74	13
Groep 3 (B + C)	384	93	84	14
Controlegroep	397	93	93	0
Referentiegroep	205	n.v.t.	n.v.t.	n.v.t.
Groningen				
Groep 1 (A)	1641	183	183	19
Groep 2 (B)	1620	144	144	40
Groep 3 (C)	1649	153	153	24
Groep 4 (A, B of C)	1647	188	188	24
Controlegroep	1635	222	222	25
Referentiegroep	146	n.v.t.	n.v.t.	n.v.t.
Nijmegen				
Groep 1 (A + C)	122	119	110	5
Groep 2 (B + C)	122	122	100	2
Controlegroep	122	120	94	8
Referentiegroep ⁴⁰	n.v.t.	n.v.t.	n.v.t.	n.v.t.
Tilburg				
Groep 1 (A + C)	193	193	193	46
Groep 2 (B)	191	191	191	75
Groep 3 (B+C)	194	194	194	53
Controlegroep	202	202	202	61
Referentiegroep	n.v.t.	n.v.t.	n.v.t.	n.v.t.
Utrecht				
Groep 1 (A)	189	189	182	13
Groep 2 (B)	188	188	176	15
Groep 3 (C)	187	187	172	12
Controlegroep	188	188	180	8
Referentiegroep	198	n.v.t.	n.v.t.	n.v.t.
Wageningen				
Groep 1 (A)	106	106	106	20
Groep 2 (B)	98	98	98	15
Groep 3 (C)	113	113	113	16
Controlegroep	93	93	93	13
Referentiegroep	n.v.t.	n.v.t.	n.v.t.	n.v.t.

⁴⁰ In deze tabel laten we cijfers zien van de aselect gekozen referentiegroepen. Nijmegen, Wageningen, en Tilburg hebben referentiegroepen samengesteld uit de personen die uitgenodigd waren voor een interventie, maar die zich niet hebben aangemeld. Deze laten wij buiten beschouwing omdat deze wegens selectie minder goed vergelijkbaar zijn met de overige groepen.

Bijlage II – Local Average Treatment Model

Om het LATE-effect te achterhalen schatten we het volgende lineaire kansmodel met zSLS:

$$y_i^t = \delta + \sum_{m=1}^M \rho_m \widehat{I}_m + X' \lambda + q + \eta, \quad (2)$$

daarbij zijn de uitkomst y_i^t , D_m , X en q hetzelfde als in vergelijking (1). \widehat{I}_m is de voorspelde deelname aan de interventie uit de onderstaande *first-stage* regressie:

$$I_m = \psi + \sum_{m=1}^M \theta_m D_m + X' \tau + q + \nu, \quad (3)$$

waarbij θ_m de *compliance rate* is.⁴¹ De coëfficiënt ρ_m in vergelijking (2) geeft de LATE-effecten weer, de effecten van de interventie op de behandelde groep.

De ITT- en LATE-effecten zijn aan elkaar gerelateerd. Het LATE-effect is het ITT-effect gedeeld door θ_m . Indien iedereen de toewijzing opvolgt, is de nalevingsgraad θ_m gelijk aan 1 en daarmee zijn de ITT- en LATE-effecten aan elkaar gelijk. Een nalevingsgraad kleiner dan 1 leidt tot een LATE-effect dat, in absolute termen, groter is dan het ITT-effect.

⁴¹ Deze wordt gemeten aan het begin van de interventie, neemt iemand die is uitgenodigd ook deel. Voor de analyses in deze notitie beschouwen we ook mensen die vroegtijdig stoppen met de interventie als deelnemer. Alternatieve analyses, waarbij deze worden beschouwd als niet-deelnemer, geven vergelijkbare resultaten.

Bijlage III – Beschrijvende Statistieken

Tabel B3.1a Beschrijvende statistieken Deventer, karakteristieken 3 maanden voor start

	Controle	Referentie	Ontheffing m/ vrijlating ink-grens	Intensivering m/ vrijlating ink-grens	Ontheffing m/ app en vrijl. ink-grens
Leeftijd	46,78 (10.18)	46,19 (9.97)	45,99 (10.16)	46,31 (10.31)	46,16 (10.20)
Man (1/0)	0.45 (0.50)	0.45 (0.50)	0.45 (0.50)	0.48 (0.50)	0.45 (0.50)
Alleenstaand (1/0)	0.42 (0.49)	0.50* (0.50)	0.44 (0.50)	0.50** (0.50)	0.47 (0.50)
Alleenstaande ouder (1/0)	0.26 (0.44)	0.23 (0.42)	0.25 (0.43)	0.23 (0.42)	0.25 (0.44)
Alleenstaanden en gehuwden (1/0)	0.11 (0.31)	0.10 (0.30)	0.12 (0.32)	0.11 (0.31)	0.11 (0.32)
Alleenstaanden en gehuwden met kinderen (1/2)	0.21 (0.41)	0.18** (0.38)	0.18 (0.39)	0.15** (0.35)	0.15** (0.36)
Overig/onbekend (1/0)	0.01 (0.11)	0.00 (0.00)	0.01 (0.11)	0.02† (0.13)	0.01† (0.12)
Autochtoon (1/0)	0.52 (0.50)	0.53 (0.50)	0.52 (0.50)	0.52 (0.50)	0.56 (0.50)
Westerse immigrant (1/0)	0.11 (0.31)	0.14 (0.34)	0.12 (0.33)	0.14 (0.35)	0.10 (0.29)
Niet-westerse immigrant (1/0)	0.37 (0.48)	0.33 (0.47)	0.36 (0.48)	0.34 (0.47)	0.34 (0.47)
Lagere opleiding (1/0)	0.54 (0.50)	0.55 (0.50)	0.54 (0.50)	0.48 (0.50)	0.53 (0.50)
Middelbare opleiding (1/0)	0.29 (0.46)	0.30 (0.46)	0.28 (0.45)	0.36* (0.48)	0.29 (0.45)
Hogere opleiding (1/0)	0.10 (0.30)	0.08 (0.27)	0.13† (0.33)	0.09 (0.28)	0.10 (0.30)
Opleiding onbekend (1/0)	0.07 (0.25)	0.07 (0.26)	0.05 (0.22)	0.07 (0.26)	0.08 (0.27)
Duur bijstand (in maanden)	64.00 (76.30)	62.07 (70.63)	61.79 (71.50)	55.95 (60.21)	61.11 (68.97)
Looninkomen in 9-3 maanden voor start (in Euro)	425.94 (1.198,56)	469.13 (1.435,78)	599.27 (1.816,53)	494.54 (1.475,74)	614.71* (1.697,13)
Aantal observaties	397	205	376	384	427

Noot: verschil met controlegroep * p < 0,10%; ** p < 0,05%; verschil met referentiegroep † p < 0,10%; †† p < 0,05%.

Tabel B3.1b Beschrijvende statistieken Groningen, karakteristieken 3 maanden voor start

	Controle	Referentie	Ontheffing	Intensivering	Vrijlating Inkomensgrens	Keuze
Leeftijd	44,80	44,66	45,09	44,40	44,42	44,75
	(10,28)	(10,97)	(10,45)	(10,44)	(10,21)	(10,25)
Man (1/0)	0,53	0,45†	0,54††	0,53†	0,52	0,54††
	(0,50)	(0,50)	(0,50)	(0,50)	(0,50)	(0,50)
Alleenstaand (1/0)	0,56	0,50	0,56	0,56	0,57†	0,57
	(0,50)	(0,50)	(0,50)	(0,50)	(0,49)	(0,50)
Alleenstaande ouder (1/0)	0,21	0,24	0,20	0,20	0,21	0,21
	(0,41)	(0,43)	(0,40)	(0,40)	(0,40)	(0,41)
Alleenstaanden en gehuwden (1/0)	0,08	0,10	0,08	0,08	0,07	0,08
	(0,27)	(0,30)	(0,27)	(0,27)	(0,25)	(0,28)
Alleenstaanden en gehuwden met kinderen (1/2)	0,14	0,14	0,15	0,14	0,14	0,13
	(0,35)	(0,35)	(0,36)	(0,35)	(0,35)	(0,33)
Overig/onbekend (1/0)	0,01	0,02	0,01	0,01	0,01	0,01
	(0,11)	(0,14)	(0,09)	(0,10)	(0,10)	(0,11)
Autochtoon (1/0)	0,56	0,66††	0,55††	0,57††	0,55††	0,55††
	(0,50)	(0,48)	(0,50)	(0,50)	(0,50)	(0,50)
Westerse immigrant (1/0)	0,11	0,11	0,11	0,10	0,12	0,11
	(0,31)	(0,31)	(0,31)	(0,30)	(0,33)	(0,31)
Niet-westerse immigrant (1/0)	0,34	0,23††	0,34††	0,33††	0,33††	0,34††
	(0,47)	(0,42)	(0,48)	(0,47)	(0,47)	(0,47)
Lagere opleiding (1/0)	0,40	0,43	0,45**	0,42	0,45**	0,42
	(0,49)	(0,50)	(0,50)	(0,49)	(0,50)	(0,49)
Middelbare opleiding (1/0)	0,37	0,35	0,35	0,36	0,35	0,38
	(0,48)	(0,48)	(0,48)	(0,48)	(0,48)	(0,48)
Hogere opleiding (1/0)	0,18	0,18	0,15**	0,17	0,17	0,17
	(0,38)	(0,38)	(0,35)	(0,38)	(0,37)	(0,37)
Opleiding onbekend (1/0)	0,05	0,04	0,05	0,05	0,04*	0,04
	(0,22)	(0,20)	(0,22)	(0,22)	(0,19)	(0,19)
Duur bijstand (in maanden)	65,36	64,72	67,41	64,25	63,86	64,58
	(62,91)	(59,21)	(65,37)	(62,93)	(60,60)	(61,72)
Looninkomen in 9-3 maanden voor start (in Euro)	298,43	308,14	272,90	242,29	252,82	311,53
	(1.112,02)	(1.009,93)	(979,13)	(896,52)	(931,01)	(1.063,03)
Aantal observaties	1.635	146	1.641	1.620	1.649	1.647

Noot: verschil met controlegroep * p < 0,10%; ** p < 0,05%; verschil met referentiegroep † p < 0,10%; †† p < 0,05%.

Tabel B3.1c Beschrijvende statistieken Nijmegen, karakteristieken 3 maanden voor start

	Controle	Ontheffing met vrijlating ink.grens	Intensivering met vrijlating ink.grens
Leeftijd	45,52 (9,66)	45,34 (9,70)	42,98** (9,97)
Man (1/0)	0,47 (0,50)	0,52 (0,50)	0,46 (0,50)
Alleenstaand (1/0)	0,58 (0,50)	0,52 (0,50)	0,61 (0,49)
Alleenstaande ouder (1/0)	0,29 (0,45)	0,34 (0,47)	0,28 (0,45)
Alleenstaanden en gehuwden (1/0)	0,02 (0,16)	0,04 (0,20)	0,02 (0,13)
Alleenstaanden en gehuwden met kinderen (1/2)	0,09 (0,29)	0,07 (0,26)	0,04 (0,20)
Overig/onbekend (1/0)	0,02 (0,13)	0,03 (0,18)	0,06* (0,23)
Autochtoon (1/0)	0,66 (0,47)	0,65 (0,48)	0,62 (0,49)
Westerse immigrant (1/0)	0,09 (0,29)	0,11 (0,32)	0,10 (0,30)
Niet-westerse immigrant (1/0)	0,25 (0,43)	0,24 (0,43)	0,28 (0,45)
Lagere opleiding (1/0)	0,24 (0,43)	0,24 (0,43)	0,21 (0,41)
Middelbare opleiding (1/0)	0,37 (0,48)	0,37 (0,48)	0,48* (0,50)
Hogere opleiding (1/0)	0,34 (0,48)	0,36 (0,48)	0,26 (0,44)
Opleiding onbekend (1/0)	0,05 (0,22)	0,03 (0,18)	0,04 (0,20)
Duur bijstand (in maanden)	54,26 (56,54)	56,62 (58,41)	49,72 (53,03)
Looninkomen in 9-3 maanden voor start (in Euro)	1.127,49 (1.938,58)	716,87* (1.354,27)	923,65 (1.619,23)
Aantal observaties	122	122	122
Noot: verschil met controlegroep * p < 0,10%; ** p < 0,05%.			

Tabel B3.1d Beschrijvende statistieken Tilburg, karakteristieken 3 maanden voor start

	Controle	Intensivering	Ontheffing m/ vrijlating ink-grens	Intensivering m/ vrijlating ink-grens
Leeftijd	44,51 (9,77)	47,06** (10,14)	45,95 (10,36)	43,77 (10,72)
Man (1/0)	0,42 (0,49)	0,39 (0,49)	0,35 (0,48)	0,37 (0,48)
Alleenstaand (1/0)	0,54 (0,50)	0,52 (0,50)	0,49 (0,50)	0,46 (0,50)
Alleenstaande ouder (1/0)	0,30 (0,46)	0,31 (0,47)	0,33 (0,47)	0,37 (0,48)
Alleenstaanden en gehuwden (1/0)	0,07 (0,25)	0,08 (0,27)	0,06 (0,24)	0,06 (0,24)
Alleenstaanden en gehuwden met kinderen (1/2)	0,07 (0,25)	0,08 (0,27)	0,09 (0,29)	0,10 (0,30)
Overig/onbekend (1/0)	0,02 (0,14)	0,01 (0,07)	0,02 (0,14)	0,01 (0,07)
Autochtoon (1/0)	0,64 (0,48)	0,60 (0,49)	0,60 (0,49)	0,59 (0,49)
Westerse immigrant (1/0)	0,12 (0,32)	0,13 (0,33)	0,10 (0,31)	0,10 (0,30)
Niet-westerse immigrant (1/0)	0,24 (0,43)	0,27 (0,45)	0,30 (0,46)	0,31* (0,47)
Lagere opleiding (1/0)	0,43 (0,50)	0,43 (0,50)	0,50 (0,50)	0,49 (0,50)
Middelbare opleiding (1/0)	0,37 (0,48)	0,36 (0,48)	0,30 (0,46)	0,32 (0,47)
Hogere opleiding (1/0)	0,15 (0,36)	0,16 (0,36)	0,16 (0,36)	0,16 (0,37)
Opleiding onbekend (1/0)	0,06 (0,24)	0,05 (0,22)	0,05 (0,22)	0,03 (0,16)
Duur bijstand (in maanden)	56,85 (57,11)	73,04** (74,09)	63,67 (69,73)	60,39 (57,19)
Looninkomen in 9-3 maanden voor start (in Euro)	573,90 (1.486,99)	255,36** (872,35)	443,82 (1.252,89)	312,44** (957,05)
Aantal observaties	202	191	193	194
Noot: verschil met controlegroep * p < 0,10%; ** p < 0,05%.				

Tabel B3.1e Beschrijvende statistieken Utrecht, karakteristieken 3 maanden voor start

	Controle	Referentie	Ontheffing	Intensivering	Vrijlating
Leeftijd	46,77	45,63	46,95	46,66	46,21
	(10,13)	(10,15)	(10,14)	(10,30)	(9,97)
Man (1/0)	0,51	0,37	0,53	0,45	0,50
	(0,50)	(0,48)	(0,50)	(0,50)	(0,50)
Alleenstaand (1/0)	0,51	0,36	0,51	0,49	0,51
	(0,50)	(0,48)	(0,50)	(0,50)	(0,50)
Alleenstaande ouder (1/0)	0,26	0,25	0,22	0,26	0,21
	(0,44)	(0,44)	(0,42)	(0,44)	(0,41)
Alleenstaanden en gehuwden (1/0)	0,05	0,09	0,08	0,06	0,05
	(0,23)	(0,28)	(0,27)	(0,24)	(0,23)
Alleenstaanden en gehuwden met kinderen (1/2)	0,15	0,27	0,17	0,18	0,20
	(0,36)	(0,45)	(0,38)	(0,38)	(0,40)
Overig/onbekend (1/0)	0,03	0,03	0,02	0,02	0,02
	(0,16)	(0,17)	(0,13)	(0,14)	(0,13)
Autochtoon (1/0)	0,35	0,37	0,40	0,37	0,37
	(0,48)	(0,49)	(0,49)	(0,48)	(0,49)
Westerse immigrant (1/0)	0,11	0,08	0,13	0,09	0,13
	(0,31)	(0,27)	(0,34)	(0,28)	(0,34)
Niet-westerse immigrant (1/0)	0,54	0,55	0,47	0,54	0,50
	(0,50)	(0,50)	(0,50)	(0,50)	(0,50)
Lagere opleiding (1/0)	0,52	0,64	0,43*	0,47	0,42*
	(0,50)	(0,48)	(0,50)	(0,50)	(0,50)
Middelbare opleiding (1/0)	0,23	0,24	0,26	0,27	0,26
	(0,42)	(0,43)	(0,44)	(0,44)	(0,44)
Hogere opleiding (1/0)	0,19	0,10	0,26*	0,22	0,26
	(0,39)	(0,30)	(0,44)	(0,41)	(0,44)
Opleiding onbekend (1/0)	0,06	0,02	0,05	0,04	0,06
	(0,24)	(0,14)	(0,21)	(0,20)	(0,24)
Duur bijstand (in maanden)	76,10	85,02	69,43	68,84	74,60
	(74,27)	(83,30)	(73,47)	(64,55)	(74,63)
Looninkomen in 9-3 maanden voor start (in Euro)	370,70	453,41	517,46	366,62	458,63
	(1.159,63)	(1.469,76)	(1.482,53)	(1.165,06)	(1.332,25)
Aantal observaties	188	198	189	188	187
Noot: verschil met controlegroep * p < 0,10%; ** p < 0,05%.					

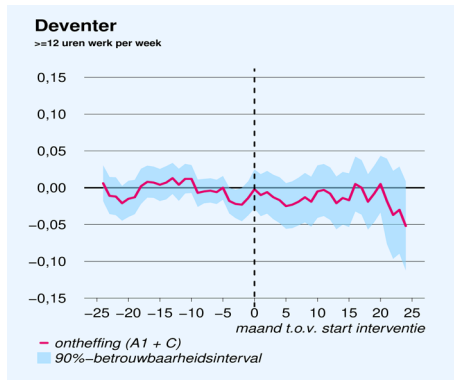
Tabel B3.1f Beschrijvende statistieken Wageningen, karakteristieken 3 maanden voor start

	Controle	Ontheffing	Intensivering	Vrijlating Ink.grens
Leeftijd	42,65	37,83**	41,27	40,65
	(12,72)	(12,38)	(12,80)	(11,26)
Man (1/0)	0,47	0,46	0,52	0,52
	(0,50)	(0,50)	(0,50)	(0,50)
Alleenstaand (1/0)	0,48	0,46	0,44	0,53
	(0,50)	(0,50)	(0,50)	(0,50)
Alleenstaande ouder (1/0)	0,20	0,22	0,19	0,24
	(0,41)	(0,41)	(0,40)	(0,43)
Alleenstaanden en gehuwden (1/0)	0,04	0,08	0,09	0,04
	(0,20)	(0,28)	(0,29)	(0,21)
Alleenstaanden en gehuwden met kinderen (1/2)	0,25	0,23	0,26	0,19
	(0,43)	(0,42)	(0,44)	(0,39)
Overig/onbekend (1/0)	0,02	0,01	0,02	0,00
	(0,15)	(0,10)	(0,14)	(0,00)
Autochtoon (1/0)	0,35	0,45	0,37	0,46
	(0,48)	(0,50)	(0,48)	(0,50)
Westerse immigrant (1/0)	0,11	0,05	0,11	0,13
	(0,31)	(0,21)	(0,32)	(0,34)
Niet-westerse immigrant (1/0)	0,54	0,50	0,52	0,41*
	(0,50)	(0,50)	(0,50)	(0,49)
Lagere opleiding (1/0)	0,41	0,42	0,35	0,33
	(0,49)	(0,50)	(0,48)	(0,47)
Middelbare opleiding (1/0)	0,23	0,19	0,26	0,25
	(0,42)	(0,39)	(0,44)	(0,43)
Hogere opleiding (1/0)	0,23	0,25	0,19	0,27
	(0,42)	(0,43)	(0,40)	(0,44)
Opleiding onbekend (1/0)	0,14	0,14	0,20	0,16
	(0,35)	(0,35)	(0,41)	(0,37)
Duur bijstand (in maanden)	40,78	39,71	45,47	35,85
	(50,77)	(62,12)	(67,49)	(48,84)
Looninkomen in 9-3 maanden voor start (in Euro)	650,54	553,34	576,04	422,14
	(1.900,43)	(1.448,24)	(1.803,97)	(1.103,70)
Aantal observaties	93	106	98	113
Noot: verschil met controlegroep * p < 0,10%; ** p < 0,05%.				

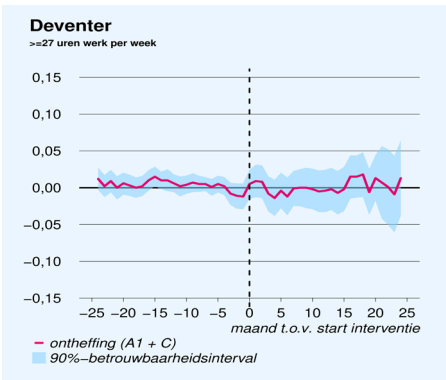
Bijlage IV – Placebo: ITT-effecten over de tijd

Figuur B4.1a-1f Deventer ITT-effecten over de tijd meer dan 12 uur werk per week (links) en 27 uur werk per week (rechts)

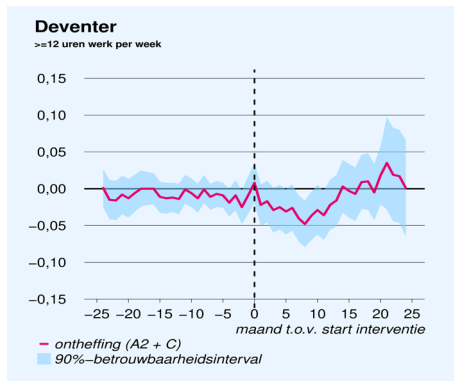
1a) ontheffing



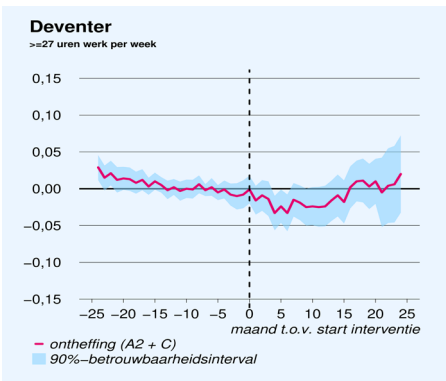
1b) ontheffing



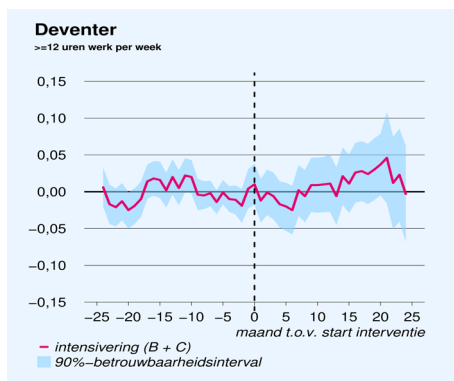
1c) ontheffing met app



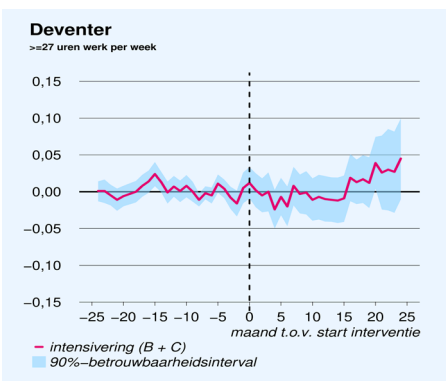
1d) ontheffing met app



1e) intensivering



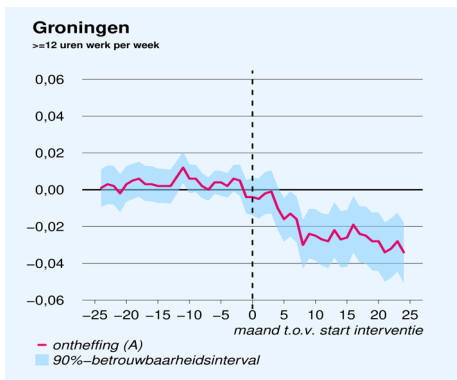
1f) intensivering



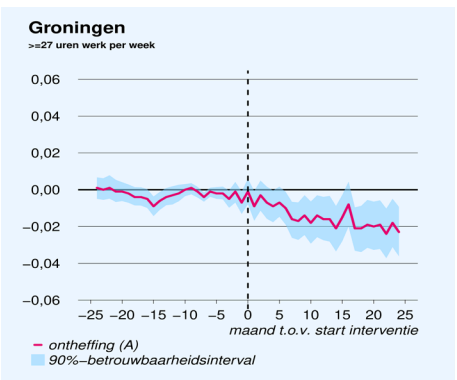
Noot: A1 = Ontheffing, A2 = Ontheffing met app, B = Intensivering, C = Vrijlating inkomensgrens.

Figuur B4.2a-2h Groningen ITT-effecten over de tijd meer dan 12 uur werk per week (links) en 27 uur werk per week (rechts)

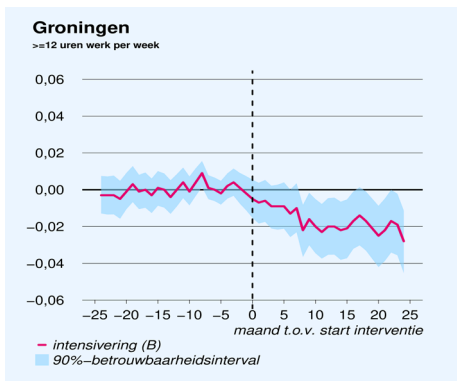
2a) ontheffing



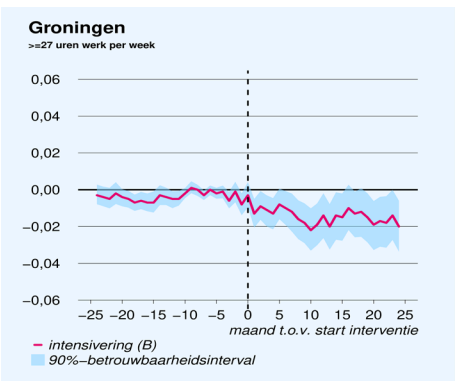
2b) ontheffing



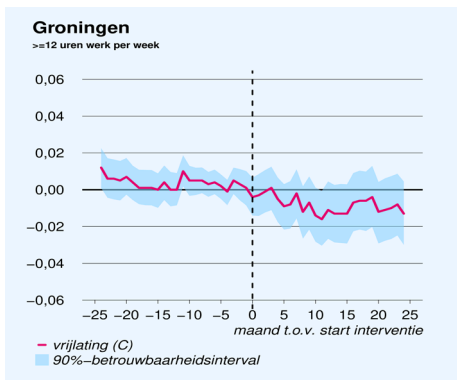
2c) intensivering



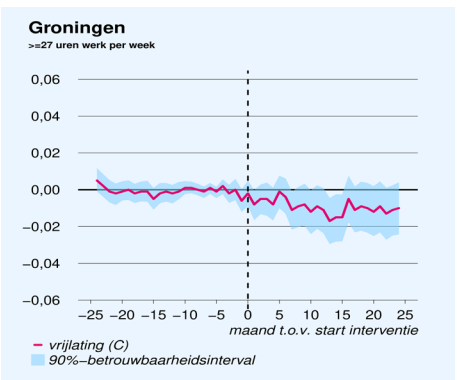
2d) intensivering



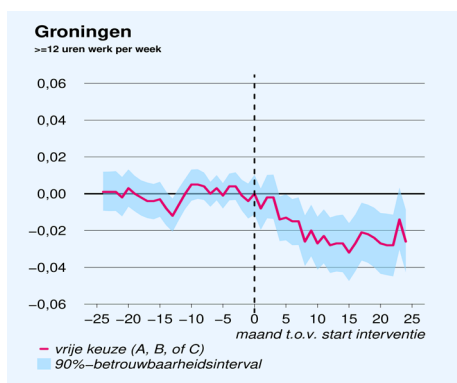
2e) vrijlating



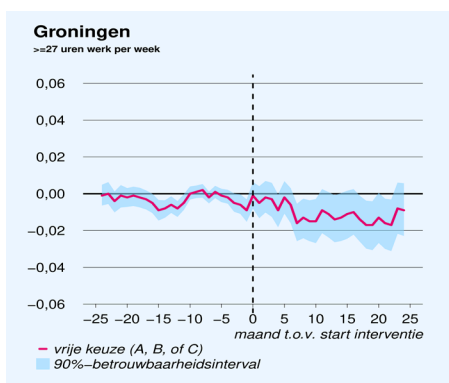
2f) vrijlating



2g) vrije keuze



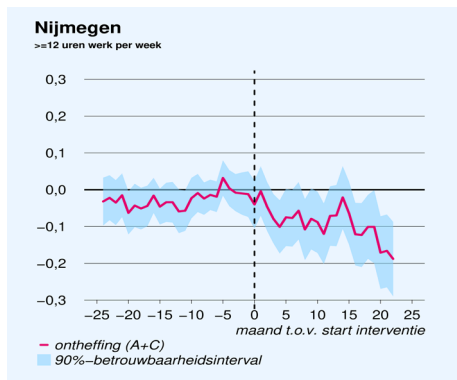
2h) vrije keuze



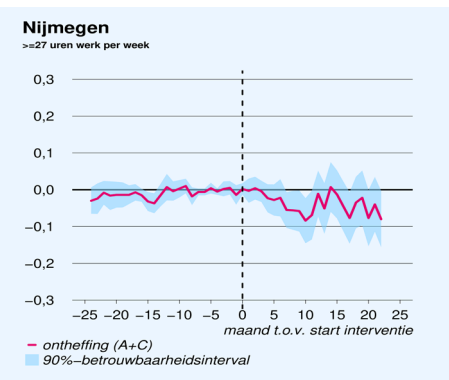
Noot: A = Ontheffing, B = Intensivering, C = Vrijlating inkomensgrens.

Figuur B4.3a-3d Nijmegen ITT-effecten over de tijd meer dan 12 uur werk per week (links) en 27 uur werk per week (rechts)

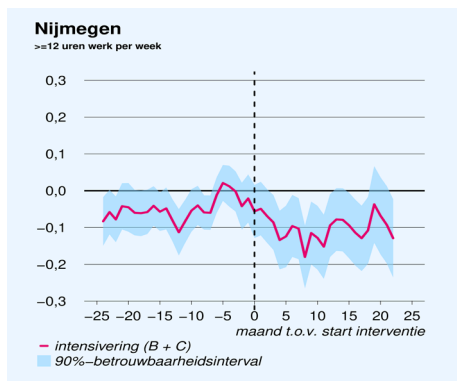
3a) ontheffing



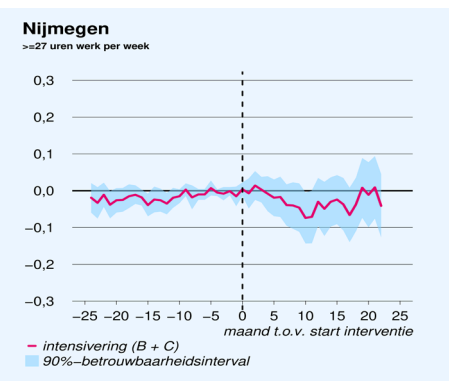
3b) ontheffing



3c) intensivering



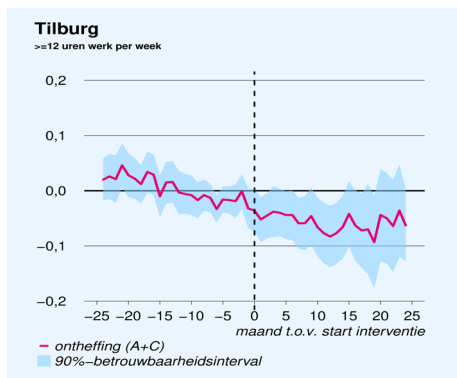
3d) intensivering



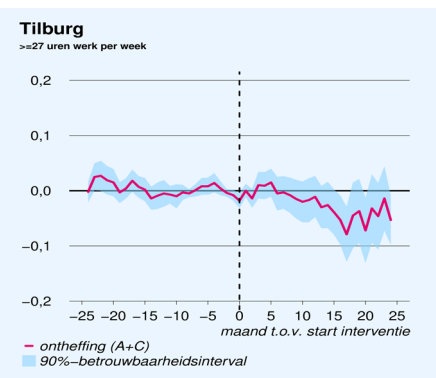
Noot: A = Ontheffing, B = Intensivering, C = Vrijlating inkomensgrens.

Figuur B4.4a-4f Tilburg ITT-effecten over de tijd meer dan 12 uur werk per week (links) en 27 uur werk per week (rechts)

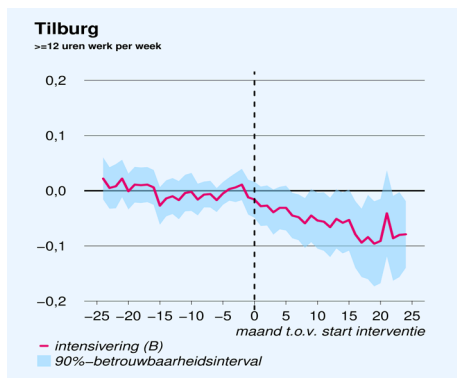
4a) ontheffing



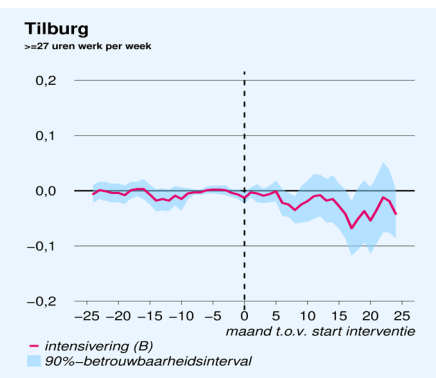
4b) ontheffing



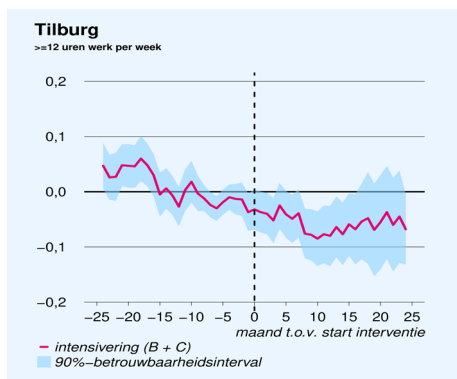
4c) intensivering



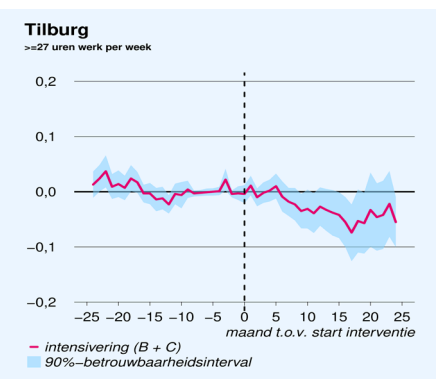
4d) intensivering



4e) intensivering met vrijlating inkomensgrens



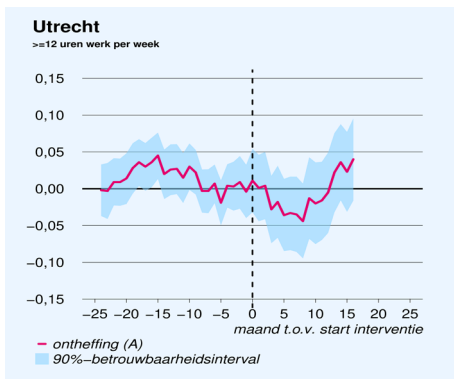
4f) intensivering met vrijlating inkomensgrens



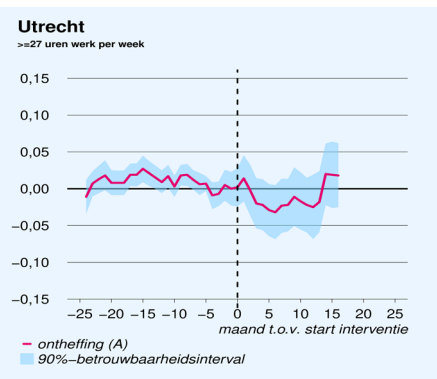
Noot: A = Ontheffing, B = Intensivering, C = Vrijlating inkomensgrens.

Figuur B4.5a-5f Utrecht ITT-effecten over de tijd meer dan 12 uur werk per week (links) en 27 uur werk per week (rechts)

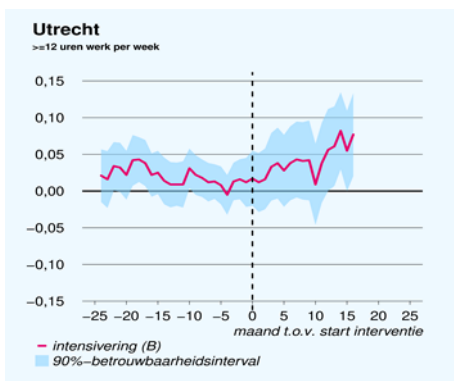
5a) ontheffing



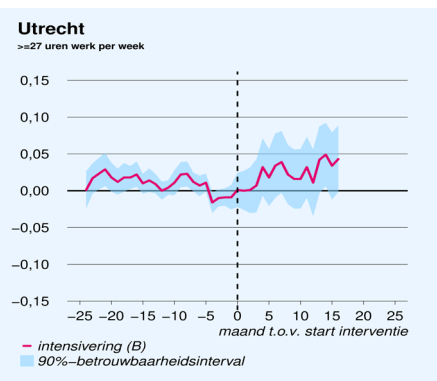
5b) ontheffing



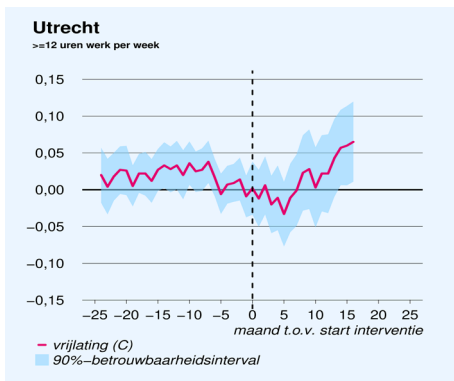
5c) intensivering



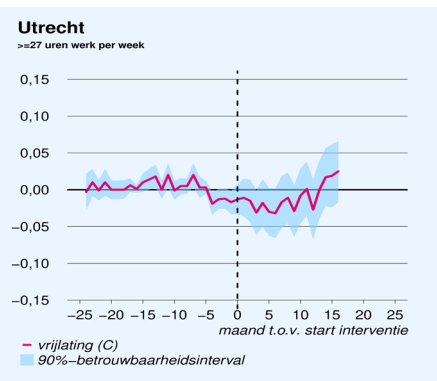
5d) intensivering



5e) vrijlating



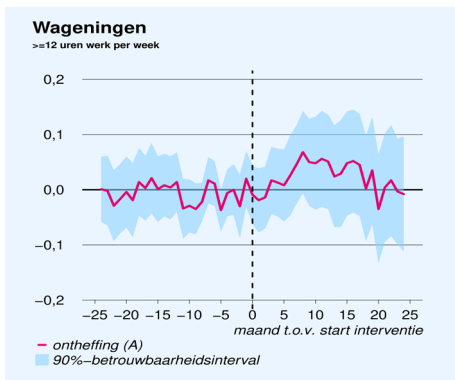
5f) vrijlating



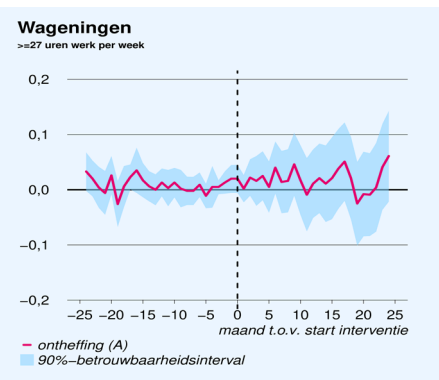
Noot: A = Ontheffing, B = Intensivering, C = Vrijlating inkomensgrens.

Figuur B4.6a-6f Wageningen ITT-effecten over de tijd meer dan 12 uur werk per week (links) en 27 uur werk per week (rechts)

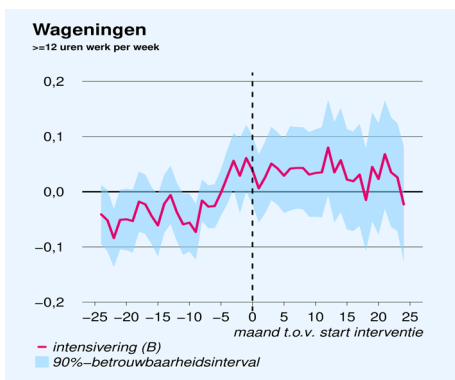
6a) ontheffing



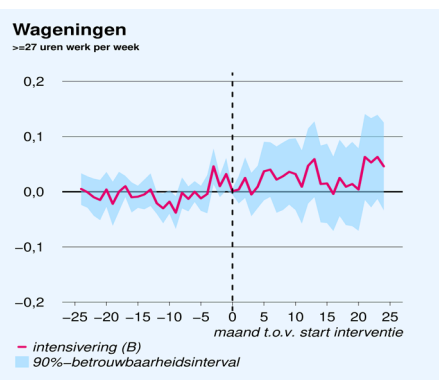
6b) ontheffing



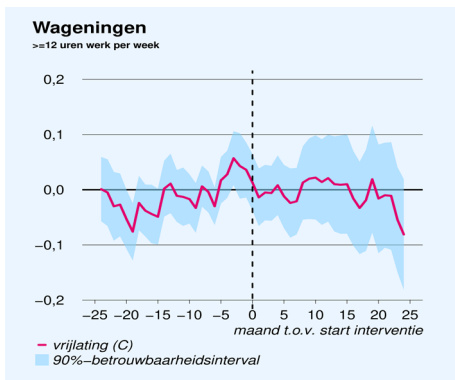
6c) intensivering



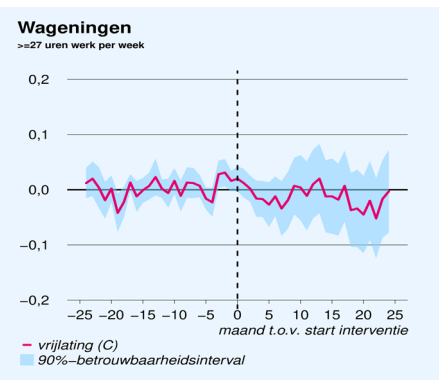
6d) intensivering



6e) vrijlating



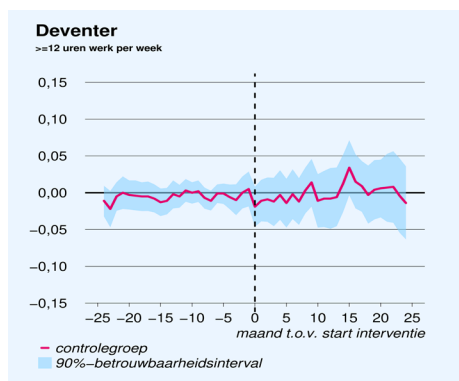
6f) vrijlating



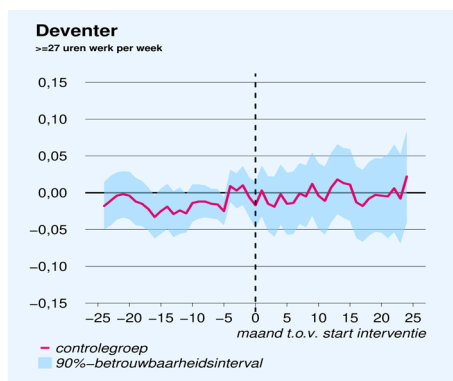
Noot: A = Ontheffing, B = Intensivering, C = Vrijlating inkomensgrens.

Figuur B4.7a en 7b Deventer ITT-effecten over de tijd meer dan 12 uur werk per week (links) en 27 uur werk per week (rechts)

7a) Controle t.o.v. referentiegroep

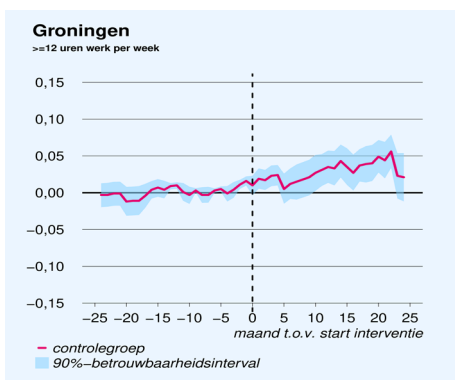


7b) Controle t.o.v. referentiegroep

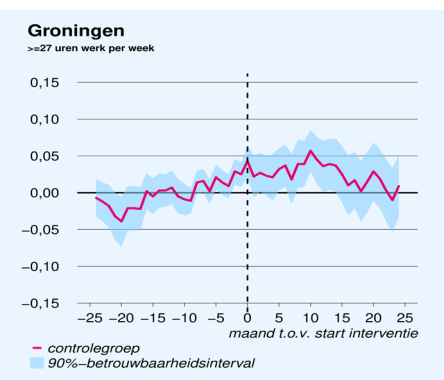


Figuur B4.8a en 8b Groningen ITT-effecten over de tijd meer dan 12 uur werk per week (links) en 27 uur werk per week (rechts)

8a) Controle t.o.v. referentiegroep



8b) Controle t.o.v. referentiegroep



Bijlage V – ITT-effecten

Tabel B5.1a Intention-to-Treat effecten, >27 uur werk per week, meet moment 24 maanden na start - Utrecht, Wageningen, Nijmegen en Tilburg

	Utrecht	Wageningen	Nijmegen	Tilburg
Ontheffing	0,018	0,061		
	(0,026)	(0,050)		
Intensivering	0,043	0,046		-0,042
	(0,027)	(0,048)		(0,027)
Vrijstelling Inkomensgrens	0,025	-0,003		
	(0,025)	(0,045)		
Keuze interventie				
Ontheffing (met vrijstelling inkomensgrens)			-0,080*	-0,053**
			(0,046)	(0,027)
Intensivering (met vrijstelling inkomensgrens)			-0,041	-0,055**
			(0,052)	(0,028)
Ontheffing (met app en vrijstelling inkomensgrens)				
Aantal observaties	752	375	289	683
Gemiddelde controle groep	0,059	0,103	0,165	0,102
	(0,235)	(0,306)	(0,373)	(0,304)

Noot: * $p < 0,10$; ** $p < 0,05$. Deze tabel laat coëfficiënten en standaardfouten () zien van schattingen uit een lineair kansmodel. De specificatie bevat de volgende controlevariabelen: geslacht (dummy), hoogste opleidingsniveau (4 categorieën: laag, middelbaar, hoog, onbekend), migratieachtergrond (4 categorieën: geen, westers, niet-westers, onbekend), huishoudtype (5 categorieën: alleenstaande, alleenstaande ouder, gehuwden en samenwonenden, gehuwden en samenwonend met kinderen, anders), leeftijd in jaren, duur van de lopende uitkering in maanden, cumulatief looninkomen 9 tot 3 maanden voor start interventie, en fixed effects voor het startmoment van experiment. Het laatst beschikbare meetmoment in Utrecht en Tilburg is 16 maanden na start, en in Nijmegen 22 maanden na start. De schattingen voor die gemeentes zijn voor die meetmomenten.

Tabel B5.1b Intention-to-Treat effecten, >12 uur werk per week, meet moment 24 maanden na start - Utrecht, Wageningen, Nijmegen en Tilburg

	Utrecht	Wageningen	Nijmegen	Tilburg
Ontheffing	0,040	-0,008		
	(0,034)	(0,063)		
Intensivering	0,077**	-0,023		-0,079**
	(0,034)	(0,063)		(0,037)
Vrijstelling Inkomensgrens	0,065**	-0,081		
	(0,033)	(0,061)		
Keuze interventie				
Ontheffing (met vrijstelling inkomensgrens)			-0,188**	-0,063
			(0,061)	(0,039)
Intensivering (met vrijstelling inkomensgrens)			-0,129**	-0,068*
			(0,064)	(0,039)
Ontheffing (met app en vrijstelling inkomensgrens)				
Aantal observaties	752	375	289	683
Gemiddelde controle groep	0,117	0,287	0,392	0,25
	(0,322)	(0,455)	(0,491)	(0,434)
<p>Noot: * $p < 0,10$; ** $p < 0,05$. Deze tabel laat coëfficiënten en standaardfouten () zien van schattingen uit een lineair kansmodel. De specificatie bevat de volgende controlevariabelen: geslacht (dummy), hoogste opleidingsniveau (4 categorieën: laag, middelbaar, hoog, onbekend), migratieachtergrond (4 categorieën: geen, westers, niet-westers, onbekend), huishoudtype (5 categorieën: alleenstaande, alleenstaande ouder, gehuwden en samenwonenden, gehuwden en samenwonend met kinderen, anders), leeftijd in jaren, duur van de lopende uitkering in maanden, cumulatief looninkomen 9 tot 3 maanden voor start interventie, en fixed effects voor het startmoment van experiment. Het laatst beschikbare meetmoment in Utrecht en Tilburg is 16 maanden na start, en in Nijmegen 22 maanden na start. De schattingen voor die gemeentes zijn voor die meetmomenten.</p>				

Tabel B5.2a Intention-to-Treat effecten, >27 uur werk per week, meet moment 24 maanden na start – Deventer en Groningen

	Deventer	Groningen
Ontheffing		-0,023**
		(0,008)
Intensivering		-0,020**
		(0,008)
Vrijstelling Inkomensgrens		-0,010
		(0,009)
Keuze interventie		-0,009
		(0,009)
Ontheffing (met vrijstelling inkomensgrens)	0,013	
	(0,031)	
Intensivering (met vrijstelling inkomensgrens)	0,045	
	(0,033)	
Ontheffing (met app en vrijstelling inkomensgrens)	0,020	
	(0,032)	
Aantal observaties	695	8.190
Gemiddelde controle groep	0,092	0,079
	(0,290)	(0,270)

Noot: * $p < 0,10$; ** $p < 0,05$. Deze tabel laat coëfficiënten en standaardfouten () zien van schattingen uit een lineair kansmodel. De specificatie bevat de volgende controlevariabelen:

geslacht (dummy), hoogste opleidingsniveau (4 categorieën: laag, middelbaar, hoog, onbekend), migratieachtergrond (4 categorieën: geen, westers, niet-westers, onbekend), huishoudtype (5 categorieën: alleenstaande, alleenstaande ouder, gehuwden en samenwonenden, gehuwden en samenwonend met kinderen, anders), leeftijd in jaren, duur van de lopende uitkering in maanden, cumulatief looninkomen 9 tot 3 maanden voor start interventie, en fixed effects voor het startmoment van experiment.

Tabel B5.2b Intention-to-Treat effecten, >12 uur werk per week, meet moment 24 maanden na start - Deventer en Groningen

	Deventer	Groningen
Ontheffing		-0,034**
		(0,010)
Intensivering		-0,028**
		(0,010)
Vrijstelling Inkomensgrens		-0,013
		(0,010)
Keuze interventie		-0,026**
		(0,010)
Ontheffing (met vrijstelling inkomensgrens)	-0,052	
	(0,037)	
Intensivering (met vrijstelling inkomensgrens)	-0,003	
	(0,040)	
Ontheffing (met app en vrijstelling inkomensgrens)	0,001	
	(0,040)	
Aantal observaties	695	8.190
Gemiddelde controle groep	0,209	0,139
	(0,408)	(0,347)

Noot: * p < 0,10; ** p < 0,05. Deze tabel laat coëfficiënten en standaardfouten () zien van schattingen uit een lineair kansmodel. De specificatie bevat de volgende controlevariabelen: geslacht (dummy), hoogste opleidingsniveau (4 categorieën: laag, middelbaar, hoog, onbekend), migratieachtergrond (4 categorieën: geen, westers, niet-westers, onbekend), huishoudtype (5 categorieën: alleenstaande, alleenstaande ouder, gehuwden en samenwonenden, gehuwden en samenwonend met kinderen, anders), leeftijd in jaren, duur van de lopende uitkering in maanden, cumulatief looninkomen 9 tot 3 maanden voor start interventie, en fixed effects voor het startmoment van experiment.

Tabel B5.3a Intention-to-Treat effecten t.o.v. referentiegroep, >27 uur werk per week, meetmoment 24 maanden na start - Deventer en Groningen

	Deventer	Groningen
Ontheffing (met vrijstelling inkomensgrens)	-0,001	
	(0,030)	
Intensivering (met vrijstelling inkomensgrens)	0,030	
	(0,032)	
Intensivering (met app en vrijstelling inkomensgrens)	0,008	
	(0,030)	
Ontheffing		-0,002
		(0,020)
Intensivering		0,001
		(0,020)
Vrijstelling Inkomensgrens		0,011
		(0,020)
Keuze interventie		0,012
		(0,020)
Aantal observaties	738	6.701
Gemiddelde controle groep	0,092	0,079
	(0,290)	(0,270)
<p>Noot: * p < 0,10; ** p < 0,05. Deze tabel laat coëfficiënten en standaardfouten () zien van schattingen uit een lineair kansmodel. De specificatie bevat de volgende controlevariabelen: geslacht (dummy), hoogste opleidingsniveau (4 categorieën: laag, middelbaar, hoog, onbekend), migratieachtergrond (4 categorieën: geen, westers, niet-westers, onbekend), huishoudtype (5 categorieën: alleenstaande, alleenstaande ouder, gehuwden en samenwonenden, gehuwden en samenwonend met kinderen, anders), leeftijd in jaren, duur van de lopende uitkering in maanden, cumulatief looninkomen 9 tot 3 maanden voor start interventie, en fixed effects voor het startmoment van experiment.</p>		

**Tabel B5.3b Intention-to-Treat effecten t.o.v. referentiegroep, >12 uur werk per week, meetmoment 24 maanden na start
- Deventer en Groningen**

	Deventer	Groningen
Ontheffing (met vrijstelling inkomensgrens)	-0,001	
	(0,035)	
Intensivering (met vrijstelling inkomensgrens)	0,020	
	(0,036)	
Intensivering (met app en vrijstelling inkomensgrens)	0,022	
	(0,037)	
Ontheffing		-0,027
		(0,025)
Intensivering		-0,020
		(0,025)
Vrijstelling Inkomensgrens		-0,005
		(0,025)
Keuze interventie		-0,018
		(0,025)
Aantal observaties	738	6.701
Gemiddelde controle groep	0,209	0,139
	(0,408)	(0,347)
<p>Noot: * p < 0,10; ** p < 0,05. Deze tabel laat coëfficiënten en standaardfouten () zien van schattingen uit een lineair kansmodel. De specificatie bevat de volgende controlevariabelen: geslacht (dummy), hoogste opleidingsniveau (4 categorieën: laag, middelbaar, hoog, onbekend), migratieachtergrond (4 categorieën: geen, westers, niet-westers, onbekend), huishoudtype (5 categorieën: alleenstaande, alleenstaande ouder, gehuwden en samenwonenden, gehuwden en samenwonend met kinderen, anders), leeftijd in jaren, duur van de lopende uitkering in maanden, cumulatief looninkomen 9 tot 3 maanden voor start interventie, en fixed effects voor het startmoment van experiment.</p>		

Bijlage VI - LATE-effecten

Tabel B6.1a Local Average Treatment Effecten, >27 uur werk per week, meetmoment 24 maanden na start - Utrecht, Wageningen, Nijmegen en Tilburg

	Utrecht	Wageningen	Nijmegen	Tilburg
Ontheffing	0,019	0,075		
	(0,027)	(0,060)		
Intensivering	0,046	0,055		-0,069
	(0,029)	(0,057)		(0,045)
Vrijstelling Inkomensgrens	0,027	-0,002		
	(0,027)	(0,050)		
Ontheffing (met vrijstelling inkomensgrens)			-0,086*	-0,069**
			(0,048)	(0,035)
Intensivering (met vrijstelling inkomensgrens)			-0,049	-0,075**
			(0,061)	(0,037)
Aantal observaties	752	375	289	683
Gemiddelde controle groep	0,059	0,103	0,165	0,102
	(0,235)	(0,306)	(0,373)	(0,304)
First Stage F-stat interventie 1	1943	147	562	82
First Stage F-stat interventie 2	967	156	241	184
First Stage F-stat interventie 3	714	277		156

Noot: * $p < 0,10$; ** $p < 0,05$. Deze tabel laat coëfficiënten en standaardfouten () zien van schattingen uit een 2SLS lineair kansmodel. De specificatie bevat de volgende controlevariabelen: geslacht (dummy), hoogste opleidingsniveau (4 categorieën: laag, middelbaar, hoog, onbekend), migratieachtergrond (4 categorieën: geen, westers, niet-westers, onbekend), huishoudtype (5 categorieën: alleenstaande, alleenstaande ouder, gehuwden en samenwonenden, gehuwden en samenwonend met kinderen, anders), leeftijd in jaren, duur van de lopende uitkering in maanden, cumulatief looninkomen 9 tot 3 maanden voor start interventie, en fixed effects voor het startmoment van experiment. Het laatst beschikbare meetmoment in Utrecht en Tilburg is 16 maanden na start, en in Nijmegen 22 maanden na start. De schattingen voor die gemeentes zijn voor die meetmomenten. Iedereen die gestart is met een interventie wordt beschouwd als deelnemer.

Tabel B6.1b Local Average Treatment Effecten, >12 uur werk per week, meetmoment 24 maanden na start - Utrecht, Wageningen, Nijmegen en Tilburg

	Utrecht	Wageningen	Nijmegen	Tilburg
Ontheffing	0,041	-0,009		
	(0,035)	(0,076)		
Intensivering	0,082**	-0,027		-0,132**
	(0,036)	(0,074)		(0,061)
Vrijstelling Inkomensgrens	0,071**	-0,091		
	(0,035)	(0,067)		
Ontheffing (met vrijstelling inkomensgrens)			-0,204**	-0,082
			(0,064)	(0,051)
Intensivering (met vrijstelling inkomensgrens)			-0,156**	-0,093*
			(0,075)	(0,052)
Aantal observaties	752	375	289	683
Gemiddelde controle groep	0,117	0,287	0,392	0,25
	(0,322)	(0,455)	(0,491)	(0,434)
First Stage F-stat interventie 1	1943	147	562	82
First Stage F-stat interventie 2	967	156	241	184
First Stage F-stat interventie 3	714	277		156

Noot: * $p < 0,10$; ** $p < 0,05$. Deze tabel laat coëfficiënten en standaardfouten () zien van schattingen uit een 2SLS lineair kansmodel. De specificatie bevat de volgende controlevariabelen: geslacht (dummy), hoogste opleidingsniveau (4 categorieën: laag, middelbaar, hoog, onbekend), migratieachtergrond (4 categorieën: geen, westers, niet-westers, onbekend), huishoudtype (5 categorieën: alleenstaande, alleenstaande ouder, gehuwden en samenwonenden, gehuwden en samenwonend met kinderen, anders), leeftijd in jaren, duur van de lopende uitkering in maanden, cumulatief looninkomen 9 tot 3 maanden voor start interventie, en fixed effects voor het startmoment van experiment. Het laatst beschikbare meetmoment in Utrecht en Tilburg is 16 maanden na start, en in Nijmegen 22 maanden na start. De schattingen voor die gemeentes zijn voor die meetmomenten. Iedereen die gestart is met een interventie wordt beschouwd als deelnemer.

Tabel B6.2a Local Average Treatment Effecten, >27 uur werk per week, meetmoment 24 maanden na start – Deventer en Groningen

	Deventer	Groningen
Ontheffing		-0,197**
		(0,072)
Intensivering		-0,221**
		(0,095)
Vrijstelling Inkomensgrens		-0,106
		(0,092)
Keuze interventie		-0,074
		(0,076)
Ontheffing (met vrijstelling inkomensgrens)	0,04	
	(0,095)	
Intensivering (met vrijstelling inkomensgrens)	0,201	
	(0,148)	
Intensivering (met app en vrijstelling inkomensgrens)	0,099	
	(0,161)	
Aantal observaties	695	8.190
Gemiddelde controle groep	0,092	0,079
	(0,290)	(0,270)
First Stage F-stat interventie 1	27	52
First Stage F-stat interventie 2	19	40
First Stage F-stat interventie 3	13	43

Noot: * $p < 0,10$; ** $p < 0,05$. Deze tabel laat coëfficiënten en standaardfouten () zien van schattingen uit een 2SLS lineair kansmodel. De specificatie bevat de volgende controlevariabelen: geslacht (dummy), hoogste opleidingsniveau (4 categorieën: laag, middelbaar, hoog, onbekend), migratieachtergrond (4 categorieën: geen, westers, niet-westers, onbekend), huishoudtype (5 categorieën: alleenstaande, alleenstaande ouder, gehuwden en samenwonenden, gehuwden en samenwonend met kinderen, anders), leeftijd in jaren, duur van de lopende uitkering in maanden, cumulatief looninkomen 9 tot 3 maanden voor start interventie, en fixed effects voor het startmoment van experiment. Iedereen die gestart is met een interventie wordt beschouwd als deelnemer.

Tabel B6.2b Local Average Treatment Effecten, >12 uur werk per week, meetmoment 24 maanden na start – Deventer en Groningen

	Deventer	Groningen
Ontheffing		-0,297**
		(0,088)
Intensivering		-0,315**
		(0,118)
Vrijstelling Inkomensgrens		-0,132
		(0,111)
Keuze interventie		-0,223**
		(0,091)
Ontheffing (met vrijstelling inkomensgrens)	-0,160	
	(0,114)	
Intensivering (met vrijstelling inkomensgrens)	-0,010	
	(0,176)	
Ontheffing (met app en vrijstelling inkomensgrens)	0,007	
	(0,201)	
Aantal observaties	695	8.190
Gemiddelde controle groep	0,209	0,139
	(0,408)	(0,347)
First Stage F-stat interventie 1	27	52
First Stage F-stat interventie 2	19	40
First Stage F-stat interventie 3	13	43

Noot: * $p < 0,10$; ** $p < 0,05$. Deze tabel laat coëfficiënten en standaardfouten () zien van schattingen uit een 2SLS lineair kansmodel. De specificatie bevat de volgende controlevariabelen: geslacht (dummy), hoogste opleidingsniveau (4 categorieën: laag, middelbaar, hoog, onbekend), migratieachtergrond (4 categorieën: geen, westers, niet-westers, onbekend), huishoudtype (5 categorieën: alleenstaande, alleenstaande ouder, gehuwden en samenwonenden, gehuwden en samenwonend met kinderen, anders), leeftijd in jaren, duur van de lopende uitkering in maanden, cumulatief looninkomen 9 tot 3 maanden voor start interventie, en fixed effects voor het startmoment van experiment. Iedereen die gestart is met een interventie wordt beschouwd als deelnemer.

Bijlage VII – Gevoeligheidsanalyses (ITT)

Tabel B7.1a Gevoeligheidsanalyse - ITT-effecten, >32 uur werk per week, meetmoment 24 maanden na start - Utrecht, Wageningen, Nijmegen en Tilburg

	Utrecht	Wageningen	Nijmegen	Tilburg
Ontheffing	-0,016	0,053		
	(0,021)	(0,044)		
Intensivering	0,016	0,061		-0,022
	(0,023)	(0,044)		(0,022)
Vrijstelling Inkomensgrens	0,027	0,013		
	(0,023)	(0,040)		
Keuze interventie				
Ontheffing (met vrijstelling inkomensgrens)			-0,061	-0,030
			(0,041)	(0,022)
Intensivering (met vrijstelling inkomensgrens)			-0,048	-0,036
			(0,047)	(0,022)
Ontheffing (met app en vrijstelling inkomensgrens)				
Aantal observaties	752	375	289	683
Gemiddelde controle groep	0,048	0,069	0,134	0,068
	(0,214)	(0,255)	(0,342)	(0,253)

Noot: * $p < 0,10$; ** $p < 0,05$. Deze tabel laat coëfficiënten en standaardfouten () zien van schattingen uit een lineair kansmodel. De specificatie bevat de volgende controlevariabelen: geslacht (dummy), hoogste opleidingsniveau (4 categorieën: laag, middelbaar, hoog, onbekend), migratieachtergrond (4 categorieën: geen, westers, niet-westers, onbekend), huishoudtype (5 categorieën: alleenstaande, alleenstaande ouder, gehuwden en samenwonenden, gehuwden en samenwonend met kinderen, anders), leeftijd in jaren, duur van de lopende uitkering in maanden, cumulatief looninkomen 9 tot 3 maanden voor start interventie, en fixed effects voor het startmoment van experiment. Het laatst beschikbare meetmoment in Utrecht en Tilburg is 16 maanden na start, en in Nijmegen 22 maanden na start. De schattingen voor die gemeentes zijn voor die meetmomenten.

Tabel B7.1b Gevoeligheidsanalyse - ITT-effecten, >32 uur werk per week, meetmoment 24 maanden na start – Deventer en Groningen

	Deventer	Groningen
Ontheffing		-0,018**
		(0,007)
Intensivering		-0,017**
		(0,008)
Vrijstelling Inkomensgrens		-0,007
		(0,008)
Keuze interventie		-0,009
		(0,008)
Ontheffing (met vrijstelling inkomensgrens)	0,013	
	(0,026)	
Intensivering (met vrijstelling inkomensgrens)	0,032	
	(0,028)	
Ontheffing (met app en vrijstelling inkomensgrens)	0,005	
	(0,025)	
Aantal observaties	695	8.190
Gemiddelde controle groep	0,061	0,062
	(0,241)	(0,242)

Noot: * $p < 0,10$; ** $p < 0,05$. Deze tabel laat coëfficiënten en standaardfouten () zien van schattingen uit een lineair kansmodel. De specificatie bevat de volgende controlevariabelen: geslacht (dummy), hoogste opleidingsniveau (4 categorieën: laag, middelbaar, hoog, onbekend), migratieachtergrond (4 categorieën: geen, westers, niet-westers, onbekend), huishoudtype (5 categorieën: alleenstaande, alleenstaande ouder, gehuwden en samenwonenden, gehuwden en samenwonend met kinderen, anders), leeftijd in jaren, duur van de lopende uitkering in maanden, cumulatief looninkomen 9 tot 3 maanden voor start interventie, en fixed effects voor het startmoment van experiment.

Tabel B7.2a Gevoeligheidsanalyse - ITT-effecten, >8 uur werk per week, meetmoment 24 maanden na start - Utrecht, Wageningen, Nijmegen en Tilburg

	Utrecht	Wageningen	Nijmegen	Tilburg
Ontheffing	0,039	0,004		
	(0,034)	(0,064)		
Intensivering	0,077**	-0,027		-0,085**
	(0,035)	(0,063)		(0,037)
Vrijstelling Inkomensgrens	0,086**	-0,025		
	(0,034)	(0,064)		
Keuze interventie				
Ontheffing (met vrijstelling inkomensgrens)			-0,186**	-0,052
			(0,061)	(0,040)
Intensivering (met vrijstelling inkomensgrens)			-0,123*	-0,047
			(0,066)	(0,039)
Ontheffing (met app en vrijstelling inkomensgrens)				
Aantal observaties	752	375	289	683
Gemiddelde controle groep	0,122	0,299	0,443	0,256
	(0,329)	(0,460)	(0,499)	(0,437)

Noot: * $p < 0,10$; ** $p < 0,05$. Deze tabel laat coëfficiënten en standaardfouten () zien van schattingen uit een lineair kansmodel. De specificatie bevat de volgende controlevariabelen: geslacht (dummy), hoogste opleidingsniveau (4 categorieën: laag, middelbaar, hoog, onbekend), migratieachtergrond (4 categorieën: geen, westers, niet-westers, onbekend), huishoudtype (5 categorieën: alleenstaande, alleenstaande ouder, gehuwden en samenwonenden, gehuwden en samenwonend met kinderen, anders), leeftijd in jaren, duur van de lopende uitkering in maanden, cumulatief looninkomen 9 tot 3 maanden voor start interventie, en fixed effects voor het startmoment van experiment. Het laatst beschikbare meetmoment in Utrecht en Tilburg is 16 maanden na start, en in Nijmegen 22 maanden na start. De schattingen voor die gemeentes zijn voor die meetmomenten.

Tabel B7.2b Gevoeligheidsanalyse - ITT-effecten, >8 uur werk per week, meetmoment 24 maanden na start – Deventer en Groningen

	Deventer	Groningen
Ontheffing		-0,034**
		(0,010)
Intensivering		-0,022**
		(0,011)
Vrijstelling Inkomensgrens		-0,009
		(0,011)
Keuze interventie		-0,025**
		(0,011)
Ontheffing (met vrijstelling inkomensgrens)	-0,022	
	(0,037)	
Intensivering (met vrijstelling inkomensgrens)	0,005	
	(0,040)	
Ontheffing (met app en vrijstelling inkomensgrens)	0,027	
	(0,039)	
Aantal observaties	695	8.190
Gemiddelde controle groep	0,209	0,155
	(0,408)	(0,362)

Noot: * $p < 0,10$; ** $p < 0,05$. Deze tabel laat coëfficiënten en standaardfouten () zien van schattingen uit een lineair kansmodel. De specificatie bevat de volgende controlevariabelen: geslacht (dummy), hoogste opleidingsniveau (4 categorieën: laag, middelbaar, hoog, onbekend), migratieachtergrond (4 categorieën: geen, westers, niet-westers, onbekend), huishoudtype (5 categorieën: alleenstaande, alleenstaande ouder, gehuwden en samenwonenden, gehuwden en samenwonend met kinderen, anders), leeftijd in jaren, duur van de lopende uitkering in maanden, cumulatief looninkomen 9 tot 3 maanden voor start interventie, en fixed effects voor het startmoment van experiment.

Tabel B7.3a Gevoeligheidsanalyse - ITT-effecten, aantal gewerkte uren per maand, meetmoment 24 maanden na start - Utrecht, Wageningen, Nijmegen en Tilburg

	Utrecht	Wageningen	Nijmegen	Tilburg
Ontheffing	2,735	3,175		
	(4,237)	(7,695)		
Intensivering	6,734	5,241		-11,020**
	(4,259)	(7,716)		(4,203)
Vrijstelling Inkomensgrens	7,724*	-3,939		
	(4,285)	(7,215)		
Keuze interventie				
Ontheffing (met vrijstelling inkomensgrens)			-21,379**	-10,298**
			(7,555)	(4,394)
Intensivering (met vrijstelling inkomensgrens)			-14,832*	-9,411**
			(8,503)	(4,471)
Ontheffing (met app en vrijstelling inkomensgrens)				
Aantal observaties	752	375	289	683
Gemiddelde controle groep	15,571	31,986	48,265	29,147
	(41,892)	(50,010)	(61,263)	(51,370)
<p>Noot: * $p < 0,10$; ** $p < 0,05$. Deze tabel laat coëfficiënten en standaardfouten () zien van schattingen uit een OLS-model. De specificatie bevat de volgende controlevariabelen: geslacht (dummy), hoogste opleidingsniveau (4 categorieën: laag, middelbaar, hoog, onbekend), migratieachtergrond (4 categorieën: geen, westers, niet-westers, onbekend), huishoudtype (5 categorieën: alleenstaande, alleenstaande ouder, gehuwden en samenwonenden, gehuwden en samenwonend met kinderen, anders), leeftijd in jaren, duur van de lopende uitkering in maanden, cumulatief looninkomen 9 tot 3 maanden voor start interventie, en fixed effects voor het startmoment van experiment. Het laatst beschikbare meetmoment in Utrecht en Tilburg is 16 maanden na start, en in Nijmegen 22 maanden na start. De schattingen voor die gemeentes zijn voor die meetmomenten.</p>				

Tabel B7.3b Gevoeligheidsanalyse - ITT-effecten, aantal gewerkte uren per maand, meetmoment 24 maanden na start – Deventer en Groningen

	Deventer	Groningen
Ontheffing		-4,358**
		(1,319)
Intensivering		-3,591**
		(1,358)
Vrijstelling Inkomensgrens		-1,801
		(1,376)
Keuze interventie		-2,985**
		(1,380)
Ontheffing (met vrijstelling inkomensgrens)	-3,040	
	(4,755)	
Intensivering (met vrijstelling inkomensgrens)	3,327	
	(5,355)	
Ontheffing (met app en vrijstelling inkomensgrens)	0,771	
	(4,948)	
Aantal observaties	695	8.190
Gemiddelde controle groep	25,352	18,407
	(50,439)	(45,773)

Noot: * $p < 0,10$; ** $p < 0,05$. Deze tabel laat coëfficiënten en standaardfouten () zien van schattingen uit een OLS-model. De specificatie bevat de volgende controlevariabelen: geslacht (dummy), hoogste opleidingsniveau (4 categorieën: laag, middelbaar, hoog, onbekend), migratieachtergrond (4 categorieën: geen, westers, niet-westers, onbekend), huishoudtype (5 categorieën: alleenstaande, alleenstaande ouder, gehuwden en samenwonenden, gehuwden en samenwonend met kinderen, anders), leeftijd in jaren, duur van de lopende uitkering in maanden, cumulatief looninkomen 9 tot 3 maanden voor start interventie, en fixed effects voor het startmoment van experiment.

Tabel B7.4a Gevoeligheidsanalyse - ITT-effecten, >70% WML, meetmoment 24 maanden na start - Utrecht, Wageningen, Nijmegen en Tilburg

	Utrecht	Wageningen	Nijmegen	Tilburg
Ontheffing	0,046	0,008		
	(0,030)	(0,054)		
Intensivering	0,047	0,013		-0,083**
	(0,029)	(0,053)		(0,031)
Vrijstelling Inkomensgrens	0,031	-0,045		
	(0,027)	(0,051)		
Keuze interventie				
Ontheffing (met vrijstelling inkomensgrens)			-0,126**	-0,075**
			(0,055)	(0,033)
Intensivering (met vrijstelling inkomensgrens)			-0,093	-0,058*
			(0,058)	(0,035)
Ontheffing (met app en vrijstelling inkomensgrens)				
Aantal observaties	752	375	289	683
Gemiddelde controle groep	0,069	0,161	0,237	0,159
	(0,254)	(0,370)	(0,428)	(0,367)

Noot: * $p < 0,10$; ** $p < 0,05$. Deze tabel laat coëfficiënten en standaardfouten () zien van schattingen uit een lineair kansmodel. De specificatie bevat de volgende controlevariabelen: geslacht (dummy), hoogste opleidingsniveau (4 categorieën: laag, middelbaar, hoog, onbekend), migratieachtergrond (4 categorieën: geen, westers, niet-westers, onbekend), huishoudtype (5 categorieën: alleenstaande, alleenstaande ouder, gehuwden en samenwonenden, gehuwden en samenwonend met kinderen, anders), leeftijd in jaren, duur van de lopende uitkering in maanden, cumulatief looninkomen 9 tot 3 maanden voor start interventie, en fixed effects voor het startmoment van experiment. Het laatst beschikbare meetmoment in Utrecht en Tilburg is 16 maanden na start, en in Nijmegen 22 maanden na start. De schattingen voor die gemeentes zijn voor die meetmomenten.

Tabel B7.4b Gevoeligheidsanalyse - ITT-effecten, >70% WML, meetmoment 24 maanden na start – Deventer en Groningen

	Deventer	Groningen
Ontheffing		-0,025**
		(0,009)
Intensivering		-0,020**
		(0,009)
Vrijstelling Inkomensgrens		-0,015
		(0,009)
Keuze interventie		-0,011
		(0,009)
Ontheffing (met vrijstelling inkomensgrens)	-0,030	
	(0,035)	
Intensivering (met vrijstelling inkomensgrens)	0,020	
	(0,038)	
Ontheffing (met app en vrijstelling inkomensgrens)	-0,009	
	(0,036)	
Aantal observaties	695	8.190
Gemiddelde controle groep	0,141	0,096
	(0,349)	(0,295)
<p>Noot: * $p < 0,10$; ** $p < 0,05$. Deze tabel laat coëfficiënten en standaardfouten () zien van schattingen uit een lineair kansmodel. De specificatie bevat de volgende controlevariabelen: geslacht (dummy), hoogste opleidingsniveau (4 categorieën: laag, middelbaar, hoog, onbekend), migratieachtergrond (4 categorieën: geen, westers, niet-westers, onbekend), huishoudtype (5 categorieën: alleenstaande, alleenstaande ouder, gehuwden en samenwonenden, gehuwden en samenwonend met kinderen, anders), leeftijd in jaren, duur van de lopende uitkering in maanden, cumulatief looninkomen 9 tot 3 maanden voor start interventie, en fixed effects voor het startmoment van experiment.</p>		

Tabel B7.5a Gevoeligheidsanalyse - ITT-effecten, >50% WML, meetmoment 24 maanden na start - Utrecht, Wageningen, Nijmegen en Tilburg

	Utrecht	Wageningen	Nijmegen	Tilburg
Ontheffing	0,040	0,016		
	(0,033)	(0,061)		
Intensivering	0,061*	0,029		-0,079**
	(0,033)	(0,062)		(0,033)
Vrijstelling Inkomensgrens	0,044	-0,053		
	(0,031)	(0,058)		
Keuze interventie				
Ontheffing (met vrijstelling inkomensgrens)			-0,169**	-0,081**
			(0,060)	(0,035)
Intensivering (met vrijstelling inkomensgrens)			-0,143**	-0,061*
			(0,062)	(0,037)
Ontheffing (met app en vrijstelling inkomensgrens)				
Aantal observaties	752	375	289	683
Gemiddelde controle groep	0,101	0,218	0,33	0,205
	(0,302)	(0,416)	(0,473)	(0,405)
<p>Noot: * $p < 0,10$; ** $p < 0,05$. Deze tabel laat coëfficiënten en standaardfouten () zien van schattingen uit een lineair kansmodel. De specificatie bevat de volgende controlevariabelen: geslacht (dummy), hoogste opleidingsniveau (4 categorieën: laag, middelbaar, hoog, onbekend), migratieachtergrond (4 categorieën: geen, westers, niet-westers, onbekend), huishoudtype (5 categorieën: alleenstaande, alleenstaande ouder, gehuwden en samenwonenden, gehuwden en samenwonend met kinderen, anders), leeftijd in jaren, duur van de lopende uitkering in maanden, cumulatief looninkomen 9 tot 3 maanden voor start interventie, dummies voor de kans op meer dan 27 uur werk 24, 16, en 8 maanden voor de start van het experiment, en fixed effects voor het startmoment van experiment. Het laatst beschikbare meetmoment in Utrecht en Tilburg is 16 maanden na start, en in Nijmegen 22 maanden na start. De schattingen voor die gemeentes zijn voor die meetmomenten.</p>				

Tabel B7.5b Gevoeligheidsanalyse - ITT-effecten, >50% WML, meetmoment 24 maanden na start – Deventer en Groningen

	Deventer	Groningen
Ontheffing		-0,033**
		(0,010)
Intensivering		-0,027**
		(0,010)
Vrijstelling Inkomensgrens		-0,015
		(0,010)
Keuze interventie		-0,024**
		(0,010)
Ontheffing (met vrijstelling inkomensgrens)	-0,030	
	(0,036)	
Intensivering (met vrijstelling inkomensgrens)	0,007	
	(0,039)	
Ontheffing (met app en vrijstelling inkomensgrens)	0,004	
	(0,038)	
Aantal observaties	695	8.190
Gemiddelde controle groep	0,178	0,122
	(0,384)	(0,328)

Noot: * $p < 0,10$; ** $p < 0,05$. Deze tabel laat coëfficiënten en standaardfouten () zien van schattingen uit een lineair kansmodel. De specificatie bevat de volgende controlevariabelen: geslacht (dummy), hoogste opleidingsniveau (4 categorieën: laag, middelbaar, hoog, onbekend), migratieachtergrond (4 categorieën: geen, westers, niet-westers, onbekend), huishoudtype (5 categorieën: alleenstaande, alleenstaande ouder, gehuwden en samenwonenden, gehuwden en samenwonend met kinderen, anders), leeftijd in jaren, duur van de lopende uitkering in maanden, cumulatief looninkomen 9 tot 3 maanden voor start interventie, en fixed effects voor het startmoment van experiment.

Tabel B7.6a Gevoeligheidsanalyse - ITT-effecten, >27 uur werk per week met controle voor uitgangspositie, meetmoment 24 maanden na start - Utrecht, Wageningen, Nijmegen en Tilburg

	Utrecht	Wageningen	Nijmegen	Tilburg
Ontheffing	0,020	0,054		
	(0,027)	(0,050)		
Intensivering	0,043	0,047		-0,040
	(0,027)	(0,048)		(0,026)
Vrijstelling Inkomensgrens	0,025	0,000		
	(0,025)	(0,045)		
Keuze interventie				
Ontheffing (met vrijstelling inkomensgrens)			-0,074	-0,052*
			(0,047)	(0,027)
Intensivering (met vrijstelling inkomensgrens)			-0,028	-0,054**
			(0,053)	(0,027)
Ontheffing (met app en vrijstelling inkomensgrens)				
Aantal observaties	752	375	289	683
Gemiddelde controle groep	0,059	0,103	0,165	0,102
	(0,235)	(0,306)	(0,373)	(0,304)

Noot: * $p < 0,10$; ** $p < 0,05$. Deze tabel laat coëfficiënten en standaardfouten () zien van schattingen uit een lineair kansmodel. De specificatie bevat de volgende controlevariabelen: geslacht (dummy), hoogste opleidingsniveau (4 categorieën: laag, middelbaar, hoog, onbekend), migratieachtergrond (4 categorieën: geen, westers, niet-westers, onbekend), huishoudtype (5 categorieën: alleenstaande, alleenstaande ouder, gehuwden en samenwonenden, gehuwden en samenwonend met kinderen, anders), leeftijd in jaren, duur van de lopende uitkering in maanden, cumulatief looninkomen 9 tot 3 maanden voor start interventie, dummies voor de kans op meer dan 27 uur werk 24, 16, en 8 maanden voor de start van het experiment, en fixed effects voor het startmoment van experiment. Het laatst beschikbare meetmoment in Utrecht en Tilburg is 16 maanden na start, en in Nijmegen 22 maanden na start. De schattingen voor die gemeentes zijn voor die meetmomenten.

Tabel B7.6b Gevoeligheidsanalyse - ITT-effecten, >27 uur werk per week met controle voor uitgangspositie, meetmoment 24 maanden na start – Deventer en Groningen

	Deventer	Groningen
Ontheffing		-0,022**
		(0,008)
Intensivering		-0,018**
		(0,008)
Vrijstelling Inkomensgrens		-0,010
		(0,009)
Keuze interventie		-0,008
		(0,009)
Ontheffing (met vrijstelling inkomensgrens)	0,014	
	(0,031)	
Intensivering (met vrijstelling inkomensgrens)	0,043	
	(0,033)	
Ontheffing (met app en vrijstelling inkomensgrens)	0,019	
	(0,032)	
Aantal observaties	695	8.190
Gemiddelde controle groep	0,092	0,079
	(0,290)	(0,270)
<p>Noot: * $p < 0,10$; ** $p < 0,05$. Deze tabel laat coëfficiënten en standaardfouten () zien van schattingen uit een lineair kansmodel. De specificatie bevat de volgende controlevariabelen: geslacht (dummy), hoogste opleidingsniveau (4 categorieën: laag, middelbaar, hoog, onbekend), migratieachtergrond (4 categorieën: geen, westers, niet-westers, onbekend), huishoudtype (5 categorieën: alleenstaande, alleenstaande ouder, gehuwden en samenwonenden, gehuwden en samenwonend met kinderen, anders), leeftijd in jaren, duur van de lopende uitkering in maanden, cumulatief looninkomen 9 tot 3 maanden voor start interventie, dummies voor de kans op meer dan 27 uur werk 24, 16, en 8 maanden voor de start van het experiment, en fixed effects voor het startmoment van experiment.</p>		

Tabel B7.7a Gevoeligheidsanalyse - ITT-effecten, >12 uur werk per week met controle voor uitgangspositie, meetmoment 24 maanden na start - Utrecht, Wageningen, Nijmegen en Tilburg

	Utrecht	Wageningen	Nijmegen	Tilburg
Ontheffing	0,031	-0,012		
	(0,033)	(0,063)		
Intensivering	0,068**	-0,018		-0,08**
	(0,033)	(0,063)		(0,036)
Vrijstelling Inkomensgrens	0,058*	-0,075		
	(0,033)	(0,061)		
Keuze interventie				
Ontheffing (met vrijstelling inkomensgrens)			-0,181**	-0,064
			(0,061)	(0,039)
Intensivering (met vrijstelling inkomensgrens)			-0,108*	-0,071*
			(0,065)	(0,039)
Ontheffing (met app en vrijstelling inkomensgrens)				
Aantal observaties	752	375	289	683
Gemiddelde controle groep	0,117	0,287	0,392	0,25
	(0,322)	(0,455)	(0,491)	(0,434)
<p>Noot: * $p < 0,10$; ** $p < 0,05$. Deze tabel laat coëfficiënten en standaardfouten () zien van schattingen uit een lineair kansmodel. De specificatie bevat de volgende controlevariabelen: geslacht (dummy), hoogste opleidingsniveau (4 categorieën: laag, middelbaar, hoog, onbekend), migratieachtergrond (4 categorieën: geen, westers, niet-westers, onbekend), huishoudtype (5 categorieën: alleenstaande, alleenstaande ouder, gehuwden en samenwonenden, gehuwden en samenwonend met kinderen, anders), leeftijd in jaren, duur van de lopende uitkering in maanden, cumulatief looninkomen 9 tot 3 maanden voor start interventie, dummies voor de kans op meer dan 12 uur werk 24, 16, en 8 maanden voor de start van het experiment, en fixed effects voor het startmoment van experiment. Het laatst beschikbare meetmoment in Utrecht en Tilburg is 16 maanden na start, en in Nijmegen 22 maanden na start. De schattingen voor die gemeentes zijn voor die meetmomenten.</p>				

Tabel B7.7b Gevoeligheidsanalyse - ITT-effecten, >12 uur werk per week met controle voor uitgangspositie, meetmoment 24 maanden na start – Deventer en Groningen

	Deventer	Groningen
Ontheffing		-0,035**
		(0,010)
Intensivering		-0,028**
		(0,010)
Vrijstelling Inkomensgrens		-0,014
		(0,010)
Keuze interventie		-0,025**
		(0,010)
Ontheffing (met vrijstelling inkomensgrens)	-0,052	
	(0,037)	
Intensivering (met vrijstelling inkomensgrens)	-0,006	
	(0,040)	
Ontheffing (met app en vrijstelling inkomensgrens)	0,001	
	(0,040)	
Aantal observaties	695	8.190
Gemiddelde controle groep	0,209	0,139
	(0,408)	(0,347)
<p>Noot: * $p < 0,10$; ** $p < 0,05$. Deze tabel laat coëfficiënten en standaardfouten () zien van schattingen uit een lineair kansmodel. De specificatie bevat de volgende controlevariabelen: geslacht (dummy), hoogste opleidingsniveau (4 categorieën: laag, middelbaar, hoog, onbekend), migratieachtergrond (4 categorieën: geen, westers, niet-westers, onbekend), huishoudtype (5 categorieën: alleenstaande, alleenstaande ouder, gehuwden en samenwonenden, gehuwden en samenwonend met kinderen, anders), leeftijd in jaren, duur van de lopende uitkering in maanden, cumulatief looninkomen 9 tot 3 maanden voor start interventie, dummies voor de kans op meer dan 12 uur werk 24, 16, en 8 maanden voor de start van het experiment, en fixed effects voor het startmoment van experiment.</p>		