



Evaluatie experimenten Participatiewet: effecten op brede baten

Een ontheffing van verplichtingen in de bijstand leidt soms tot meer gebruik van Wmo-begeleiding. Intensievere begeleiding leidt in sommige gevallen tot meer gebruik van Wmo-dagbesteding. Ook bij ruimere bijverdienmogelijkheden stijgt soms het Wmo-gebruik. Meer Wmo-ondersteuning wijst op hogere kosten, maar kan ook maatschappelijke baten opleveren, die uiteindelijk bijdragen aan participatie in de samenleving. We vinden geen bewijs voor effecten op gezondheid en problematische schulden.

Het ministerie van Sociale Zaken en Werkgelegenheid (SZW) heeft het CPB gevraagd om onderzoek te doen naar de effecten van de experimenten Participatiewet op brede baten: Wmo-gebruik, gezondheid en problematische schulden.

CPB - november 2022

Timo Verlaat
Alice Zulkarnain

Brede baten bijstandsexperimenten

Gemeenten hebben experimenten uitgevoerd om te kijken wat werkt om meer mensen vanuit de bijstand aan het werk te krijgen. Waren er ook effecten op brede baten: het gebruik van sociale voorzieningen (Wmo), gezondheid en schulden?



Deelnemers

Deventer*

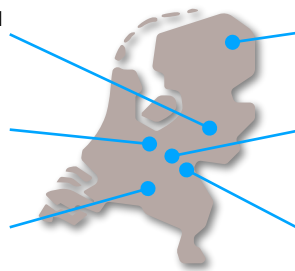
1584 uitgenodigd
364 deelnemers

Utrecht

752 uitgenodigd
710 deelnemers

Tilburg

780 uitgenodigd
780 deelnemers



Groningen*

6551 uitgenodigd
890 deelnemers

Wageningen

410 uitgenodigd
410 deelnemers

Nijmegen

366 uitgenodigd
304 deelnemers



Type experiment + effecten op brede baten



Ontheffing

Deelnemer heeft geen arbeids- en re-integratieverplichtingen



Effect op Wmo-gebruik

Utrecht: 7% van deelnemers krijgt begeleiding vanuit de Wmo, tegen 3% in de controlegroep. Dit is mogelijk op eigen initiatief, als vervanging voor het wegvallen van reguliere begeleiding



Extra begeleiding

Deelnemer krijgt intensievere begeleiding



Effect op Wmo-gebruik

Utrecht: 3% van deelnemers volgt dagbesteding, tegen 1% in de controlegroep. Dit komt mogelijk door beter zicht op behoeften van deelnemers en meer tijd om door te verwijzen naar Wmo-loket



Meer bijverdiensten

Deelnemer mag extra bijverdienen met behoud van uitkering



Effect op Wmo-gebruik

Utrecht: 4% van deelnemers volgt dagbesteding na het experiment, tegen 0,5% in de controlegroep. Mogelijk door contact met de inkomstenafdeling aan het einde van het experiment



Meer Wmo-gebruik wijst op hogere kosten, maar kan ook maatschappelijke baten opleveren, die uiteindelijk bijdragen aan participatie in de samenleving

In de andere vijf gemeenten konden we niet vaststellen of veranderingen in het Wmo-gebruik een direct gevolg van de experimenten waren



We vinden geen bewijs voor effecten op gezondheid en problematische schulden. Dit kan mede komen door databeperkingen, de korte looptijd en de omvang van de interventies



Aanbevelingen

Om meer te kunnen leren van toekomstige experimenten, kunnen deze gebaat zijn bij:



Grotere aantallen deelnemers leiden ertoe dat ook kleinere effecten betrouwbaar waar te nemen zijn



Minder experiment-groepen leiden tot grotere aantallen deelnemers per groep en beter gescheiden interventies



Uniforme opzet van interventies maakt resultaten tussen experimenten in verschillende gemeenten beter vergelijkbaar



Vergelijking van groepen voor de start maakt afwijkingen op tijd zichtbaar, waardoor ze hersteld kunnen worden

Het CPB heeft eerder de effecten op uitstroom naar betaald werk onderzocht, in 2020 ([link](#)) en 2022 ([link](#))

*ander uitnodigingsbeleid

Samenvatting

Het ministerie van Sociale Zaken en Werkgelegenheid (SZW) heeft het Centraal Planbureau (CPB) gevraagd om onderzoek te doen naar de effecten van experimenten met de Participatiewet op brede baten: het gebruik van voorzieningen uit de Wet maatschappelijke ondersteuning (Wmo), gezondheid en problematische schulden. Tussen 2017 en 2019 hebben zes gemeenten via experimenten onderzocht hoe onderdelen van de Participatiewet doeltreffender uitgevoerd kunnen worden. De experimenten bestonden uit drie interventies of combinaties daarvan: een ontheffing van arbeids- en re-integratieverplichtingen, een intensievere begeleiding en/of een vrijlating van bijverdiensten.¹ Het CPB heeft eerder de effecten op de uitstroom naar betaald werk onderzocht (De Boer et al., 2020; Verlaat en Zulkarnain, 2022). Omdat de experimenten mogelijk ook een effect hebben gehad op andere uitkomsten, analyseren we in deze notitie de effecten op brede baten (en kosten) via administratieve data.

In Utrecht leidde het experiment tot meer Wmo-begeleiding in de ontheffingsgroep en meer Wmo-dagbesteding in de intensiverings- en vrijlatingsgroepen. In de overige deelnemende gemeenten zijn de resultaten niet direct toe te schrijven aan de interventies. In vergelijking met de controlegroep lag het gebruik van Wmo-begeleiding in Utrecht tweeënhalve keer zo hoog in de ontheffingsgroep, zowel tijdens als na het experiment. Mogelijk zijn deelnemers na het wegvallen van hun reguliere contact met de klantmanager op zoek gegaan naar andere ondersteuning. Daarnaast steeg het gebruik van dagbesteding tijdens het experiment in de intensiveringsgroep en met name na het experiment in de vrijlatingsgroep. Mogelijk kregen klantmanagers door de intensievere begeleiding beter zicht op de ondersteuningsbehoeften van deelnemers. Ook hadden ze waarschijnlijk meer tijd voor doorverwijzingen naar het Wmo-loket. Een mogelijke verklaring voor de toename in de vrijlatingsgroep aan het einde van het experiment is dat er meer contact kwam met de afdeling inkomstenverrekening vanwege het aflopen van de bijverdienregeling. Ook deze afdeling is getraind in het doorverwijzen naar het Wmo-loket.

Meer gebruik van Wmo-diensten brengt hogere kosten met zich mee, maar kan ook maatschappelijke baten opleveren die op den duur bijdragen aan participatie in de samenleving. Alleen mensen met een Wmo-indicatie komen in aanmerking voor Wmo-diensten. Meestal hebben deze personen een grotere afstand tot de arbeidsmarkt, omdat ze te maken hebben met multiproblematiek. Als mensen via de Wmo de benodigde ondersteuning krijgen die ze anders niet hadden ontvangen, kan dit de afstand tot werk op termijn verkleinen. Bovendien is dagbesteding ook een vorm van participatie voor mensen die (nog) niet volledig mee kunnen draaien op de arbeidsmarkt, waaraan zij baten kunnen ontlenen door bijvoorbeeld meer persoonlijke aandacht en ontwikkeling.

De resultaten op gezondheid en problematische schulden zijn beperkt. Dit kan mede komen door databeperkingen, de korte looptijd van het experiment en de omvang van de interventies. We vinden geen overtuigende resultaten voor effecten op medicijn- en zorggebruik. Door tegengestelde mechanismen zijn effecten mogelijk klein en moeilijk te detecteren. Ook zijn data deels in brede categorieën samengevat en alleen op jaarbasis beschikbaar, terwijl de experimenten vaak midden in het jaar zijn gestart. Verder waren de experimenten mogelijk te kort om te resulteren in veranderingen in medicijn- en zorggebruik. We vinden daarnaast geen bewijs voor effecten op wanbetaling van de zorgpremie, wat we als indicator gebruiken voor problematische schulden. Ook deze data zijn alleen op jaarbasis beschikbaar. Bovendien waren de effecten op werkhervatting veelal beperkt, wat de mogelijkheid om schulden af te lossen beperkt.

¹ In de reguliere situatie mogen bijstandsgerechtigden maximaal 25% van hun inkomsten uit arbeid houden, tot een maximum van circa 200 euro per maand. Deze vrijlating van inkomsten geldt voor maximaal zes maanden. Voor de vrijlatingsinterventie tijdens de experimenten werd deze regeling verruimd.

Om de causale effecten op brede baten goed in kaart te brengen is er meer onderzoek nodig, met grotere aantallen deelnemers, een langere looptijd en meer gedetailleerdere data. In dit onderzoek vinden we enkele effecten op brede baten. Om causale effecten in de toekomst nog beter empirisch vast te kunnen leggen, zijn goed opgezette experimenten met grotere aantallen nodig (zie het kader 'Lessen voor toekomstige experimenten'). Als het tijd kost voordat effecten zich uiten, is onderzoek gebaat bij een langere looptijd. Verder kunnen data met een hogere frequentie en gedetailleerdere informatie bijdragen aan een nauwkeurige meting van mogelijke effecten.

Lessen voor toekomstige experimenten

Experimenten kunnen belangrijke inzichten opleveren voor *evidence-based* beleid. Met een experiment kunnen de effecten van specifieke beleidsinterventies worden getest, voordat ze volledig worden uitgerold (Koning, 2011; Van der Klaauw, 2014). Ook kunnen de causale effecten van (wijzigingen van) beleid worden bepaald, doordat deelnemers aselekt worden toegewezen aan controle- en interventiegroepen. Dit is van groot belang, omdat alleen causale effecten een oorzakelijk verband tussen een beleidsinterventie en een uitkomst aantonen.

Uit de experimenten met de Participatiewet zijn een aantal belangrijke lessen te trekken voor toekomstige experimenten. Gemeenten hebben voor het eerst op basis van het experimenteerartikel Participatiewet op grote schaal experimenten uitgevoerd om te onderzoeken welke interventies het beste werken voor de bijstandspopulatie. Uit de analyse volgen een aantal aanbevelingen voor toekomstige experimenten:

1. **Experimenteer met grotere aantallen.** Een groter aantal deelnemers in de controle- en interventiegroepen zorgt ervoor dat ook kleinere effecten met voldoende betrouwbaarheid van nuleffecten kunnen worden onderscheiden. Ook maken grotere aantallen het mogelijk om analyses op subgroepen uit te voeren.
2. **Beperk het aantal interventiegroepen.** Het aantal deelnemers per groep kan onder andere worden verhoogd door minder interventies tegelijk te testen. Bij minder interventies is het bovendien minder waarschijnlijk dat er fouten worden gemaakt bij de uitvoering, of dat interventies niet strikt gescheiden zijn, bijvoorbeeld omdat klantmanagers meerdere interventies tegelijkertijd moeten uitvoeren.
3. **Zorg voor uniforme interventies bij experimenten in meerdere gemeenten.** Uniforme interventies maken resultaten beter vergelijkbaar wanneer experimenten in verschillende gemeenten plaatsvinden. Daarnaast kunnen data makkelijker worden gepoold, waardoor er grotere aantallen ontstaan.
4. **Zorg voor voldoende verschil met de reguliere situatie.** Als de geteste interventie weinig verschilt van de reguliere situatie, zijn de effecten naar verwachting klein en moeilijk waar te nemen met een beperkt aantal deelnemers.
5. **Overweeg een referentiegroep.** Er kunnen gedragseffecten ontstaan als deelnemers in de controlegroep weten dat zij meedoen aan een experiment en de reguliere behandeling krijgen. Met een referentiegroep, waarin deelnemers de reguliere behandeling krijgen maar geen informatie hebben ontvangen, kunnen deze effecten worden geanalyseerd en uitgesloten (Duflo et al., 2007). Hierbij is wel van belang dat deze groep door hetzelfde randomisatieproces tot stand is gekomen als de overige experimentengroepen. Ook een ruim aantal deelnemers is een voorwaarde.

6. **Kies een geschikte randomisatievolgorde.** Als deelname aan het experiment verplicht is, is het vaak raadzaam om een aselechte streekproef of de hele doelgroep te randomiseren. De gemeten effecten zijn dan in theorie van toepassing op de doelgroep, wat leidt tot zogenoemde externe validiteit. Als deelname aan het experiment vrijwillig is, stijgt het risico op (selectieve) uitval. Personen kunnen dan beter eerst worden uitgenodigd voordat er gerandomiseerd wordt. Dit verkleint het risico op uitval en gedragseffecten, wat de zogenoemde interne validiteit vergroot. De gemeten effecten zijn echter minder informatief voor de doelgroep, omdat de geworven groep selectief is.
7. **Gebruik een geschikte randomisatiemethode.** Een simpele loting leidt naar verwachting tot experimentgroepen die vergelijkbaar zijn met betrekking tot waarneembare en niet-waarneembare kenmerken. Bij kleine aantallen bestaat er echter de kans dat de groepen per toeval niet vergelijkbaar zijn. Deze kans kan verkleind worden door een zogenoemde gestratificeerde randomisatie of een *pair-wise* randomisatie toe te passen (Bruhn en McKenzie, 2009). Deze methoden houden onder andere rekening met bepaalde achtergrondkenmerken of uitkomsten voor de start, waardoor de groepen wat betreft de gekozen factoren beter vergelijkbaar zijn.
8. **Controleer de vergelijkbaarheid van groepen voor de start van het experiment.** Als er informatie over achtergrondkenmerken beschikbaar is, kan de vergelijkbaarheid van de controle- en interventiegroepen voor de start van het experiment worden onderzocht. Het is raadzaam om ook informatie te verzamelen over de uitkomsten voor de start van een experiment (bijvoorbeeld arbeidsdeelname en Wmo-gebruik voor de start) en deze te gebruiken voor de vergelijkbaarheidsanalyse. Bij onvoldoende vergelijkbaarheid kan een herrandomisatie overwogen worden. Hiervoor bestaan verschillende methoden (Bruhn en McKenzie, 2009).
9. **Zorg voor een eenduidige behandeling volgens protocol.** Klantmanagers hebben hart voor de zaak en vinden het mogelijk lastig om personen uit te sluiten van een bepaalde behandeling. Interventies kunnen echter alleen causaal op hun werkzaamheid worden getest als zij volgens het protocol zijn uitgevoerd. Effecten worden bepaald door het verschil met de reguliere situatie. Het is daarom belangrijk dat de reguliere behandeling in de controlegroep niet verandert en goed omschreven is.
10. **Verzamel tijdens het experiment informatie over de uitvoering.** Afwijkingen van het experimentele protocol worden op tijd zichtbaar gemaakt door data over bijvoorbeeld contactmomenten en ingezette dienstverlening. Via deze data kan ook worden bepaald hoe groot het verschil met de reguliere situatie was voor deelnemers in de interventiegroep.
11. **Overweeg langere looptijden.** Een langere looptijd maakt het mogelijk om interventies ook voor nieuwe instroom te testen, in plaats van alleen het zittende bestand. Nieuwe instroom kan er ook bij helpen om de deelnemersaantallen vergroten. Het is dan wel belangrijk dat de vergelijkbaarheid van de controle- en interventiegroepen gewaarborgd blijft en dat de duur van het experiment ook voor de latere instroom lang genoeg is. Daarnaast zijn langere looptijden van belang om effecten te kunnen meten die langer duren om te ontwikkelen, zoals het vinden van een baan voor mensen met een grotere afstand tot de arbeidsmarkt. De experimenten met de Participatiewet konden niet langer lopen door een tijdslimiet in het experimenteerartikel van de wet. Toekomstige experimenten kunnen gebaat zijn bij een aanpassing, zodat langere looptijden mogelijk zijn.

1 Inleiding

Zes gemeenten hebben gedurende twee jaar bijstandsexperimenten uitgevoerd om te onderzoeken of er mogelijkheden zijn om de Participatiewet doeltreffender uit te voeren. In 2017 heeft de minister van Sociale Zaken en Werkgelegenheid (SZW), op grond van het Tijdelijk besluit experimenten Participatiewet, zes gemeenten aangewezen als experimenteergemeenten: Deventer, Groningen,² Nijmegen, Tilburg, Utrecht³ en Wageningen. Het doel van de experimenten was om mogelijkheden te onderzoeken om onderdelen van de Participatiewet doeltreffender uit te voeren. De volgende interventies of combinaties daarvan werden onderzocht: een ontheffing van arbeids- en re-integratieverplichtingen, een intensievere begeleiding en/of een langere en hogere vrijlating van bijverdiensten. De experimenten liepen tot en met december 2019.

Eerdere analyses wezen uit dat de effecten op uitstroom naar betaald werk beperkt waren en dat de resultaten alleen in Utrecht, Deventer en Wageningen geïnterpreteerd konden worden als causale effecten van de interventies. Het Centraal Planbureau (CPB) heeft op verzoek van het ministerie van SZW tweemaal een kwantitatieve analyse uitgevoerd naar de effecten van de experimenten op de uitstroom naar betaald werk. De eerste analyse werd uitgevoerd in 2020, direct na afloop van de experimenten (De Boer et al., 2020). De tweede analyse vond plaats in 2022, met focus op effecten op de langere termijn en op uitstroom naar zelfstandig ondernemerschap (Verlaat en Zulkarnain, 2022). In Utrecht leidden een intensievere begeleiding en ruimere bijverdienmogelijkheden tijdens het experiment soms tot een hogere kans op gedeeltelijke of volledige uitstroom naar betaald werk (gedefinieerd als meer dan 12 uur werk per week). Na afloop van het experiment zagen we in Utrecht een hogere uitstroom in de ontheffingsgroep en soms in de groep met een intensievere begeleiding. In Wageningen was er na afloop van het experiment soms een lagere kans op volledige uitstroom naar betaald werk (gedefinieerd als meer dan 27 uur werk per week) in de groep met ruimere bijverdienmogelijkheden, welke na het experiment weer werden opgeheven. Uit de eerdere analyses bleek verder dat de resultaten in Groningen, Nijmegen en Tilburg niet direct te interpreteren zijn als het gevolg van de interventies, omdat de controle- en interventiegroepen niet vergelijkbaar waren.

De academische literatuur suggereert dat arbeidsmarktbeleid tot bredere baten kan leiden, ook bij mensen met een grotere afstand tot de arbeidsmarkt. Daarom heeft het ministerie van SZW het CPB gevraagd om te kijken naar de effecten van de experimenten op het gebruik van voorzieningen uit de Wet maatschappelijke ondersteuning (Wmo), gezondheid en problematische schulden. De theorie uit de wetenschappelijke literatuur beschrijft dat er relaties kunnen zijn tussen het vinden van werk en iemands gezondheid of schuldensituatie. Verder is er ook empirisch bewijs voor verbanden tussen werk en gezondheid en tussen werk en Wmo-gebruik onder mensen met een grotere afstand tot de arbeidsmarkt. Daarnaast kunnen sommige van de geteste interventies, namelijk het wegnemen van verplichtingen en het krijgen van extra begeleiding, ook een direct effect op deze uitkomsten hebben. Daarom heeft het ministerie van SZW het CPB gevraagd om te onderzoeken of de experimenten mogelijk ook andere effecten hebben gehad dan op de uitstroom naar betaald werk. Ook in de gemeentelijke eindrapportages is hieraan veel aandacht besteed.⁴

In tegenstelling tot eerder onderzoek naar brede baten van de gemeenten zelf, leggen we de focus in deze notitie niet op enquêtedata maar op administratieve data, die ruim voor de start van de experimenten beschikbaar zijn voor alle deelnemers. Administratieve data reiken meestal verder terug dan enquêtedata, waardoor verschillen in uitkomsten voor de start van een experiment beter kunnen worden onderzocht.

² De gemeente Groningen werkte samen met de gemeente Ten Boer.

³ De gemeente Utrecht werkte samen met de gemeente Zeist.

⁴ Zie Betkó et al., 2020; Edzes et al., 2020; Gramberg en De Swart, 2020; Muffels et al., 2020a; Muffels et al., 2020b; Verlaat et al., 2020. We bedanken de betrokken onderzoekers voor hun deelname aan een klankbordgroep en voor hun feedback en suggesties.

Daarnaast zijn deze data beschikbaar voor alle bijstandsgerechtigden en niet alleen voor deelnemers die reageren op de vragenlijsten, wat selectieve respons tegengaat. Bovendien wordt administratieve informatie niet beïnvloed door sociaal wenselijke antwoorden. Een beperking van de administratieve data is dat deze alleen feitelijke uitkomsten bevatten en geen subjectieve belevingen uitvragen. Ook moet er, afhankelijk van de onderzoeksvraag en de beschikbaarheid van data, in sommige gevallen beschikbare informatie als indicator worden gebruikt voor iets waarover geen data voorhanden is (ook *proxy* genoemd).

De opbouw van de notitie is als volgt. Hoofdstuk 2 geeft een kort overzicht van de experimenten in de verschillende gemeenten. Hoofdstuk 3 bespreekt de theoretische achtergrond en de mogelijk te verwachten effecten. Hoofdstuk 4 beschrijft de gebruikte data. Hoofdstuk 5 presenteert beschrijvende statistieken en analyseert de vergelijkbaarheid van de controle- en interventiegroepen. Hoofdstuk 6 bespreekt de empirische methode en hoofdstuk 7 de empirische resultaten. Tot slot geeft hoofdstuk 8 de hoofdconclusies, enkele kanttekeningen bij de analyses en mogelijkheden voor toekomstig onderzoek.

2 Beschrijving experimenten Participatiewet

In Deventer, Groningen, Nijmegen, Tilburg, Utrecht en Wageningen zijn experimenten uitgevoerd op onderdelen van de Participatiewet. Op grond van het Tijdelijk besluit experimenten Participatiewet heeft de minister van SZW deze zes gemeenten aangewezen als experimenteergemeenten. Dit gaf deze gemeenten de bevoegdheid om bij wijze van experiment gedurende maximaal twee jaar af te wijken op onderdelen van de Participatiewet.

De experimenten bestonden grofweg uit drie typen interventies of een combinatie daarvan: een ontheffing van de arbeids- en re-integratieverplichtingen, een intensievere begeleiding en een langere en hogere vrijlating van bijverdiensten. De gemeenten konden de volgende specifieke interventies inzetten bij bijstandsgerechtigden in de experimenten:

- Een tijdelijke ontheffing van de arbeids- en re-integratieverplichtingen (ontheffingsgroep).⁵
- Een tijdelijke intensivering van de begeleiding (intensiveringsgroep).
- Een vrijlating van de inkomsten uit arbeid tot maximaal 50% van deze inkomsten, met een maximum van circa 200 euro per maand (vrijlatingsgroep).⁶ De vrijlating gold gedurende de hele looptijd van de experimenten in plaats van de reguliere zes maanden.
- 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 deze bespreken we hierna in meer detail (zie ook paragraaf 2.2 in De Boer et al., 2020).

In de ontheffings- en intensiveringsgroepen ontvingen deelnemers een andere dienstverlening, terwijl de ruimere bijverdienmogelijkheid in de vrijlatingsgroep alleen een administratieve verandering inhield. In de meeste gemeenten ontvingen deelnemers in de ontheffingsgroep alleen op eigen initiatief

⁵ In Deventer was er een variant waarbij de ontheffingsgroep digitale begeleiding ontving via een app (zie ook De Boer et al., 2020).

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

gemeentelijke re-integratiedienstverlening (Deventer, Groningen, Nijmegen en Utrecht), waardoor er tijdens het experiment weinig contact was tussen klantmanagers en deelnemers in deze groep.⁷ In de andere gemeenten kregen deelnemers met een ontheffing begeleiding bij het leren van zelfredzaamheid, waardoor contact met de sociale dienst soms bleef bestaan. Voor alle gemeenten geldt dat klantmanagers in de intensiveringsgroepen meer tijd een ruimte hadden voor persoonlijk contact en maatwerk. Deels werd het dienstverleningspakket voor deze groep aangevuld met nieuwe instrumenten. In Utrecht en Tilburg was er een vaste klantmanager in de intensiveringsgroep, terwijl de reguliere dienstverlening geen toewijzing van bijstandsgerechtigden aan vaste klantmanagers kent. De vrijlatingsgroep behield de reguliere dienstverlening.

Deelnemers zijn per randomisatie ingedeeld in de verschillende experimentgroepen. Personen zijn aselect toegewezen aan verschillende interventiegroepen (een uitnodiging voor het experiment en een aangepaste behandeling) en aan een controlegroep (een uitnodiging voor het experiment, maar in principe geen andere behandeling dan regulier gebruikelijk). In de gemeenten Deventer en Groningen was er ook een groep aselect getrokken personen die geen uitnodiging voor het experiment kreeg en de reguliere behandeling behield (de zogenoemde referentiegroep). Deelname aan het experiment was vrijwillig; personen konden er dus voor kiezen om niet (meer) deel te nemen aan het experiment. Met behulp van de administratieve data konden we zowel de personen die wel als niet (meer) deelnamen aan het experiment volgen over de tijd.

In Groningen en Deventer zijn de bijstandsgerechtigden eerst gerandomiseerd en vervolgens uitgenodigd, terwijl ze in de overige vier gemeenten eerst zijn uitgenodigd en vervolgens gerandomiseerd. In Groningen en Deventer is eerst de randomisatie uitgevoerd en zijn personen vervolgens uitgenodigd. Daarbij is in Groningen in de uitnodiging vermeld in welke groep men is ingedeeld. In Deventer is de groepsindeling tijdens een intakegesprek bekendgemaakt. In de vier andere gemeenten zijn personen eerst uitgenodigd en daarna pas aselect ingedeeld. Ondanks dat in alle gemeenten doorgaans dezelfde groepen bijstandsgerechtigden werden uitgenodigd, verschillen de deelnemersbestanden licht tussen de gemeenten. Dit kan bijvoorbeeld komen door een andere samenstelling van bijstandspopulaties of andere keuzes omtrent deelnamecriteria.⁸ Ten slotte verschilden de startmomenten tussen de gemeenten en gebruikten sommige gemeenten meerdere startmomenten, of nodigden zij ook nieuwe instroom uit voor deelname (zie ook tabel 2.1).

⁷ In Utrecht konden deelnemers in de ontheffingsgroep zelf kiezen, maar een ruime meerderheid (94%) koos ervoor om geen begeleiding meer te ontvangen. In Groningen en Deventer was dienstverlening wel mogelijk, maar werd deze niet actief aangeboden.

⁸ In Wageningen waren jongeren onder de 27 jaar bijvoorbeeld toegelaten tot het experiment, terwijl dat in andere gemeenten niet het geval was. Ook de selectie op bijstandsduur verschilde tussen gemeenten. Terwijl deelnemers in Groningen minstens 26 weken een uitkering moesten hebben, lag de grens in Utrecht, Deventer en Nijmegen bij minimaal 10 tot 12 weken. Wageningen kende geen selectie op bijstandsduur. In Tilburg lag de grens in feite bij 12 maanden, doordat bijstandsgerechtigden met een kortere duur begeleid werden door Sagenn/Diamant-groep en daardoor niet mee konden doen aan het experiment.

Tabel 2.1 Overzichtstabel startmomenten en duur experiment, per gemeente

	Deventer	Groningen	Nijmegen	Tilburg ^b	Utrecht	Wageningen ^b
Start experiment	1 oktober '17 1 februari '18 1 juli '18	1 november '17	1 december '17 1 april '18	1 november '17 en continu t/m 1 oktober '18	1 juni '18	1 oktober '17 en continu t/m 1 september '18
Startgroep 1	Zittend bestand met ≥12 weken uitkering (tranche 1)	Zittend bestand met ≥26 weken uitkering	Zittend bestand	Zittend bestand	Zittend bestand met ≥10 weken uitkering	Zittend bestand
Startgroep 2	Zittend bestand met ≥12 weken uitkering (tranche 2)		Zittend bestand	Nieuwe instroom tussen februari '18 en oktober '18		Nieuwe instroom tussen november '17 en september '18
Startgroep 3	Nieuwe instroom tussen augustus '17 en mei '18					
Einde experiment ^a	31 december '19	31 december '19	31 december '19	31 december '19	31 december '19	31 december '19
Totale duur	17-27 maanden	27 maanden	21-25 maanden	Maximaal 26 maanden	19 maanden	Maximaal 27 maanden

^a De oorspronkelijke einddatum was 30 september 2019. De experimenten werden met drie maanden verlengd om het onderzoek goed af te kunnen ronden en de terugkeer naar de reguliere situatie goed vorm te kunnen geven. ^b Tilburg (Wageningen) kende verschillende startmomenten tot aan 1 oktober (1 september) 2018, omdat naast het zittende bestand ook elke twee weken is geworven en gerandomiseerd onder nieuwe instroom in de bijstand.

3 Theoretische achtergrond

In dit hoofdstuk bespreken we hoe de geteste interventies Wmo-gebruik, gezondheid en schuldenproblematiek kunnen hebben beïnvloed. Wetenschappelijke literatuur wijst erop dat er verbanden kunnen bestaan tussen de geteste interventies en andere sociaal-economische uitkomsten. We bespreken per domein de theoretische mechanismen die naar verwachting een rol spelen bij de relaties tussen de experimenten en Wmo-gebruik, gezondheid en problematische schulden.

De theorie is niet altijd eenduidig en daarom bespreken we, indien beschikbaar, ook empirisch bewijs uit de literatuur. Uit die bespreking blijkt dat de verschillende onderliggende mechanismen dikwijls een tegengestelde richting kunnen hebben. Het is daarom een empirische vraag wat de richting van het effect in de praktijk is. De empirische literatuur rondom Wmo-gebruik en problematische schulden is beperkt; waar beschikbaar bespreken we deze. De verbanden tussen werk en fysieke en mentale gezondheid zijn uitgebreid empirisch onderzocht, ook specifiek voor personen met een grotere afstand tot de arbeidsmarkt. Deze literatuur bespreken we eveneens in dit hoofdstuk.

3.1 Wmo-gebruik

Voor Wmo-gebruik volgen wij Van Eijkel et al. (2020) en richten we ons op twee voorzieningen: begeleiding en dagbesteding inclusief arbeidsmatige activering. De Wet maatschappelijke ondersteuning (Wmo) biedt via maatwerkvoorzieningen hulp en ondersteuning die op de persoonlijke situatie is afgestemd. Maatwerkvoorzieningen worden ingedeeld in vier hoofdgroepen: ondersteuning thuis, hulp bij het huishouden, verblijf en opvang, en hulpmiddelen en diensten. Zoals in een eerder CPB-onderzoek (Van Eijkel

et al., 2020) richten we ons op twee voorzieningen uit de eerste categorie.⁹ Begeleiding vanuit de Wmo biedt ondersteuning in het dagelijkse leven, bijvoorbeeld bij het onderhouden van sociale contacten, bij het uitoefenen van activiteiten en hobby's, bij administratie en geldzaken, of bij de ontwikkeling van vaardigheden. Onder dagbesteding vallen activiteiten om de dag zinvol of prettig te vullen, meestal onder begeleiding van professionals. Ook arbeidsmatige activering, waarbij het gaat om werk met veel begeleiding, hoort bij deze voorziening.¹⁰

Een ontheffing van arbeids- en re-integratieverplichtingen kan theoretisch leiden tot zowel meer als minder gebruik van Wmo-begeleiding en -dagbesteding. Bij minder contact kan de behoefte aan ondersteuning toenemen, maar ook kunnen doorverwijzingen naar Wmo-diensten juist afnemen. In sommige gemeenten was er nagenoeg geen contact meer tussen deelnemers in de ontheffingsgroep en klantmanagers (zie hoofdstuk 2). Mogelijk nam de behoefte aan ondersteuning hierdoor toe en gingen deelnemers zelf op zoek naar andere diensten, zoals via het Wmo-loket of door dit aan te geven bij hun (voormalige) klantmanagers. Uit gesprekken met de gemeenten kwam naar voren dat klantmanagers deelnemers niet op structurele wijze hebben aangemeld voor Wmo-diensten. Maar mogelijk zijn enkele deelnemers, naar aanleiding van een verzoek, gewezen op de beschikbaarheid van Wmo-diensten. Anderzijds kan beperkt contact met de klantmanager ertoe leiden dat behoefte aan Wmo-ondersteuning minder vaak gesignaleerd wordt. Aan het eind van het experiment kan deze behoefte wel weer gesignaleerd worden, wegens terugkeer naar de reguliere dienstverlening. Wel verwachten we dat personen die een Wmo-indicatie hebben gekregen gebruik blijven maken van de Wmo-diensten.

Een intensievere begeleiding kan in theorie resulteren in meer Wmo-begeleiding en -dagbesteding, als klantmanagers beter inzicht krijgen in de behoeften van de cliënten. Bepaalde onderdelen, zoals Wmo-begeleiding, worden vaak gebruikt door mensen met een psychische of psychosomatische problematiek (Van Eijkel et al., 2020). Tijdens de experimenten kregen klantmanagers meer tijd om de cliënten in de intensiveringsgroep te leren kennen en om hun behoeften in kaart te brengen. Hierdoor werd de behoefte aan additionele (geestelijke) zorg of vormen van activering mogelijk meer inzichtelijk. Daarnaast is er in sommige gemeenten intensiever samengewerkt met de sociale wijkteams, die verantwoordelijk zijn voor de Wmo-diensten, wat de afstand tot doorverwijzing en het verkrijgen van een Wmo-indicatie korter maakt. Naar verwachting blijven de deelnemers die een Wmo-indicatie hebben gekregen ook na het einde van het experiment nog een tijd gebruikmaken van Wmo-diensten.

Eerder empirisch onderzoek heeft vastgesteld dat het vinden van een baan kan leiden tot minder gebruik van Wmo-ondersteuning (Van Eijkel et al., 2020), omdat er naast een baan minder tijd en mogelijk minder behoefte is. Personen hebben na het vinden van een baan minder tijd om gebruik te maken van Wmo-diensten. Ook hebben ze om die reden minder behoefte aan dagbesteding. Daarnaast kan de behoefte aan Wmo-begeleiding afnemen als de baan financiële stress verlicht en de mentale gesteldheid verbetert. Eerder CPB-onderzoek toonde al aan dat mensen met een arbeidsbeperking minder vaak een beroep doen op Wmo-ondersteuning als zij werk hebben (Van Eijkel et al., 2020). Voor de ontheffingsgroep en de intensiveringsgroep kan dit indirecte effect via werkhervatting nog bijdragen aan het totale effect op Wmo-gebruik. Voor de vrijlatingsgroep verwachten we alleen een effect via het vinden van werk.

⁹ Van Eijkel et al. (2020) richten zich ook op de voorziening beschermd wonen uit de categorie verblijf en opvang. We laten deze voorziening buiten beschouwing omdat een eerste dataverkenning heeft uitgewezen dat deze voorziening nauwelijks wordt ingezet in de zes deelnemende gemeenten.

¹⁰ Arbeidsmatige dagbesteding vanuit de Wmo is vergelijkbaar met beschermt werk vanuit de Participatiewet. In tegenstelling tot beschermt werk, waarbij de werknemer een dienstbetrekking heeft en loon ontvangt, betalen personen voor arbeidsmatige dagbesteding een eigen bijdrage.

In sommige gemeenten waren Wmo-diensten onderdeel van het dienstverleningspakket voor de intensiveringsgroep, wat de afstand tot deze diensten verkleint. In twee gemeenten behoorden Wmo-diensten tot de instrumenten voor de intensiveringsgroep. In Nijmegen konden klantmanagers deelnemers doorverwijzen naar een traject voor arbeidsmatige activering. Ook in Tilburg konden klantmanagers Wmo-diensten (waaronder dagbesteding) inzetten als ze dat nodig vonden. Naar verwachting verkleint deze aanpak de afstand tot doorverwijzingen en het verkrijgen van een Wmo-indicatie nog verder.

3.2 Gezondheid

Een ontheffing van arbeids- en re-integratieverplichtingen kan in theorie zowel positieve als negatieve effecten hebben op de fysieke en mentale gezondheid. Enerzijds kan een ontheffing stress verminderen, anderzijds kan men zich ook in de steek gelaten voelen. Deelnemers die de arbeids- en re-integratieverplichtingen als last ervaren, kunnen zich in de ontheffingsgroep bevrijd voelen, met een vermindering van stress als gevolg. Voor deelnemers die het contact met de klantmanager daarentegen aangenaam vonden, kan het wegvallen van contact juist stress veroorzaken. Via deze twee kanalen kan een ontheffing de gezondheid positief of negatief beïnvloeden, omdat stress gerelateerd is aan een scala van gezondheidsproblemen (Yaribeygi et al., 2017). Ten slotte zouden effecten juist weer om kunnen draaien bij de terugkeer naar de reguliere situatie aan het einde van het experiment.

Ook een intensievere begeleiding kan tegengestelde effecten hebben op de gezondheid. Sommigen kunnen zich gesteund voelen, wat stress vermindert, terwijl anderen de interventie juist als betuttelend kunnen ervaren en daardoor meer stress ondervinden. De mechanismen achter de relatie tussen intensivering en gezondheid spiegelen de mechanismen achter ontheffing en gezondheid. Sommige personen vinden de intensievere begeleiding een welkome ondersteuning en ervaren daardoor minder stress in hun leven. Anderen hebben geen behoefte aan meer contact en vinden dit juist stressvol. Daarnaast kan deze interventie ook indirect positief zijn voor de gezondheid, wanneer intensievere begeleiding tot meer ondersteuning via andere sociale voorzieningen leidt, waaronder Wmo of schuldhulpverlening. Deze indirecte effecten kunnen ook na het einde van het experiment aanhouden, omdat ze mogelijk voor structurele veranderingen zorgen.

We verwachten in theorie geen gezondheidseffecten van een ruimere bijverdienregeling, behalve via het vinden van betaald werk. Van de mogelijkheid om meer bij te verdienen verwachten we geen directe effecten op de gezondheid. Maar omdat een hoger inkomen gerelateerd is aan positieve gezondheidseffecten, kan deze interventie via het vinden van betaald werk wel een additionele positieve uitwerking hebben op de gezondheid.¹¹

Betaald werk vinden kan in theorie zowel positieve als negatieve gezondheidseffecten hebben. Een baan hebben verhoogt het inkomen, vermindert (financiële) stress (Mullainathan en Shafir, 2013; Haushofer en Fehr, 2014) en leidt tot meer fysieke activiteit.¹² Deze aspecten zijn positief voor de fysieke en mentale gezondheid. Anderzijds kan een baan ook juist een bron van stress zijn (Marcatto et al., 2016) en verhoogt het de kans op een werkgerelateerd ongeval, met negatieve gevolgen voor de fysieke gezondheid. Het gebruik van

¹¹ In de literatuur is geconstateerd dat er een positieve relatie bestaat tussen een hoger inkomen en een goede gezondheid of langer leven (Frijters et al., 2005; Cutler et al., 2006; Chetty et al., 2016). Meerdere factoren kunnen hierbij een rol spelen, waaronder ongelijkheid, economische en sociale stress, verschillen in toegang tot gezondheidszorg, goede huisvesting en gezonde voeding (Woolf en Braveman, 2011; Chetty et al., 2016; French et al., 2019).

¹² Stress is gerelateerd aan verschillende gezondheidsproblemen (Yaribeygi et al., 2017). Mensen met een baan bewegen meer ten opzichte van mensen zonder baan, ongeacht de aard van het beroep. Inactiviteit is gerelateerd aan overgewicht, hart- en vaatziekten, een hoge bloeddruk en chronische aandoeningen. Daarom kan het vinden van een baan positief zijn voor de gezondheid (Van Domelen et al., 2011).

verslavende middelen en investeringen in de gezondheid kunnen zowel toe- als afnemen na het vinden van een baan en hierdoor tegengestelde effecten hebben op de fysieke en mentale gezondheid.¹³

De empirische literatuur vindt vaak positieve effecten van werkhervatting op de mentale gezondheid, maar geen effecten op de fysieke gezondheid. De literatuur laat echter ook zien dat slechte werkomstandigheden en baanonzekerheid de gezondheid weer kunnen verslechteren. Huber et al. (2011) bestuderen de gezondheidseffecten van werkhervatting onder Duitse bijstandsgerechtigden. Zij vinden significante positieve effecten op de mentale gezondheid, maar niet op de algemene gezondheidsstatus. Van Eijkel et al. (2020) bestuderen het effect van het sluiten van de sociale werkvoorzieningen na de invoering van de Participatiewet. Dit leidde tot een lagere baankans voor de doelgroep van deze voorzieningen. Hun onderzoek laat zien dat het hebben van een baan voor mensen met arbeidsbeperkingen tot minder ggz-gebruik leidt, maar geen effect heeft op het gebruik van fysieke gezondheidszorg. De literatuur laat echter ook zien dat de kwaliteit van een baan belangrijk is. Mensen met een precaire arbeidsmarktpositie hebben vaak te maken met baanonzekerheid en slechte werkomstandigheden. Zulke factoren kunnen de gezondheid weer verslechteren (Fletcher et al., 2011; Caroli, 2016; Ravesteijn et al., 2018).

De empirische bevindingen over de gezondheidseffecten van actief arbeidsmarktbeleid zijn niet eenduidig. In recente literatuur over de gezondheidseffecten van actief arbeidsmarktbeleid wordt over het algemeen geconcludeerd dat deze effecten positief zijn voor zowel mentale als fysieke gezondheid en welzijn in het algemeen (Caliendo et al., 2022; Coutts et al., 2014; Puig-Barrachina et al., 2010). In oudere literatuur is daarentegen een negatief verband te zien tussen trainingen en de mentale gezondheid (Creed et al., 1998; Machin en Creed, 2003). Daarnaast zijn er in de empirische literatuur verschillende gezondheidseffecten van beleidswijzigingen in de bijstand onderzocht. Bolhaar et al. (2019) vinden geen effect op de totale uitgaven aan gezondheidszorg (huisarts of fysieke gezondheidszorg) na de invoering van een verplichte sollicitatieperiode van vier weken voor bijstandsaanvragers in Amsterdam, maar wel op mentale gezondheidsuitgaven.¹⁴ Wickham et al. (2020) onderzochten de effecten van langere wachttijden, wijzigingen in de uitkeerfrequentie en lagere uitkeringen in de bijstand in het Verenigd Koninkrijk. Zij vonden statistisch significante negatieve effecten op de mentale gezondheid.

De gezondheidseffecten van de experimenten zijn een empirische vraag, doordat er een gebrek is aan een eenduidige richting uit de theoretische en empirische literatuur. In theorie zijn er vaak zowel positieve als negatieve mechanismen van de interventies denkbaar, zowel direct als indirect via het vinden van betaald werk. Ook de empirische bevindingen over de gezondheidseffecten van betaald werk en actief arbeidsmarktbeleid zijn niet eenduidig. Daarnaast blijkt uit de eerdere notities dat uitstroom naar betaald werk beperkt is, waardoor de tegengestelde directe gezondheidsmechanismen waarschijnlijk een grotere rol spelen. Om deze redenen is empirisch bewijs nodig om de effecten te bepalen.

3.3 Problematische schulden

Betaald werk vinden kan in theorie positief zijn voor de financiële situatie van een persoon met schulden, hoewel er beperkt empirisch bewijs is voor deze relatie. Betaald werk kan het inkomen

¹³ Enerzijds kunnen stress op het werk en een hoger inkomen er toe leiden dat men meer verslavende middelen consumeert. Anderzijds rookt of drinkt men na het vinden van een baan mogelijk minder, omdat het op het werk niet is toegestaan en de werkloosheidsstress is verdwenen (Ruhm en Black, 2002; Ruhm, 2005; Saffer et al., 2016; Everding en Marcus, 2020; Gibson et al., 2020). Als een goede gezondheid en gezondheidsinvesteringen normale goederen zijn, zal de vraag ernaar stijgen na het vinden van een baan (Dardanoni en Wagstaff, 1990; Chang 1996). Aan de andere kant is er is minder tijd om bijvoorbeeld gezond te koken en te sporten (Payne et al., 2002; Ruhm, 2005; Shukri et al., 2016).

¹⁴ Bolhaar et al. (2019) vermoeden dat het mechanisme achter deze relatie via het vinden van betaald werk loopt.

verhogen. Een grotere financiële armslag verkleint de noodzaak om schulden te maken en maakt het mogelijk om bestaande schulden af te lossen (Houwing en Guiaux, 2015; Van Eijkel et al., 2020). Dat kan de schuldenproblematiek verkleinen. Een belangrijke kanttekening is dat schulden de prikkel om te gaan werken kunnen verminderen, omdat er beslag gelegd kan worden op het inkomen als onderdeel van een schuldregeling (Houwing en Guiaux, 2015).¹⁵

Intensievere begeleiding kan in theorie een positieve uitwerking hebben op personen met problematische schulden, als financiële problemen sneller gesignaleerd worden en mensen doorverwezen worden naar de schuldhulpverlening. Maatwerkondersteuning in de intensiveringsgroep kan een positieve uitwerking hebben op het bijhouden van administratie en op het tijdig betalen van rekeningen. Ook worden financiële problemen mogelijk eerder gesignaleerd en worden mensen sneller doorverwezen naar de schuldhulpverlening. In de ontheffingsgroep zou het tegengestelde mechanisme een rol kunnen spelen. De effecten kunnen blijvend zijn als ze gedreven worden door permanente gedragsveranderingen.

Het vrijlaten van inkomsten heeft naar verwachting geen direct effect op problematische schulden, behalve via het vinden van betaald werk. We verwachten dat de mogelijkheid om meer bij te verdienen geen effect heeft op schuldenproblematiek. De mensen die in deze groep werk vinden, zullen wel iets sneller financiële verlichting voelen door hun bijverdiensten.

4 Data

Dit hoofdstuk bespreekt in het kort de gebruikte data en de uitkomstmaten om de effecten op Wmo-gebruik, gezondheid en problematische schulden te bepalen. We bespreken de gebruikte administratieve data en de gehanteerde definities voor Wmo-gebruik, gezondheid en problematische schulden. Daarnaast bespreken we de informatie over achtergrondkenmerken en de enquêtedata die de gemeenten hebben aangeleverd.

Alle gemeenten hebben geanonimiseerde data over hun experimenten aangeleverd, die gekoppeld zijn aan informatie uit verschillende databronnen van het Centraal Bureau voor de Statistiek (CBS). Elke gemeente heeft data aangeleverd over de bijstandsgerechtigden die betrokken waren bij het experiment. Deze gegevens bevatten informatie over de toewijzing aan controle- en interventiegroepen, wanneer de interventies zijn uitgevoerd, of bijstandsgerechtigden hebben ingestemd met deelname aan het experiment en of zij zijn gestopt met een interventie. Deze data zijn door het CBS geanonimiseerd en gekoppeld aan de CBS-microdata met behulp van een persoonlijk identificatienummer (CBS RINPERSOON-nummer).

Om Wmo-gebruik te meten, combineren we informatie van gemeenten over de inzet van maatwerkvoorzieningen met gegevens over de eigen bijdrage. We gebruiken twee CBS-databestanden met gegevens over Wmo-gebruik. Het eerste bestand (*wmobus*) bevat informatie over Wmo-maatwerkvoorzieningen. Gemeenten leveren deze informatie aan het CBS. Het tweede bestand bevat gegevens over personen die een eigen bijdrage hebben betaald voor hun voorziening (*gebwmotab*). We combineren deze twee gegevensbronnen, omdat in beide bronnen een deel van de populatie kan ontbreken.¹⁶ De data uit het eerste bestand zijn beschikbaar tot en met het jaar 2021 en de data uit het tweede bestand tot en met het jaar

¹⁵ Beslaglegging op het inkomen geeft ook administratieve lasten voor de werkgever, waardoor personen met schulden minder aantrekkelijke werknemers zijn (Houwing en Guiaux, 2015).

¹⁶ Personen kunnen in de *wmobus* voorkomen en niet in de *gebwmotab*, als er geen eigen bijdrage is betaald. Personen kunnen in de *gebwmotab* voorkomen en niet in de *wmobus*, als gemeenten registraties nog niet verwerkt hebben op het moment dat zij de data naar het CBS sturen.

2019. Op basis van informatie over de begin- en einddatums en het type voorziening construeren we twee indicatorvariabelen (ja/nee), die aangeven of een persoon in een betreffende maand gebruik heeft gemaakt van respectievelijk begeleiding of dagbesteding.

Aan de hand van data over geneesmiddelenverstrekkingen definiëren we een indicatorvariabele voor de algehele gezondheid. We gebruiken het CBS-databestand *medicijntab*, dat jaarlijkse gegevens bevat over alle geneesmiddelenverstrekkingen die vergoed worden uit de verplichte basisverzekering. Het bestand is beschikbaar tot en met het jaar 2020. Zoals in een eerder CPB-onderzoek (Van Eijkel et al., 2020) definiëren we een indicatorvariabele, die aangeeft of er in het betreffende jaar minstens één keer geneesmiddelen zijn verstrekt aan de betreffende persoon.

Om mentale gezondheid te meten maken we gebruik van dezelfde data, maar deze keer beperkt tot geneesmiddelen voor psychische klachten. Het bestand *medicijntab* bevat gegevens over geneesmiddelenverstrekkingen ingedeeld naar geneesmiddelengroepen.¹⁷ We creëren een indicatorvariabele die aangeeft of een persoon in het betreffende jaar geneesmiddelen uit één of meer van de volgende vier subgroepen heeft gebruikt: antipsychotica (N05A), anxiolytica (N05B), hypnotica en sedativa (N05C), antidepressiva (N06A). De keuze voor deze vier subgroepen baseren we op eerder onderzoek naar de relatie tussen werk(loosheid) en gezondheid (Van Eijkel et al., 2020; Yildiz et al., 2020).

Als alternatieve gezondheidsmaten gebruiken we de uitgaven aan huisartsenzorg en een indicatorvariabele voor ggz-gebruik. Het CBS-databestand *zvwzorgkostentab* bevat jaarlijkse informatie over zorgkosten die verzekerd zijn via de basisverzekering. Als indicator voor de algehele gezondheid meten we de uitgaven aan huisartsenzorg in euro's.¹⁸ Om mentale gezondheid te meten, gebruiken we een indicatorvariabele die aangeeft of er in het betreffende jaar ggz-uitgaven zijn gemaakt.¹⁹ Dit gaat om kosten van zowel de generalistische basis-ggz als de specialistische ggz binnen de basisverzekering. Omdat de zorgkostendata bij het CBS alleen beschikbaar zijn tot en met het jaar 2019, richten we onze aandacht in de hoofdanalyses op de gezondheidsmaten van de voorgaande twee alinea's.

We gebruiken data over wanbetaling van de zorgpremie als proxy voor problematische schulden. De CBS-data bevatten geen informatie over schulden. Wel zijn er jaarlijkse data beschikbaar over personen die onder de wanbetalersregeling van de Zorgverzekeringswet (Zvw) vallen (*wanbzwvtab*). Personen komen in deze regeling terecht als zij meer dan zes maanden geen zorgverzekeringspremie hebben betaald. Vaak blijkt wanbetaling van de zorgpremie een signaal te zijn dat iemand in de financiële moeilijkheden zit (zie Jungmann en Werksma, 2012; Posthumus et al., 2019; Douven et al., 2019). We volgen eerder CPB-onderzoek (Van Eijkel et al., 2020; Roos et al., 2021) en gebruiken deze data als proxy voor problematische schulden. Onze uitkomstmaat is een indicatorvariabele die aangeeft of een persoon in het betreffende jaar (op peildatum 31 december) onder de wanbetalersregeling viel.²⁰

We gebruiken informatie over achtergrondkenmerken van deelnemers uit verschillende CBS-databestanden. Ten eerste verzamelen we data uit de Basisregistratie Personen (*gbapersoontab*) met informatie over leeftijd, geslacht en migratieachtergrond (vier categorieën: geen, westers, niet-westers, onbekend). Daarnaast gebruiken we gegevens over het hoogst behaalde opleidingsniveau (vier categorieën: laag,

¹⁷ De data zijn ingedeeld naar geneesmiddelengroepen op vier posities van de ATC-code. De ATC-code is een classificatiesysteem voor geneesmiddelen van de World Health Organization (WHO). Zie de WHO-website voor meer informatie over de ATC-indeling: www.whooc.no/atc.

¹⁸ Deze kosten omvatten het inschrijftarief, de kosten van consulten, eerstelijnsondersteuning en eventuele overige kosten die door de huisarts gedeclareerd zijn voor de betreffende persoon.

¹⁹ We gebruiken een indicatorvariabele in plaats van de zorgkosten in euro's, omdat zowel extreme waarden als de waarde 0 (door geen ggz-gebruik) vaak voorkomen in de data.

²⁰ In het bestand zitten geen personen die binnen een kalenderjaar de regeling in- en weer uitstromen.

middelbaar, hoog, onbekend) uit het bestand *hoogsteop1tab*. Uit het bestand *gbahuishoudensbus* gebruiken we informatie over het type huishouden (vijf categorieën: alleenstaande, alleenstaande ouder, gehuwden en samenwonenden zonder kinderen, gehuwden en samenwonenden met kinderen, overig). Ook meten we de duur van de lopende uitkering in de maanden voor de start van het experiment met data uit het bestand *bijstanduitkeringtab*. Als laatste berekenen we het cumulatieve looninkomen in de negen tot drie maanden voor de start van het experiment met gegevens uit het bestand *spolis*. Hiermee sluiten we aan op de vorige CPB-notities.

Gemeenten hebben ook enquêtedata aangeleverd die gekoppeld zijn aan de CBS-data. In alle gemeenten zijn op ten minste drie meetmomenten enquêtes afgenomen: aan het begin, halverwege en tegen het einde van het experiment. De enquêtes werden naar personen gestuurd die op dat moment deelnamen aan het experiment. Het niet invullen van de enquêtes had geen consequenties. Om deze twee redenen zijn de enquêtedata niet voor alle ingedeelde deelnemers beschikbaar (zie tabel B6.2 en B6.3 in bijlage VI voor responscijfers per gemeente). Wegens afstemming onder de betrokken partijen zijn de meeste vragen in alle gemeenten identiek. De enquêtedata zijn door het CBS geanonimiseerd en gekoppeld aan de overige databestanden. Aan de hand van vragen over ondernomen activiteiten definiëren we twee indicatorvariabelen, die dagbesteding in de enquêtedata te meten.²¹ Daarnaast gebruiken we zelfgerapporteerde informatie over ervaren gezondheid en welbevinden. De schuldsituatie brengen we in kaart met een indicatorvariabele gebaseerd op een vraag naar de financiële situatie van het huishouden. Meer informatie over de uitkomstmaten op basis van de enquêtedata is te vinden in bijlage VI.

²¹ De enquêtes bevatten geen vragen die het gebruik van begeleiding vanuit de Wmo meten of als proxy hiervoor kunnen dienen.

Prevalenties in de steekproeven, doelgroepen en bevolking

De prevalenties voor medicijngebruik, problematische schulden en Wmo-gebruik zijn deels hoger in de steekproeven en doelgroepen van de experimenten vergeleken met de Nederlandse bevolking. Onder 18-tot 65-jarigen in de Nederlandse bevolking gebruikten in 2016 rond 66% geneesmiddelen in het algemeen (inclusief voor psychische klachten) en 9% geneesmiddelen voor psychische klachten. Ongeveer 2% van de bevolking was wanbetaler van de zorgpremie, een *proxy* voor problematische schulden. Rond 1% ontving Wmo-begeleiding en ongeveer 0,3% deed mee aan dagbesteding. Tabel 4.1 toont de prevalenties voor deelnemers aan het experiment (steekproef voor de randomisatie) en de groep die in aanmerking kwam voor het experiment (doelgroep). De prevalenties onder deze bijstandsgerechtigden zijn deels hoger, wat hun kwetsbare positie onderstreept.

De groep deelnemers heeft soms een wat betere gezondheidspositie dan de doelgroep, wat verklaard kan worden door vrijwillige deelname. Van de personen die in de steekproef zitten in Nijmegen, Tilburg, Utrecht en Wageningen gebruikte in 2016 zo'n 18-27% geneesmiddelen voor psychische klachten. In de respectievelijke doelgroepen lag de prevalentie meestal enkele procentpunten hoger (23-27%), met uitzondering van Tilburg. Het is aannemelijk dat deelnemers met een betere gezondheid eerder geneigd waren om zich op te geven voor een experiment. In Deventer en Groningen werd de hele doelgroep ingedeeld in experimentgroepen, waardoor de steekproef gelijk is aan de doelgroep.

Ook de financiële situatie van deelnemers aan de experimenten is gemiddeld vaak wat beter vergeleken met de hele doelgroep. Voor Wmo-gebruik is er weinig verschil. De kans om in de wanbetalingsregeling te zitten, lag onder personen in de steekproef in 2016 gemiddeld tussen de 7 en 15%. Voor personen in de doelgroep was de kans gemiddeld enkele procentpunten hoger (9-13%), behalve in Tilburg. Het aandeel personen met begeleiding of dagbesteding vanuit de Wmo verschilt in de meeste gemeenten weinig tussen de steekproef en de doelgroep.

Tabel 4.1 Prevalenties steekproef en doelgroep in 2016, per gemeente

	Deventer	Groningen	Nijmegen	Tilburg	Utrecht	Wageningen
Steekproef						
Geneesmiddelen	78%	68%	75%	80%	77%	73%
Geneesmiddelen (psych.)	23%	17%	21%	27%	25%	18%
Wanbetaling zorgpremie	12%	9%	9%	15%	7%	7%
Wmo-begeleiding	11%	7%	11%	14%	3%	12%
Wmo-dagbesteding	3%	1%	2%	1%	1%	<1%
Doelgroep						
Geneesmiddelen	78%	68%	76%	78%	79%	74%
Geneesmiddelen (psych.)	23%	17%	25%	25%	27%	23%
Wanbetaling zorgpremie	12%	9%	11%	13%	11%	9%
Wmo-begeleiding	11%	7%	11%	11%	3%	14%
Wmo-dagbesteding	3%	1%	3%	1%	1%	2%

5 Beschrijvende statistieken en vergelijkbaarheid groepen

Om de gemeten effecten toe te kunnen schrijven aan de interventies is het belangrijk dat de controle- en interventiegroepen per gemeente goed vergelijkbaar zijn. Randomisatie zorgt ervoor dat de controle- en interventiegroepen van een experiment, zoals ontstaan bij de toewijzing, naar verwachting vergelijkbaar zijn op zowel waargenomen als niet-waargenomen kenmerken. Het kan echter gebeuren dat de randomisatie door toeval niet tot vergelijkbare groepen leidt. Dit risico is groter bij een laag aantal deelnemers per groep. Ook fouten in het randomisatieproces kunnen tot niet-vergelijkbare groepen leiden. Ten slotte kunnen er alleen al door deelname aan een experiment gedragseffecten ontstaan. Dit zijn drie mogelijke oorzaken waardoor verschillen tijdens en na het experiment niet geïnterpreteerd kunnen worden als het causale effect van de geteste interventie.

We zagen eerder al dat de experimentgroepen met betrekking tot achtergrondkenmerken in alle gemeenten doorgaans goed vergelijkbaar zijn. In eerdere notities onderzochten we de vergelijkbaarheid van de groepen op basis van persoons- en huishoudkenmerken, looninkomen voor de start van het experiment en bijstandshistorie. Uit de analyses bleek dat deze achtergrondkenmerken in de controle- en interventiegroepen in de verschillende gemeenten redelijk goed overeenkwamen, zoals je zou verwachten bij een gerandomiseerd experiment (zie paragraaf 4.2 in De Boer et al., 2020).

De effecten in drie gemeenten zijn echter niet direct toe te schrijven aan de interventies. De arbeidsdeelname in de controlegroepen in Nijmegen en Tilburg lag al voor de start van het experiment hoger dan in de interventiegroepen. In Groningen vertoonde de controlegroep een afwijkend patroon van snel stijgende uitstroom naar werk direct na de start. In de vorige notities bleek dat er in Nijmegen en Tilburg verschillen waren in de kans op betaald werk voor de start van het experiment (zie paragraaf 4.3 in De Boer et al., 2020). Daarnaast was er in Groningen kort na de start van het experiment een opvallend afwijkende trend in de controlegroep: een snel stijgende uitstroom naar werk vergeleken met alle interventiegroepen én de aselekt getrokken referentiegroep. Om deze redenen konden we de geschatte effecten op uitstroom naar betaald werk in deze drie gemeenten niet direct toeschrijven aan de interventies. Zoals beschreven in hoofdstuk 3 kunnen er verbanden zijn tussen werk en de uitkomstmaten die we in deze notitie bestuderen. De effecten op brede baten in Groningen, Nijmegen en Tilburg zijn om de hiervoor genoemde redenen waarschijnlijk ook niet direct causaal te interpreteren. Daarom baseren we onze conclusies in deze notitie op de resultaten in Utrecht, Deventer en Wageningen. De resultaten voor de overige drie gemeenten laten we voor de volledigheid zien in de bijlage IV.

Voor Wmo-gebruik zijn er ook in Deventer en Wageningen deels grote verschillen voor start, waardoor de resultaten voor deze uitkomsten vermoedelijk niet geïnterpreteerd kunnen worden als een direct effect van de interventies. Figuur 5.3 en 5.4 tonen het percentage deelnemers met Wmo-begeleiding respectievelijk -dagbesteding in Deventer, in de twee jaar voor de start van de interventies tot één jaar daarna.²² We zien dat het aandeel deelnemers met begeleiding al voor de start van het experiment lager ligt in de ontheffingsgroep met vrijlating en in de intensiveringsgroep met vrijlating ten opzichte van de controlegroep. Het verschil loopt op tot rond 4%-punt bij de start van het experiment. Ook het percentage

²² De zwarte gestippelde lijnen in de grafiek geven de start en het einde van het experiment aan. Het grijze vlak markeert de maanden waarin de eerste volledige lockdown werd ingesteld (maart 2020 tot en met juni 2020). Alle figuren tonen descriptieve uitkomsten, dat wil zeggen zonder correctie voor achtergrondkenmerken of andere verschillen tussen de groepen.

deelnemers met dagbesteding is, in de meeste maanden voor de start, lager in deze twee groepen. Figuur 5.5 en 5.6 tonen dezelfde grafische vergelijking voor Wageningen. Het percentage deelnemers met begeleiding ligt al voor de start van het experiment een stuk hoger in de ontheffingsgroep en het verschil met de controlegroep neemt verder toe vanaf tien maanden voor de start. De vrijlatingsgroep laat een vergelijkbaar patroon zien. Bij de start van het experiment is de kans op begeleiding in beide groepen rond de 9%-punt hoger. Een statistische analyse wijst uit dat de verschillen voor start in beide gemeenten merendeels ook statistisch significant zijn (zie tabel B8.1 in bijlage VIII).²³ De verschillen zorgen ervoor dat de geschatte effecten voor Wmo-gebruik in deze twee gemeenten vermoedelijk niet volledig kunnen worden toegeschreven aan de interventies. We gaan ervan uit dat de verschillen door toeval zijn ontstaan, omdat eerdere vergelijkingen van de groepen op basis van achtergrondkenmerken en arbeidsdeelname voor de start van de experimenten geen systematische verschillen hebben aangetoond (De Boer et al., 2020; Verlaat en Zulkarnain 2022).

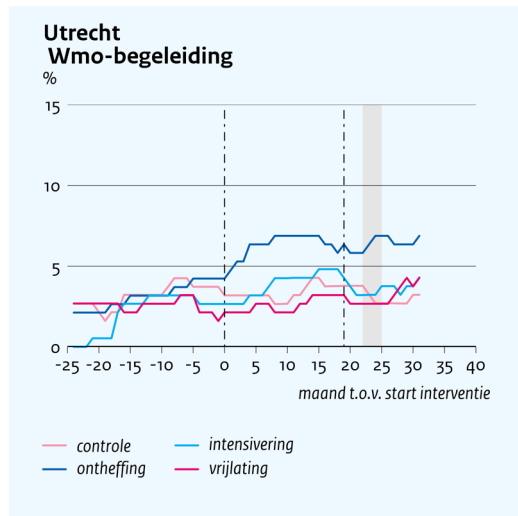
De uitkomsten voor gezondheid en problematische schulden zijn in de groepen in Utrecht, Deventer en Wageningen voor de start van het experiment redelijk goed vergelijkbaar. Alleen in Deventer zien we een afwijkende trend voor de ontheffingsgroep met vrijlating en app bij schuldenproblematiek. Figuur 5.7 tot en met 5.12 toont een grafische vergelijking voor medicijngebruik over de tijd, in het algemeen en voor psychische klachten. De verschillen tussen de controle- en interventiegroepen in de jaren 2014 tot en met 2016 zijn over het algemeen niet statistisch significant, met uitzondering voor de groep intensivering met vrijlating in Deventer (zie tabel B8.2 in bijlage VIII). In deze groep is de kans op geneesmiddelengebruik rond 8%-punt lager vergeleken met de controlegroep (zie ook figuur 5.9). Figuur 5.13 tot en met 5.15 laat dezelfde grafische vergelijking voor wanbetaling van de zorgpremie zien. De verschillen in de jaren 2014 tot en met 2016 zijn over het algemeen niet statistisch significant (zie tabel B8.4 in bijlage VIII). In Deventer valt echter op dat de ontheffingsgroep met vrijlating en app vanaf 2017 een afwijkende trend volgt, met hogere kansen op wanbetaling van de zorgpremie (zie figuur 5.14).²⁴ Een aanvullende analyse wijst uit dat dit patroon wordt bepaald door de groep deelnemers die is toegewezen aan de interventie, maar niet daadwerkelijk heeft deelgenomen (zie figuur B2.16 en B2.17 in bijlage II).²⁵ De geschatte effecten voor problematische schulden kunnen voor deze groep daarom niet direct worden toegeschreven aan de interventie.

²³ Voor deze analyse schatten we voor elke maand een regressiemodel met de uitkomstvariabele als afhankelijke variabele en indicatorvariabelen (ja/nee) voor de interventiegroepen als onafhankelijke variabelen. Wel corrigeren we voor het startmoment. We voeren deze analyse voor elke gemeente apart uit.

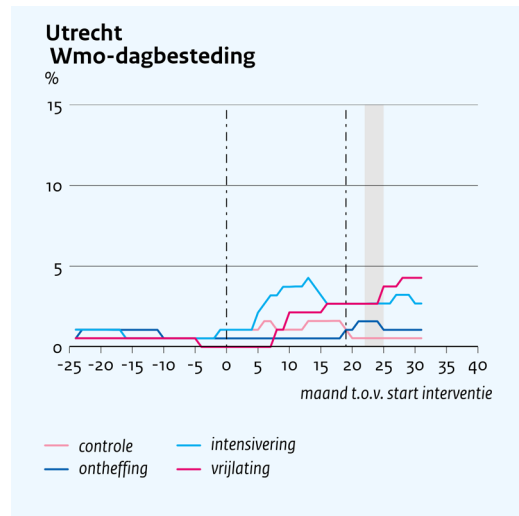
²⁴ In Deventer was er een variant waarbij de ontheffingsgroep digitale begeleiding ontving via een app (zie ook De Boer et al., 2020).

²⁵ In Deventer werden bijstandsgerechtigden eerst gerandomiseerd en vervolgens uitgenodigd voor deelname. Rond 75% van de deelnemers die aan een interventie zijn toegewezen, heeft deze niet daadwerkelijk ontvangen.

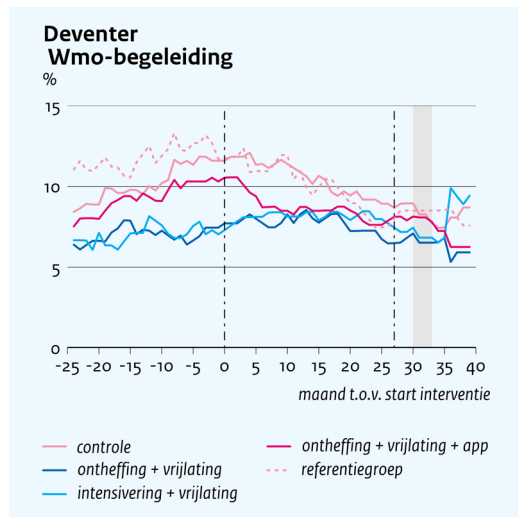
Figuur 5.1 Utrecht, Wmo-begeleiding



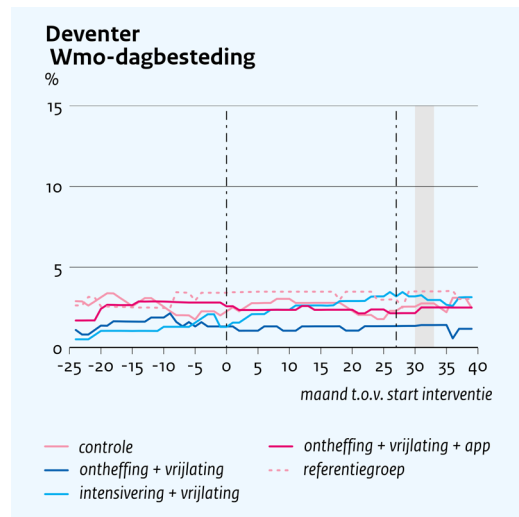
Figuur 5.2 Utrecht, Wmo-dagbesteding



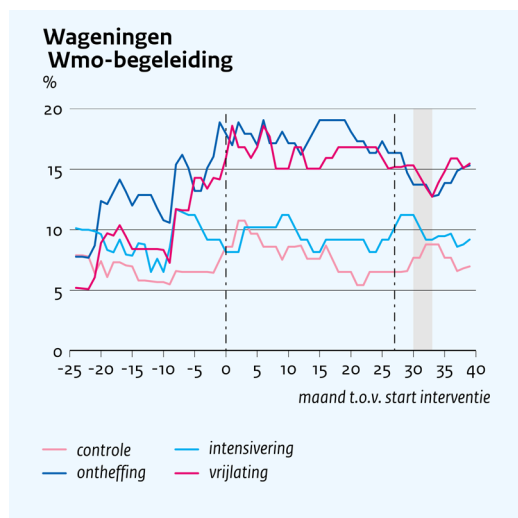
Figuur 5.3 Deventer, Wmo-begeleiding



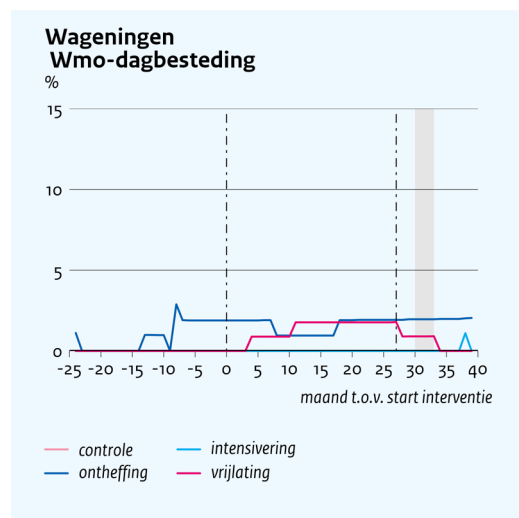
Figuur 5.4 Deventer, Wmo-dagbesteding



Figuur 5.5 Wageningen, Wmo-begeleiding

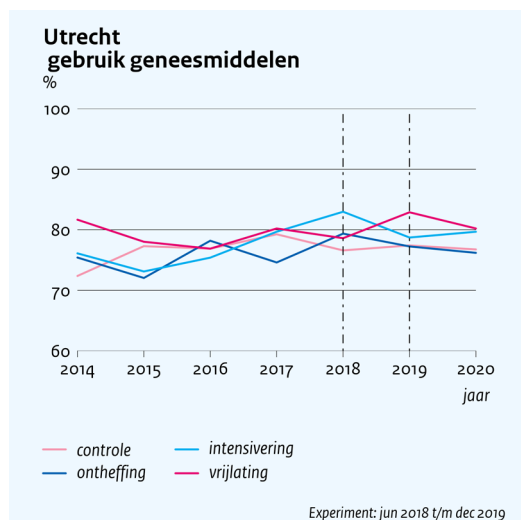


Figuur 5.6 Wageningen, Wmo-dagbesteding

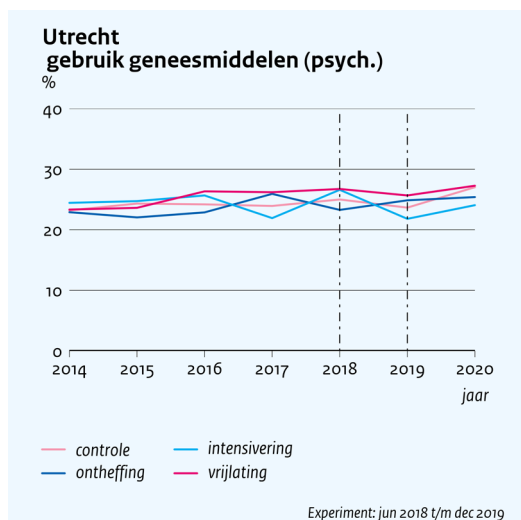


Noten: De zwarte gestippelde lijnen geven de start en het einde van de experimenten aan. Het grijze vlak markeert de maanden waarin de eerste volledige lockdown werd ingesteld (maart 2020 tot en met juni 2020).

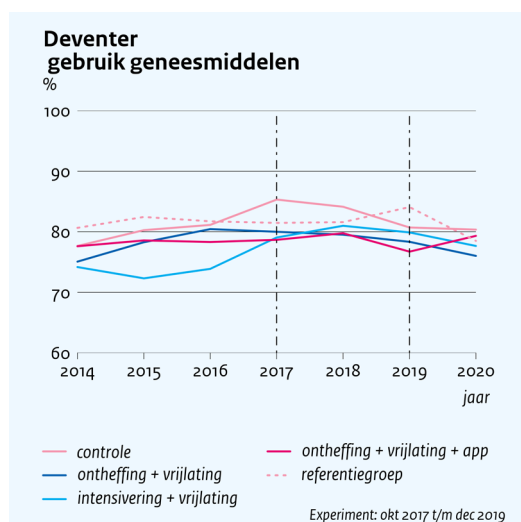
Figuur 5.7 Utrecht, gebruik geneesmiddelen



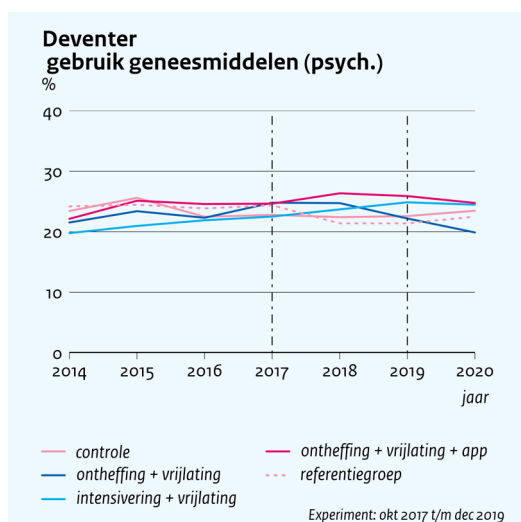
Figuur 5.8 Utrecht, gebruik geneesmiddelen (psych.)



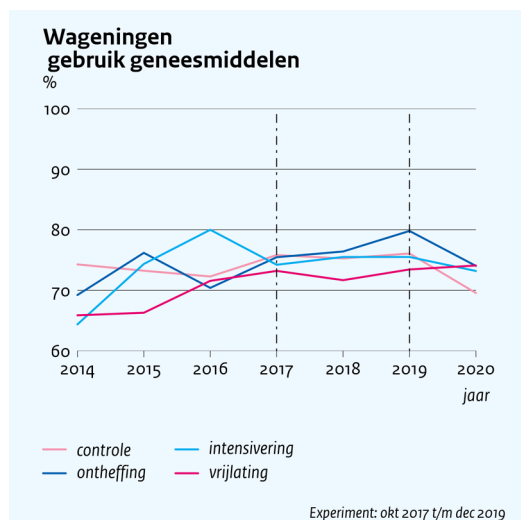
Figuur 5.9 Deventer, gebruik geneesmiddelen



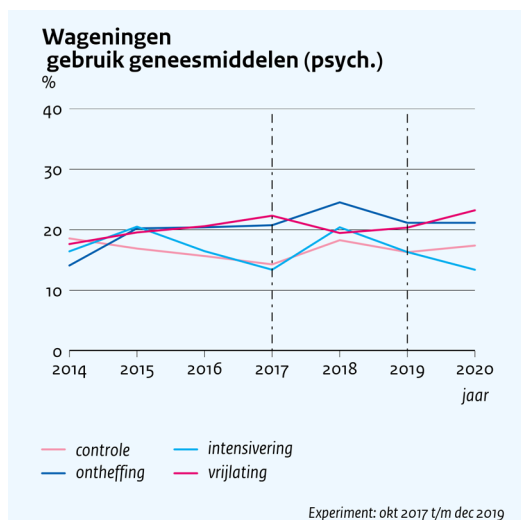
Figuur 5.10 Deventer, gebruik geneesmiddelen (psych.)



Figuur 5.11 Wageningen, gebruik geneesmiddelen

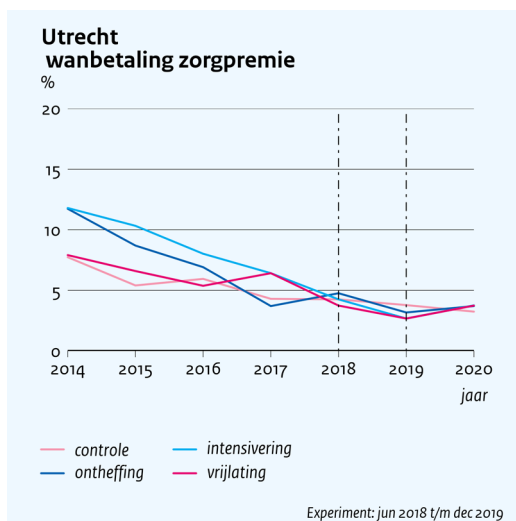


Figuur 5.12 Wageningen, gebruik geneesmiddelen (psych.)

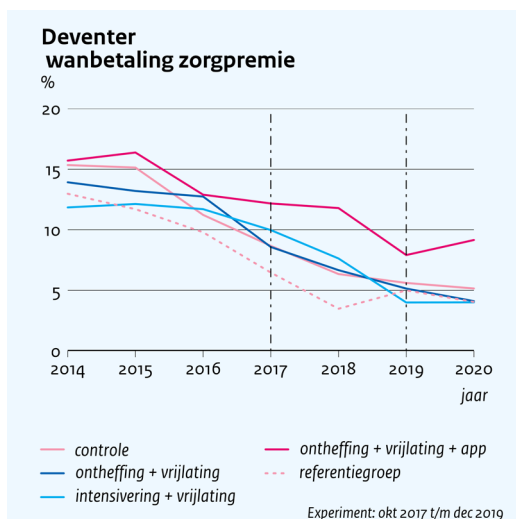


Noten: De zwarte gestippelde lijnen geven het start- en eindjaar van de experimenten aan.

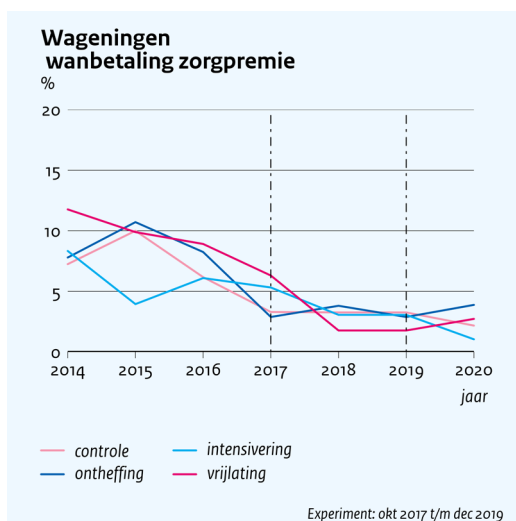
Figuur 5.13 Utrecht, wanbetaling zorgpremie



Figuur 5.14 Deventer, wanbetaling zorgpremie



Figuur 5.15 Wageningen, wanbetaling zorgpremie



Noten: De zwarte gestippelde lijnen geven het start- en eindjaar van de experimenten aan.

6 Methode

We gebruiken een *intention-to-treat* (ITT)-regressie om de effecten van de interventies op brede baten te bepalen. Dit betekent dat we de daadwerkelijke deelname aan de interventies buiten beschouwing laten en de uitkomsten vergelijken op basis van de toewijzing aan een interventiegroep of de controlegroep. Bij voldoende vergelijkbaarheid van de controle- en interventiegroepen kan het geschatte effect worden geïnterpreteerd als het causale effect van de toewijzing aan een interventie (ITT-effect). In theorie zorgt een aselechte toewijzing voor vergelijkbare groepen. Het kan echter gebeuren dat de groepen door toeval van elkaar verschillen.

De hoofdspecificatie van het regressiemodel houdt rekening met verschillen in achtergrondkenmerken en verschillen in de uitkomsten voor de start van de experimenten.²⁶ We gebruiken het volgende lineaire kansmodel om de ITT-effecten te schatten (Angrist en Pischke, 2009):

$$y_i^t = \alpha^t + \sum_{m=1}^M \beta_m^t D_{im} + X_i' \gamma + \Omega_i' \delta + q + \varepsilon, \quad (1)$$

waarbij de uitkomst y_i^t een indicatorvariabele is, die gelijk is aan 1 als persoon i (a) gebruikmaakt van Wmo-begeleiding; (b) gebruikmaakt van Wmo-dagbesteding; (c) geneesmiddelen gebruikt; (d) geneesmiddelen voor psychische klachten gebruikt; of (d) onder de wanbetalersregeling van de Zvw valt op meetmoment t . Voor Wmo-gebruik meten we uitkomsten op maandelijks niveau; t duidt dan de maand na de start van de interventie aan. Data over gezondheid en problematische schulden zijn alleen op jaarbasis beschikbaar; t staat in dat geval voor een kalenderjaar. D_{im} is een indicatorvariabele gelijk aan 1 als persoon i aan interventiegroep m is toegewezen. De vector X_i bevat persoons- en huishoudkenmerken, gemeten drie maanden voor de start van de interventie, waaronder geslacht, hoogst behaalde opleiding, migratieachtergrond, type huishouden, leeftijd, duur van de lopende uitkering en cumulatief looninkomen in de periode van negen tot drie maanden voor start van de interventie. Omdat deze kenmerken voor de start van de interventie zijn gemeten, veranderen ze niet met de tijd. We kiezen ervoor om persoons- en huishoudkenmerken op te nemen als controlevariabelen om te corrigeren voor kleine verschillen ondanks randomisatie en om de precisie van onze schattingen te vergroten. Omdat we ook rekening willen houden met mogelijke verschillen in uitkomsten voor de start van de experimenten, bevat de vector Ω_i de waarde van de uitkomstmaat op enkele meetmomenten voor de start.²⁷ Voor jaarlijkse data zijn dat de uitkomsten van 2014 tot en met 2016, voor maandelijks data de uitkomsten in maand 8, 16 en 24 voor de start van het experiment. Ten slotte bevat q de *fixed effects* voor het startkwartaal van de interventie en is ε de storingsterm.²⁸

De coëfficiënten β_m^t geven de ITT-effecten. De ITT-effecten kunnen in theorie geïnterpreteerd worden als de causale effecten van de toewijzing aan de interventie. Als iedereen de toewijzing opvolgt, is dit effect gelijk aan het causale effect van *deelname* aan de interventie. Als veel deelnemers de toewijzing niet opvolgen, en dus niet deelnemen aan de interventie, kan het effect van toewijzing en deelname ver uit elkaar liggen. Het bepalen van het causale effect van toewijzing aan de interventie is ook informatief wanneer niet iedereen de toewijzing

²⁶ We gebruiken een vergelijkbare specificatie voor analyses met de gemeentelijke enquêtedata; deze bespreken we in bijlage VI.

²⁷ In enkele gevallen is er geen informatie beschikbaar over uitkomsten voor de start van het experiment, bijvoorbeeld omdat een deelnemer toen nog in de buitenland woonde. In deze gevallen vervangen we de missende waarde met de waarde 0 en nemen we een additionele controlevariabele op die aangeeft dat de informatie niet beschikbaar is. Op deze manier raken we geen observaties kwijt door missende informatie voor controlevariabelen (McKenzie, 2012).

²⁸ We houden rekening met het startkwartaal van de interventie, omdat sommige gemeenten meerdere startmomenten kenden (zie hoofdstuk 2).

opvolgt. Dit effect geeft namelijk weer wat er zou gebeuren bij een implementatie van nieuw beleid, waarbij deelname vrijwillig is (omdat een deel bijvoorbeeld weigert deel te nemen).²⁹

Om de robuustheid van resultaten te bepalen, schatten we effecten ook op basis van alternatieve specificaties, waarin we geen controlevariabelen meenemen of alleen voor achtergrondkenmerken controleren. De voorgaande specificatie is onze hoofdspecificatie. Om de gevoeligheid van de resultaten te testen, schatten we effecten ook met twee alternatieve specificaties. In de eerste alternatieve specificatie nemen we wel controlevariabelen op voor persoons- en huishoudkenmerken, maar niet voor uitkomsten voor de start van het experiment. Deze specificatie is gelijk aan (1), maar zonder de vector Ω_i . In de tweede alternatieve specificatie laten we ook de controlevariabelen voor persoons- en huishoudkenmerken weg. De specificatie is dan gelijk aan (1), maar zonder de vectoren Ω_i en X_i . Idealiter veranderen de geschatte effecten niet bij een andere specificatie.

We vergelijken de interventiegroepen waar mogelijk ook met een aselekt getrokken groep bijstandsgerechtigden die nooit zijn uitgenodigd voor deelname aan het experiment (de zogenoemde referentiegroep). In Deventer en Groningen zijn er aselekt getrokken referentiegroepen. Zowel de controle- als de referentiegroep hebben geen interventie gehad, waardoor ze met elkaar vergelijkbaar zouden moeten zijn en we dezelfde ontwikkeling op de gemeten uitkomsten verwachten. In de vorige notities (De Boer et al., 2020; Verlaat en Zulkarnain, 2022) zagen we in Groningen verschillen tussen de controlegroep en de aselekt getrokken referentiegroep, waardoor verschillen tussen de controle- en een interventiegroep niet meer goed te interpreteren zijn als het causale effect van de interventie.³⁰ Er kan dan wel een vergelijking gemaakt worden ten opzichte van de aselekt getrokken referentiegroep. Als de controle- en referentiegroepen wel vergelijkbaar zijn, kan gekeken worden of de gevonden effecten ten opzichte van de controlegroep ook ten opzichte van de referentiegroep te zien zijn, wat wel de verwachting zou zijn.

Resultaten voor de periode na het einde van de experimenten moeten voorzichtig worden geïnterpreteerd, vanwege de coronapandemie en de terugkeer naar de reguliere dienstverlening. Er zijn twee redenen waarom de effectschattingen na het einde van de experimenten voorzichtig geïnterpreteerd moeten worden. Ten eerste keerden de interventiegroepen terug naar het reguliere regime, zijn de interventies soms geleidelijk afgebouwd en zijn er mogelijk nieuwe beleidsinitiatieven geïntroduceerd. Om die reden kunnen gemeten effecten na het einde van het experiment niet volledig worden toegewezen aan de interventies. Ten tweede kan de coronapandemie met de eerste lockdown vanaf maart 2020 de uitkomsten, die we in deze notitie analyseren, hebben beïnvloed.

Het aantal deelnemers dat we kunnen observeren neemt over de tijd af door verhuizing naar het buitenland of overlijden, maar uitval om deze redenen is beperkt. Voor personen die naar het buitenland verhuizen of overlijden, kunnen we geen uitkomsten observeren. Daarom vallen zij vanaf het moment van hun verhuizing of overlijden buiten de analyse. In de vorige notitie zagen we al dat uitval om deze redenen over het algemeen beperkt is (tussen de 1% en 2%) en gelijkmatig verdeeld over alle groepen. Als personen uitvallen om andere redenen, kunnen zij nog wel in de administratieve data gevolgd worden.

²⁹ Er kunnen echter problemen ontstaan als zich Hawthorne- of John Henry-effecten voordoen. Deze gedragseffecten kunnen in de controle- of interventiegroepen puur door de experimentele setting optreden, en niet door de interventie zelf. Deelnemers kunnen zich bijvoorbeeld anders gedragen dan normaal omdat zij weten dat ze geobserveerd worden. Dit fenomeen wordt een Hawthorne-effect genoemd. Een ander gedragseffect, het John Henry-effect, ontstaat als deelnemers in de controlegroep beter hun best doen om te compenseren voor het feit dat zij geen andere behandeling ontvangen. Bij het optreden van deze gedragseffecten is het causale effect van de interventie niet meer te onderscheiden van andere effecten (Duflo et al., 2007). Er is dan sprake van een combinatie van een causaal effect van de interventie op de interventiegroep en een neveneffect veroorzaakt door deelname aan het experiment.

³⁰ Verschillen tussen deze groepen zouden kunnen wijzen op een gedragseffect in de controlegroep door deelname aan een experimentele setting (John Henry-effect).

7 Resultaten

Dit hoofdstuk bespreekt de analyseresultaten voor Wmo-gebruik, gezondheid en problematische schulden. We bespreken de empirische resultaten per domein. In de tabellen laten we schattingen zien voor de effecten tijdens het experiment en een jaar na afloop van het experiment. Waar nodig verwijzen we ook naar effectschattingen die zijn gebaseerd op onze alternatieve specificaties. Deze presenteren we uitgebreid in bijlage III. In dit hoofdstuk bespreken we hoofdzakelijk resultaten op basis van de administratieve data. Resultaten op basis van de enquêtedata beschouwen we als aanvullende informatie, omdat deze door selectieve respons minder betrouwbaar kunnen zijn. De bijbehorende tabellen met de enquêteresultaten zijn opgenomen in bijlage VII. Voor Deventer en Groningen, waar een aselekt getrokken referentiegroep bestaat, rapporteren we ook effecten ten opzichte van deze groep in plaats van de controlegroep (zie bijlage V voor de bijhorende tabellen).

7.1 Wmo-gebruik

Grafische vergelijking

Een beschrijvende, grafische vergelijking over de tijd laat zien dat het Wmo-gebruik in Utrecht toeneemt in alle drie de interventiegroepen. In de ontheffingsgroep stijgt begeleiding vanuit de Wmo. Figuur 5.1 toont het percentage deelnemers met begeleiding vanuit de Wmo in de controle- en interventiegroepen in Utrecht in de periode vanaf twee jaar voor de start van het experiment tot één jaar daarna. Vóór de start van het experiment ontvangt zo'n 2 tot 3% van de deelnemers begeleiding. In de controlegroep blijft dit percentage gedurende de observatieperiode grotendeels constant, terwijl het aandeel in de ontheffingsgroep met de start van het experiment toeneemt. Eén jaar na de start is het percentage deelnemers met Wmo-begeleiding in de ontheffingsgroep met rond de 7% meer dan twee keer zo hoog dan in de controlegroep. Ook na het experiment blijft het verschil tussen de twee groepen bestaan. Voor de intensiverings- en vrijlatingsgroepen lijken er geen verschillen te ontstaan met de controlegroep.

In de intensiverings- en vrijlatingsgroepen neemt dagbesteding vanuit de Wmo toe. Figuur 5.2 laat het percentage deelnemers met dagbesteding vanuit de Wmo zien in de experimentgroepen in Utrecht. Het valt op dat dagbesteding voor de start van het experiment nauwelijks voorkomt. Na de start van het experiment zien we een toename in de intensiveringsgroep. Eén jaar na de start volgt zo'n 4% van de deelnemers in de intensiveringsgroep dagbesteding, vergeleken met 1% in de controlegroep. Ook in de vrijlatingsgroep stijgt het percentage deelnemers met dagbesteding: niet alleen tijdens, maar ook na het einde van het experiment. Twaalf maanden na het einde van het experiment volgt in de vrijlatingsgroep ongeveer 4% van de deelnemers dagbesteding, vergeleken met 0,5% in de controlegroep.

Effectschattingen

Tabel 7.1 rapporteert de geschatte effecten op Wmo-begeleiding (paneel A) en -dagbesteding (paneel B) in Utrecht, Deventer en Nijmegen. De tabel toont effectschattingen op drie meetmomenten: halverwege het experiment, de maand waarin de experimenten voor de meeste deelnemers eindigden en twaalf maanden later.³¹ De effectschattingen beschrijven het verschil in gemiddelde uitkomsten tussen een interventiegroep en de controlegroep in procentpunten, na correctie voor waargenomen kenmerken en verschillen in uitkomsten voor de start van het experiment. De getallen tussen haakjes tonen het relatieve effect, dat wil zeggen het effect in procenten ten opzichte van het gemiddelde in de controlegroep. Statistisch significante effecten zijn

³¹ In Utrecht eindigde het experiment voor alle deelnemers in maand 19.

zwart en voorzien van asterisken die het significantieniveau aanduiden. Niet-significante effecten zijn grijs weergegeven. Statistisch significante effecten die causaal geïnterpreteerd kunnen worden, zijn zwart en vetgedrukt. In elk paneel tonen de twee laatste rijen het aantal observaties en de gemiddelde uitkomsten in de controlegroep in procenten.

In Utrecht vinden we een hogere kans op Wmo-begeleiding tijdens het experiment van 4%-punt in de ontheffingsgroep (significant op 5%-niveau).³² Vergeleken met de controlegroep is de kans op begeleiding tweeënhalf keer zo hoog in de ontheffingsgroep. Ook twaalf maanden na het einde van het experiment is het geschatte effect positief en statistisch significant met 3,9%-punt (significant op 5%-niveau).³³ Deelnemers in Utrecht werden volgens de gemeente noch structureel gewezen op ondersteuningsmogelijkheden vanuit de Wmo, noch hiervoor aangemeld. Waarschijnlijk hebben deelnemers zichzelf dus op eigen initiatief gemeld bij het Wmo-loket of het sociale wijkteam. De intakegesprekken aan het begin van het experiment, waarin deelnemers over de aanstaande ontheffing en de vrijwilligheid van de dienstverlening werden geïnformeerd, kunnen hiervoor een aanleiding zijn geweest. Ook onder het reguliere regime zijn de sociale wijkteams verantwoordelijk voor de begeleiding van personen met een zeer grote afstand tot de arbeidsmarkt, waardoor de weg naar het Wmo-loket voor velen van hen vermoedelijk niet groot was.

In Utrecht is de kans op dagbesteding in de intensiveringsgroep 2 tot 3%-punt hoger tijdens en na het experiment (significant op 10%-niveau).³⁴ Het effect is een verhoging van 2,8%-punt gedurende het experiment en 2,3%-punt twaalf maanden na het experiment (beide significant op 10%-niveau). Doordat de kans op gebruik van dagbesteding in de controlegroep doorgaans laag is, zijn de relatieve effecten aanzienlijk, respectievelijk 255% en 460%.³⁵ In Utrecht was dagbesteding geen onderdeel van de aanpak in de intensiveringsgroep. Mogelijk komt de toename doordat klantmanagers dankzij de intensievere begeleiding beter zicht kregen op de ondersteuningsbehoeften van deelnemers. Verder blijkt uit gesprekken met de gemeente Utrecht dat klantmanagers door een lagere caseload meer tijd hadden voor samenwerking met sociale wijkteams, die verantwoordelijk zijn voor Wmo-diensten. Dit heeft de afstand tot Wmo-diensten vermoedelijk verkleind.

In de vrijlatingsgroep in Utrecht vinden we een hogere kans op dagbesteding van 4%-punt na het einde van het experiment (significant op 5%-niveau). Vergeleken met het gemiddelde in de controlegroep gaat het eveneens om een aanzienlijk effect van 800%. Voor de effecten van een ruimere vrijlating verwachtten we volgens de theorie a priori geen effect (zie hoofdstuk 3). Uit gesprekken met de gemeente kwam naar voren dat ook medewerkers voor de inkomstenverrekening getraind zijn om bij bepaalde hulpvragen door te verwijzen naar de sociale wijkteams. Mogelijk hadden deelnemers in de vrijlatingsgroep tijdens en tegen het einde van het experiment vaker contact met de afdeling inkomstenverrekening vanwege vragen over de (aflopende) bijverdienregels en zijn hierbij ook andere problemen aan bod gekomen.

In Deventer en Wageningen vinden we enkele statistisch significante effecten op Wmo-gebruik. Deze zijn echter vermoedelijk niet direct toe te schrijven aan de interventies, vanwege verschillen in de *pre-trends*. In Deventer is de kans op gebruik van dagbesteding in de groep intensivering met vrijlating statistisch

³² Naarmate het aantal interventies en uitkomstmaten toeneemt, stijgt ook het aantal statistische tests en de kans om per toeval statistisch significante effecten te vinden (ook *multiple comparison problem*). Een gebruikelijke oplossing voor dit probleem is het hanteren van striktere *p*-waarden, zodat de kans op toevallige resultaten onder de gewenste grenswaarde blijft. Als wij onze *p*-waarden dusdanig aanpassen (volgens de methode van Westfall en Young (1993)), zijn de resultaten op Wmo-gebruik in Utrecht niet langer statistisch significant. Er zijn grotere aantallen en preciezere effectschattingen nodig om met voldoende zekerheid te kunnen bepalen dat effecten ondanks het grote aantal statistische testen niet per toeval zijn ontstaan.

³³ Deze effectschatting is kleiner en niet meer statistisch significant als we effecten schatten op basis van onze alternatieve specificaties (zie tabel B3.3 in bijlage III).

³⁴ Een aanvullende analyse wijst uit dat ook de effecten van deelname aan de interventie (*local average treatment effect*) statistisch significant zijn voor Wmo-gebruik en overeenkomen met de gevonden ITT-effecten. Zie bijlage IX voor deze analyse.

³⁵ Uitgedrukt in aantallen gaat het bij deze effecten om ongeveer vijf personen.

significant hoger aan het einde van het experiment. Uit de analyse in hoofdstuk 5 bleek echter dat er al voor de start van het experiment statistisch significante verschillen waren tussen deze groep en de controlegroep. Ook in vergelijking met de aselect getrokken referentiegroep waren er al voor start verschillen te zien. In Wageningen is de kans op begeleiding statistisch significant hoger in de ontheffings- en de vrijlatingsgroepen aan het einde van het experiment. Ook voor deze twee groepen bleken de verschillen met de controlegroep niet tijdens het experiment, maar nagenoeg volledig vóór de start van het experiment te zijn ontstaan.

De resultaten in Groningen, Tilburg en Nijmegen zijn waarschijnlijk evenmin direct causaal te interpreteren, vanwege verschillen in baankansen voor de start van de experimenten (Nijmegen en Tilburg) of een afwijkend patroon in de controlegroep (Groningen). In Nijmegen heeft de ontheffingsgroep aan het einde van het experiment en daarna een grotere kans op begeleiding en dagbesteding vanuit de Wmo (zie tabel B4.1 in bijlage IV). In Tilburg heeft de intensiveringsgroep een grotere kans op begeleiding tijdens het experiment en minder kans op dagbesteding aan het einde van het experiment. In beide gemeenten bleek de controlegroep echter al voor de start van de experimenten gemiddeld genomen een betere uitgangspositie te hebben vergeleken met de interventiegroepen (zie hoofdstuk 5). In Tilburg zijn er ook voor Wmo-gebruik deels grote verschillen zichtbaar tussen de controle- en interventiegroepen voor de start van het experiment (zie figuur B2.5 in bijlage II). Dit maakt het problematisch om de effecten in Nijmegen en Tilburg te interpreteren als alleen het effect van de beoogde interventies. In Groningen bestaat er een aselect getrokken referentiegroep van deelnemers die nooit zijn uitgenodigd voor deelname aan het experiment. Als we deze groep als vergelijkingsgroep gebruiken, vinden we statistisch significante effecten van intensivering op dagbesteding tijdens en aan het einde van het experiment (zie tabel B5.1 in bijlage V).

De resultaten voor dagbesteding op basis van de enquêtedata zijn in lijn met de gevonden effecten op basis van de CBS-data. Tabel B7.1 in bijlage VII rapporteert effectschattingen voor het gebruik van dagbesteding op basis van de gemeentelijke enquêtedata. In Utrecht vinden we een positief effect van intensivering op de kans om te werken in een sociale werkvoorziening van 6,1%-punt (significant op 10%-niveau). Dit is een effect van 105% ten opzichte van de controlegroep, waar 5,8% van de deelnemers in een sociale werkvoorziening werkt. Het effect treedt op in de eerste vervolgvragenlijst, die in Utrecht tussen de acht en tien maanden na de start van het experiment is afgenomen. Het resultaat correspondeert met de gevonden effecten op basis van de CBS-data tijdens het experiment, waarvoor we maand 10 als meetmoment hadden gekozen. We vinden geen statistisch significante effecten als we een enquêtevraag gebruiken die specifiek naar aanmeldingen voor dagbesteding vraagt. Mogelijk steeg in Utrecht vooral de arbeidsmatige dagbesteding, waardoor respondenten hun activiteit eerder als sociale werkvoorziening hebben geïnterpreteerd en niet als een dagbestedingstraject. Voor Deventer en Wageningen vinden we geen statistisch significante effecten op Wmo-gebruik op basis van de enquêtedata.

Tabel 7.1 Effecten op Wmo-gebruik, Utrecht, Deventer, Wageningen

Maand	Utrecht			Deventer			Wageningen		
	10	19	31	14	27	39	14	27	39
Paneel A: Wmo-begeleiding (1/0)									
Ontheffing ^a	4,0**	2,6	3,9**	1,1	0,5	-1,4	6,3	7,3**	7,1*
	(148%)	(68%)	(122%)	(11%)	(6%)	(-11%)	(83%)	(112%)	(101%)
Intensivering ^b	1,8	0,9	1,7	1,0	1,3	1,0	-1,7	2,1	1,3
	(67%)	(24%)	(53%)	(10%)	(15%)	(11%)	(-13%)	(32%)	(19%)
Vrijlating	0,3	0,3	1,8				5,6	7,8**	8,2**
	(11%)	(8%)	(56%)				(74%)	(120%)	(117%)
Ontheffing met app				-1,6	-0,1	-1,6			
				(-10%)	(0%)	(-11%)			
Observaties	751	748	748	1566	1551	678	408	406	368
Controlegroep	2,7%	3,8%	3,2%	10,2%	8,7%	8,7%	7,6%	6,5%	7,0%
Paneel B: Wmo-dagbesteding (1/0)									
Ontheffing ^a	-0,5	0,1	0,9	-0,8	-0,3	0,1	0,1	1,0	1,1
	(0%)	(9%)	(180%)	(0%)	(0%)	(4%)			
Intensivering ^b	2,8*	1,9	2,3*	0,8	1,9*	0,8	-0,0	-0,1	-0,2
	(255%)	(173%)	(460%)	(29%)	(83%)	(32%)			
Vrijlating	1,2	1,7	4,0**				1,6	1,7	-0,0
	(109%)	(155%)	(800%)						
Ontheffing met app				-0,6	0,1	-0,5			
				(0%)	(4%)	(0%)			
Observaties	751	748	748	1566	1551	678	408	406	368
Controlegroep	1,1%	1,1%	0,5%	2,8%	2,3%	2,5%	0,0%	0,0%	0,0%
<p>Noten: Getoond zijn de <i>intention-to-treat</i>-effecten, gecorrigeerd voor achtergrondkenmerken, werk- en bijstandshistorie voor de start van het experiment, en verschillen in uitkomsten voor de start van het experiment. In gemeenten met meerdere startmomenten is ook gecontroleerd voor het startcohort. Effecten worden getoond in %-punt. Effecten in % ten opzichte van de gemiddelden in de controlegroep zijn getoond tussen haakjes. De deelnemers zijn aan een bepaald meetmoment toegekend door te selecteren op hetzelfde aantal maanden sinds de start van de interventies. De eerste kolom per gemeente toont de effecten halverwege het experiment, de tweede kolom de effecten aan het einde van het experiment en de derde kolom de effecten twaalf maanden na het einde van het experiment. Ontheffing = ontheffing van de verplichtingen in de bijstand. Intensivering = intensievere begeleiding naar werk. Vrijlating = hogere vrijlating voor bijverdiensten. ^a Ontheffing met vrijlating in Deventer. ^b Intensivering met vrijlating in Deventer. ** is significant op 5%-niveau, * is significant op 10%-niveau. Zie tabel B3.1 tot en met B3.3 in bijlage III voor de regressiecoëfficiënten met standaardfouten.</p>									

7.2 Gezondheid

In Utrecht en Deventer zijn de gevonden effecten op medicijngebruik sterk afhankelijk van de gekozen specificatie, waardoor de resultaten niet overtuigend toe te kennen zijn aan de interventies. In Wageningen vinden we geen statistisch significante effecten. Tabel 7.2 rapporteert de geschatte effecten op het gebruik van geneesmiddelen in het algemeen (paneel A) en het gebruik van geneesmiddelen voor psychische klachten (paneel B) in 2018, 2019 en 2020. In Utrecht is de kans op geneesmiddelengebruik in het algemeen in 2018 statistisch significant hoger in de intensiveringsgroep. In Deventer is het gebruik van geneesmiddelen voor psychische klachten statistisch significant hoger in 2018 in de twee ontheffingsgroepen en in 2019 in de intensiveringsgroep. Deze schattingen zijn echter alleen statistisch significant als we onze hoofdspecificatie gebruiken, waarin we ook corrigeren voor verschillen in uitkomsten voor de start van het experiment. De effectschattingen zijn kleiner en niet meer statistisch significant als we alleen voor verschillen in achtergrondkenmerken corrigeren, of als we helemaal niet corrigeren voor verschillen tussen de controle- en interventiegroepen (zie tabel B3.4 tot en met B3.6 in bijlage III). Hetzelfde geldt in Deventer voor een vergelijking met de aselect getrokken referentiegroep. De resultaten zijn door deze gevoeligheid voor de modelkeuze niet overtuigend toe te kennen aan de interventies. Voor Utrecht geldt bovendien dat het experiment in juni 2018 begon, waardoor deelnemers slechts de helft van dat jaar een andere behandeling hebben ontvangen. Dit maakt grote effecten door de interventie in 2018 minder waarschijnlijk.

De conclusies veranderen niet als we uitkomstmaten op basis van zorgkosten gebruiken. Tabel B3.7 en B3.8 in bijlage III rapporteren effectschattingen op basis van onze alternatieve gezondheidsmaatstaven, uitgaven voor de huisarts (paneel A) en ggz-gebruik (paneel B). In Utrecht en Wageningen zijn enkele schattingen statistisch significant, maar wederom alleen in onze hoofdspecificatie en niet in de overige twee specificaties. In Utrecht vallen de significante schattingen bovendien in het jaar 2018, waarin deelnemers maar de helft van het jaar een andere behandeling hebben ontvangen. In Deventer is de kans op ggz-gebruik statistisch significant hoger in de intensiveringsgroep in 2019. Een nadere analyse wijst echter uit dat dit effect wordt gedreven door de personen in de groep die de interventie niet daadwerkelijk hebben ontvangen. Het effect kan daarom niet worden toegeschreven aan de interventie.

De resultaten in Groningen, Tilburg en Nijmegen zijn waarschijnlijk niet direct causaal te interpreteren. In Groningen is de kans op geneesmiddelengebruik in het algemeen een enkele keer statistisch significant lager in de ontheffings- en vrijlatingsgroepen dan in de controlegroep (zie tabel B4.2 in bijlage IV). De kans op geneesmiddelengebruik voor psychische klachten is in 2019 in alle groepen behalve de vrijlatingsgroep statistisch significant hoger, waarbij effectschattingen doorgaans dezelfde grootte hebben. Als we aannemen dat het hebben van werk gerelateerd is aan minder psychische klachten, dan past het laatstgenoemde resultaat bij de bevinding uit de vorige notities dat baankansen met de start van het experiment alleen in de controlegroep opvallend snel stegen (zie ook hoofdstuk 5). De effecten zijn dan, net als de resultaten voor de uitstroom naar betaald werk, niet toe te schrijven aan de interventies, maar gerelateerd aan het afwijkende patroon van de controlegroep. We vinden geen statistisch significante effecten op medicijngebruik als we de interventiegroepen in Groningen vergelijken met de aselect getrokken referentiegroep (zie tabel B5.2 in bijlage V). In Tilburg is de kans op geneesmiddelengebruik voor psychische klachten een enkele keer statistisch significant hoger in de intensiveringsgroep. Maar ook voor de start van het experiment zijn er al verschillen (zie figuur B2.12 in bijlage II). Dat ondersteunt onze beslissing om effecten niet causaal te interpreteren.

Resultaten over gezondheid en welzijn op basis van de enquêtedata uiten zich niet in resultaten op basis van CBS-data. Dat kan komen door verschillen in de gebruikte uitkomstmaten, maar ook door de selectiviteit van de enquêtedata. Tabel B7.2 in bijlage VII rapporteert effectschattingen voor gezondheidsuitkomsten op basis van enquêtedata. In paneel A is de uitkomstmaat gebaseerd op een vraag

naar de ervaren algehele gezondheid. Dit is vertaald naar een schaal van 0 tot en met 10, waarbij een hoger aantal punten staat voor een betere gezondheidsbeleving. In paneel B is de uitkomstmaat het gemiddelde van twee vragen naar het ervaren welbevinden (ook op een schaal van 0 tot en met 10).³⁶ We vinden geen statistisch significante effecten in Utrecht en Deventer. In Wageningen scoren de intensiverings- en vrijlatingsgroepen in de vragenlijst aan het einde van het experiment statistisch significant lager op ervaren welbevinden. De effectschatting is -0,55 punten voor de intensiveringsgroep (significant op 10%-niveau) en -0,73 punten voor de vrijlatingsgroep (significant op 5%-niveau), een relatief effect van respectievelijk -7,5% en -10,0% vergeleken met de controlegroep. Op basis van de CBS-data vonden we geen bewijs voor gezondheidseffecten in Wageningen. Dat we op basis van de enquêtedata deels wel statistisch significante resultaten vinden, kan komen door het verschil in gemeten concepten. Terwijl de enquêtevragen de subjectieve gezondheidsbeleving meten, observeren we in de CBS-data feitelijk medicijn- en zorggebruik. Het is mogelijk dat een subjectieve gezondheidsbeleving een momentopname is, of dat die zich pas op de langere termijn uit in medicijn- en zorggebruik. Een tweede mogelijke verklaring is een selectieve respons op de vragenlijst, bijvoorbeeld als in de controlegroep vooral deelnemers met een betere gezondheidsbeleving de vragenlijst hebben ingevuld.

³⁶ De vragen richten zich op de tevredenheid met het leven en in hoeverre de respondenten de dingen die zij in het leven doen zinvol vinden. Voor meer informatie over de vragen: zie tabel B6.1 in bijlage VI.

Tabel 7.2 Effecten op medicijngebruik, Utrecht, Deventer, Wageningen

Jaar	Utrecht			Deventer			Wageningen		
	2018	2019	2020	2018	2019	2020	2018	2019	2020
Paneel A: Gebruik geneesmiddelen (1/0)									
Ontheffing ^a	3,4 (4%)	1,0 (1%)	0,3 (0%)	-3,8 (-4%)	-0,8 (0%)	-2,8 (-2%)	2,3 (3%)	4,7 (6%)	5,8 (8%)
Intensivering ^b	7,5** (10%)	2,4 (3%)	4,9 (6%)	1,5 (2%)	3,5 (4%)	1,5 (2%)	-1,5 (-1%)	-2,8 (-3%)	2,6 (4%)
Vrijlating	1,2 (2%)	4,5 (6%)	2,8 (4%)				-0,7 (0%)	-0,3 (0%)	6,9 (10%)
Ontheffing met app				-2,7 (-2%)	-2,6 (-2%)	-0,2 (0%)			
Observaties	752	750	748	1576	1561	1545	410	407	405
Controlegroep	76,6%	77,4%	76,8%	84,1%	80,7%	80,4%	75,3%	76,1%	69,6%
Paneel B: Gebruik geneesmiddelen psychische klachten (1/0)									
Ontheffing ^a	-0,9 (0%)	2,5 (11%)	-0,5 (0%)	3,5* (16%)	1,1 (5%)	-1,9 (-4%)	2,3 (13%)	2,3 (14%)	2,5 (14%)
Intensivering ^b	0,6 (2%)	-2,6 (-8%)	-3,3 (-11%)	3,1 (14%)	4,3* (19%)	2,8 (12%)	1,8 (10%)	-1,2 (-6%)	-4,7 (-23%)
Vrijlating	0,7 (3%)	1,4 (6%)	-0,2 (0%)				-3,6 (-16%)	0,1 (1%)	2,2 (13%)
Ontheffing met app				3,6* (16%)	2,4 (11%)	0,7 (3%)			
Observaties	752	750	748	1576	1561	1545	410	407	405
Controlegroep	25,0%	23,7%	27,0%	22,4%	22,6%	23,5%	18,3%	16,3%	17,4%

Noten: Getoond zijn de *intention-to-treat*-effecten, gecorrigeerd voor achtergrondkenmerken, werk- en bijstandshistorie voor de start van het experiment, en verschillen in uitkomsten voor de start van het experiment. In gemeenten met meerdere startmomenten is ook gecontroleerd voor het startcohort. Effecten worden getoond in %-punt. Effecten in % ten opzichte van de uitkomsten in de controlegroep zijn getoond tussen haakjes. Ontheffing = ontheffing van de verplichtingen in de bijstand. Intensivering = intensievere begeleiding naar werk. Vrijlating = hogere vrijlating voor bijverdiensten. ^a Ontheffing met vrijlating in Deventer. ^b Intensivering met vrijlating in Deventer. ** is significant op 5%-niveau, * is significant op 10%-niveau. Zie tabel B3.4 tot en met B3.6 in bijlage III voor de regressiecoëfficiënten met standaardfouten.

7.3 Problematische schulden

In Utrecht en Wageningen vinden we geen statistisch significante effecten van de interventies op wanbetaling van de zorgpremie. Tabel 7.3 rapporteert effectschattingen voor de kans om onder de wanbetalersregeling te vallen, een *proxy* voor problematische schulden. In Utrecht en Wageningen vinden we geen statistisch significante effecten op deze uitkomstmaat.

In Deventer laat de ontheffingsgroep met vrijlating en app een hogere kans op wanbetaling van de zorgpremie zien, maar dit effect kan niet direct worden toegeschreven aan de interventie. In Deventer is de kans om in de wanbetalersregeling te zitten rond 3,5%-punt hoger in de ontheffingsgroep met vrijlating en app dan in de controlegroep, zowel in 2018 als in 2020 (significat op 5%-niveau). Een vergelijking met de aselekt getrokken referentiegroep levert overeenkomstige effectschattingen op. Zoals we in hoofdstuk 5 beschreven, ontstaat dit effect door een afwijkende trend voor deze interventiegroep vanaf 2017 (zie figuur 5.14). Het patroon wordt gedreven door personen die zijn toegewezen aan de interventie, maar deze niet daadwerkelijk hebben ontvangen. Daarom kunnen de effecten niet direct worden toegeschreven aan de interventie.

De resultaten in Groningen, Tilburg en Nijmegen zijn vermoedelijk niet causaal te interpreteren. Tabel B4.4 in bijlage IV toont de resultaten voor Groningen, Tilburg en Nijmegen. In Groningen is de kans op wanbetaling een enkele keer statistisch significant hoger in de ontheffingsgroep. Bij een vergelijking met de aselekt getrokken referentiegroep vinden we geen statistisch significante effecten voor problematische schulden (zie tabel B5.4 in bijlage V). In Tilburg is de kans op wanbetaling een enkele keer hoger in de intensiveringsgroep. We nemen aan dat ook dit effect niet direct causaal te interpreteren is, omdat de controle- en interventiegroepen in Tilburg voor meerdere andere uitkomsten deels aanzienlijk van elkaar verschillen voor de start van het experiment.

We vinden geen bewijs voor effecten op schuldenproblematiek op basis van de enquêtedata. Tabel B7.3 in bijlage VII rapporteert effectschattingen voor de kans op het hebben van schulden op basis van de enquêtedata. De gebruikte uitkomstmaat is gebaseerd op een vraag naar de financiële situatie van het huishouden, waarbij 'Ik maak schulden' een van de antwoordopties was. De effectschattingen in Utrecht, Deventer en Wageningen zijn niet statistisch significant en leiden daarmee niet tot andere conclusies dan de analyses op basis van de CBS-data.

Tabel 7.3 Effecten op problematische schulden, Utrecht, Deventer, Wageningen

Jaar	Utrecht			Deventer			Wageningen		
	2018	2019	2020	2018	2019	2020	2018	2019	2020
Wanbetaling zorgpremie (1/0)									
Ontheffing ^a	-0,3 (0%)	-1,0 (-26%)	0,1 (3%)	-0,2 (0%)	-0,4 (0%)	-1,3 (-19%)	1,5 (45%)	0,3 (9%)	1,6 (73%)
Intensivering ^b	-1,4 (-23%)	-1,9 (-26%)	-0,4 (0%)	0,4 (6%)	-1,8 (-18%)	-1,5 (-19%)	0,2 (6%)	0,1 (3%)	-1,3 (-45%)
Vrijlating	-0,3 (0%)	-0,8 (0%)	0,9 (28%)				-1,4 (-30%)	-1,6 (-30%)	-0,0 (0%)
Ontheffing met app				3,7** (59%)	1,6 (29%)	3,5** (67%)			
Observaties	752	748	747	1567	1547	1536	408	407	401
Controlegroep	4,3%	3,8%	3,2%	6,3%	5,6%	5,2%	3,3%	3,3%	2,2%

Noten: Getoond zijn de *intention-to-treat*-effecten, gecorrigeerd voor achtergrondkenmerken, werk- en bijstandshistorie voor de start van het experiment, en verschillen in uitkomsten voor de start van het experiment. In gemeenten met meerdere startmomenten is ook gecontroleerd voor het startcohort. Effecten worden getoond in %-punt. Effecten in % ten opzichte van de uitkomsten in de controlegroep zijn getoond tussen haakjes. Ontheffing = ontheffing van de verplichtingen in de bijstand. Intensivering = intensievere begeleiding naar werk. Vrijlating = hogere vrijlating voor bijverdiensten. ^a Ontheffing met vrijlating in Deventer. ^b Intensivering met vrijlating in Deventer. ** is significant op 5%-niveau, * is significant op 10%-niveau. Zie tabel B3.9 tot en met B3.11 in bijlage III voor de regressiecoëfficiënten met standaardfouten.

8 Conclusie en discussie

Onderzoek naar brede baten kan nieuwe inzichten opleveren. Zo blijkt het experiment in Utrecht tot meer Wmo-begeleiding in de ontheffingsgroep en meer Wmo-dagbesteding in de intensiveringsgroep en de vrijlatingsgroep te hebben geleid. De ontheffingsgroep in Utrecht had een 4%-punt hogere kans op gebruik van Wmo-begeleiding tijdens het experiment. Ook was de kans op dagbesteding rond 2%-punt hoger tijdens en na het experiment in de intensiveringsgroep en 4%-punt hoger na het experiment in de vrijlatingsgroep. Dit is een belangrijke bevinding: het laat zien dat beleidswijzigingen in de bijstand niet alleen een impact kunnen hebben op uitstroom naar betaald werk, maar ook op andere sociaal-economische uitkomsten. Het is wel zo dat de effecten voortkomen uit een deel van de bijstandspopulatie, waardoor ze wellicht niet direct toepasbaar zijn voor alle bijstandsgerechtigden.

Meer Wmo-begeleiding en -dagbesteding kan wijzen op hogere kosten, maar kan ook maatschappelijke baten opleveren als mensen de nodige ondersteuning krijgen die zij onder de reguliere begeleiding niet hadden ontvangen. Deze ondersteuning kan uiteindelijk ook bijdragen aan participatie in de samenleving. Alleen mensen met een Wmo-indicatie komen in aanmerking voor Wmo-diensten, maar de resultaten suggereren dat zij deze anders niet hadden ontvangen. Deze personen staan in veel gevallen verder verwijderd van de arbeidsmarkt, omdat ze vaak te maken hebben met multiproblematiek. Wmo-begeleiding en -dagbesteding kunnen erbij helpen om deze problemen aan te pakken en op termijn de afstand tot werk te verkleinen. Daarnaast is dagbesteding ook een vorm van participatie in de samenleving. Als begeleiding vanuit de Participatiewet wordt vervangen door Wmo-diensten, gaat het (deels) om een verschuiving van de kosten naar mogelijk beter passende vormen van begeleiding. Maar ook als meer Wmo-gebruik gepaard gaat met hogere uitgaven voor de gemeenten, kunnen er maatschappelijke baten tegenover staan.

In Deventer, Nijmegen, Tilburg en Wageningen zijn de resultaten voor Wmo-gebruik waarschijnlijk niet direct toe te schrijven aan de interventies. In Deventer en Wageningen vinden we enkele statistisch significante effecten op Wmo-gebruik. Deze zijn echter niet direct toe te wijzen aan de interventies, omdat er voor de start al aanzienlijke verschillen waren tussen de controle- en interventiegroepen. In Nijmegen en Tilburg hadden we in de eerdere notities al geconstateerd dat de controle- en interventiegroepen teveel van elkaar verschillen om geschatte effecten direct aan de interventies toe te kunnen schrijven. Dit is ook het geval voor Wmo-gebruik. In Groningen heeft de intensiveringsgroep een statistisch significant hogere kans op dagbesteding, maar dat is alleen als we de aselekt getrokken referentiegroep gebruiken als vergelijkingsgroep.

De resultaten op gezondheid zijn beperkt, wat kan komen doordat tegenovergestelde mechanismen elkaar opheffen. Maar ook databeperkingen, de korte looptijd en de omvang van de interventies kunnen een rol hebben gespeeld. We vinden geen overtuigende resultaten van de interventies op gezondheid. De significante effecten die we vinden zijn gevoelig voor de modelkeuze. Een mogelijke verklaring voor de beperkte bevindingen is dat mechanismen elkaar opheffen, zoals wordt beschreven in de literatuur. Daardoor kunnen de effecten klein en moeilijk te detecteren zijn. Om kleine effecten te kunnen schatten is het van belang dat de aantallen groot genoeg zijn. Maar ook de beschikbare gezondheidsdata hebben de analyses beperkt. Deze data waren alleen op jaarbasis beschikbaar, terwijl de experimenten vaak in het midden van het jaar startten. Verder was het met deze data alleen mogelijk om geaggregeerde categorieën te onderzoeken. Ten slotte was de looptijd van de experimenten wellicht te kort om voor een groep met een grotere afstand tot de arbeidsmarkt te leiden tot robuuste effecten op uitstroom, en om in gezondheidseffecten te resulteren die zich uiten in medicijn- en zorggebruik.

De resultaten op schuldenproblematiek zijn eveneens beperkt, mogelijk ook veroorzaakt door databeperkingen. Bovendien waren de effecten op uitstroom naar betaald werk veelal klein, wat de

mogelijkheid tot het aflossen van schulden beperkt. We vinden ook geen bewijs voor effecten op wanbetaling van de zorgpremie, wat wij als *proxy* voor problematische schulden gebruiken. Gedetailleerdere data over alle schulden zou de analyse vollediger hebben gemaakt. Bovendien waren ook de gebruikte data slechts op jaarbasis beschikbaar, waardoor effecten mogelijk vertroebeld zijn. Ten slotte is het vinden van een betaalde baan een belangrijk mechanisme achter deze relatie en waren de effecten op uitstroom over het algemeen beperkt.

Meer in het algemeen geldt dat het lage aantal deelnemers in de controle- en interventiegroepen een reden kan zijn waardoor de effecten veelal niet statistisch significant verschillen van nul. Bovendien was het verschil met de uitgangssituatie mogelijk niet groot. Sommige effecten waren mogelijk klein vanwege tegengestelde mechanismen, beperkte effecten op uitstroom naar werk of omdat de looptijd te kort was om tot grote effecten te leiden. Bovendien was in sommige gemeenten het verschil met de uitgangssituatie vermoedelijk klein, met name voor deelnemers in de ontheffingsgroep. Met grotere aantallen neemt de geschatte standaardfout af en kunnen ook kleine effecten met voldoende betrouwbaarheid van nuleffecten worden onderscheiden. Hiervoor waren de aantallen deelnemers in de meeste gemeenten echter te klein. Tot slot zijn er ook grotere aantallen nodig om bij analyses van meerdere interventies en meerdere uitkomstmaten toevallige resultaten, die door het grote aantal statistische testen kunnen ontstaan (het zogenoemde *multiple comparison problem*), met voldoende zekerheid uit te sluiten.

Grotere aantallen deelnemers zorgen er ook voor dat groepen beter vergelijkbaar zijn, wat twijfels wegneemt over de causale interpretatie van effecten. In vier van de zes gemeenten bleek dat de groepen al voor de start van de experimenten verschilden in hun uitkomsten, waardoor gemeten verschillen tijdens en na het experiment niet direct aan de interventies konden worden toegeschreven. De kans op dergelijke verschillen voor de start wordt kleiner naarmate de aantallen toenemen.

Er is meer onderzoek nodig om causale effecten op brede baten beter empirisch vast te leggen. Naast grotere aantallen kan vervolgonderzoek ook gebaat zijn bij een langere looptijd en gedetailleerdere data, bijvoorbeeld gezondheidsdata op maandbasis, data over medische diagnoses of informatie over schulden. Ondanks de beperkingen laat het huidige onderzoek zien dat arbeidsmarktbeleid ook effecten kan hebben op bredere maatschappelijke baten. Er is echter meer onderzoek nodig om causale effecten op brede baten nog beter empirisch in kaart te brengen. Goed opgezette experimenten met onder andere grotere aantallen zijn hiervoor van belang (zie ook het kader 'Lessen voor toekomstige experimenten'). Daarnaast kan onderzoek naar brede baten gebaat zijn bij een langere looptijd. Voor sommige uitkomsten kost het immers tijd voordat effecten zich uiten. Dat is vooral het geval als het mechanisme via werkherovating loopt en mensen deels ver verwijderd zijn van de arbeidsmarkt. Ten slotte kunnen data met een hogere frequentie en met gedetailleerdere informatie de detecteerbaarheid van mogelijke effecten vergroten.

Literatuur

- Angrist, J. en J. Pischke, 2009, *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton: Princeton University Press.
- Betkó, J., N. Spierings, M. Gesthuizen en P. Scheepers, 2020, *Rapportage experiment Participatiewet gemeente Nijmegen*, Nijmegen: Radboud Universiteit.
- Bolhaar, J., N. Ketel en B. van der Klaauw, 2019, Job Search Periods for Welfare Applicants: Evidence from a Randomized Experiment, *American Economic Journal: Applied Economics*, vol. 11, nr. 1, pag. 92–125 ([link](#)).
- Bruhn, M. en D. McKenzie, 2009, In Pursuit of Balance: Randomization in Practice in Development Field Experiments, *American Economic Journal: Applied Economics*, vol. 1, nr. 4, pag. 200–232 ([link](#)).
- Caliendo, M., R. Mahlstedt, G.J. van den Berg en J. Vikström, 2022, Side Effects of Labor Market Policies, *The Scandinavian Journal of Economics*, vol. 124, nr. 3 ([link](#)).
- Caroli, E. en M. Godard, 2016, Does job insecurity deteriorate health?, *Health Economics*, vol. 25, nr. 2, pag. 131–147 ([link](#)).
- Chang, F.-R., 1996, Uncertainty and investment in health, *Journal of Health Economics*, vol. 15, nr. 3, pag. 369–376 ([link](#)).
- Chetty, R., M. Stepner, S. Abraham, S. Lin, B. Scuderi, N. Turner, A. Bergeron en D. Cutler, 2016, The Association Between Income and Life Expectancy in the United States, 2001-2014, *JAMA*, vol. 315, nr. 16, pag. 1750–1766 ([link](#)).
- Coutts, A.P., D. Stuckler en D.J. Cann, 2014, The Health and Wellbeing Effects of Active Labor Market Programs, in: C.L. Cooper en F.A. Huppert (red.), *Wellbeing: A Complete Reference Guide, Interventions and Policies to Enhance Wellbeing*, John Wiley & Sons vol. 6, pag. 1–18 ([link](#)).
- Creed, P.A., R.E. Hicks en M.A. Machin, 1998, Behavioural plasticity and mental health outcomes for long-term unemployed attending occupational training programmes, *Journal of Occupational and Organizational Psychology*, vol. 71, nr. 2, pag. 171–191 ([link](#)).
- Cutler, D., A. Deaton en A. Lleras-Muney, 2006, The Determinants of Mortality, *Journal of Economic Perspectives*, vol. 20, nr. 3, pag. 97–120 ([link](#)).
- Dardanoni, V. en A. Wagstaff, 1990, Uncertainty and the demand for medical care, *Journal of Health Economics*, vol. 9, nr. 1, pag. 23–38 ([link](#)).
- De Boer, H.-W., J. Bolhaar, E. Jongen en A. Zulkarnain, 2020, *Evaluatie experimenten participatiewet: Effecten op de uitstroom naar werk*, Den Haag: Centraal Planbureau ([link](#)).
- Domelen, D.R. van, A. Koster, P. Caserotti, R.J. Brychta, K.Y. Chen, J.J. McClain, R.P. Troiano, D. Berrigan en T.B. Harris, 2011, Employment and Physical Activity in the U.S., *American Journal of Preventive Medicine*, vol. 41, nr. 2, pag. 136–145 ([link](#)).
- Douven, R., A. Zeilstra, A.-F. Roos en A. Verrips, 2019, *Verbeteropties gemeentepolis*, Den Haag: Centraal Planbureau ([link](#)).
- Duflo, E., R. Glennerster en M. Kremer, 2007, Chapter 61 Using Randomization in Development Economics Research: A Toolkit, in: T.P. Schultz en J.A. Strauss (red.), *Handbook of Development Economics*, Elsevier, vol. 4, pag. 3895–3962 ([link](#)).

- Edzes, A., R. Rijnks, K. Kloosterman en V. Venhorst, 2020, *Bijstand op Maat: Beleidsrapport*, URSI-onderzoeksrapport, nr. 366, Groningen: Rijksuniversiteit Groningen, Faculteit Ruimtelijke Wetenschappen ([link](#)).
- Eijkkel, R. van, S. Gerritsen, K. Sadiraj en M. Versantvoort, 2020, *De brede baten van werk*, Den Haag: Centraal Planbureau en Sociaal en Cultureel Planbureau ([link](#)).
- Everding, J. en J. Marcus, 2020, The effect of unemployment on the smoking behavior of couples, *Health Economics*, vol. 29, nr. 2, pag. 154–170 ([link](#)).
- Fletcher, J.M., J.L. Sindelar en S. Yamaguchi, 2011, Cumulative effects of job characteristics on health, *Health Economics*, vol. 20, nr. 5, pag. 553–570 ([link](#)).
- French, S.A., C.C. Tangney, M.M. Crane, Y. Wang en B.M. Appelhans, 2019, Nutrition quality of food purchases varies by household income: the SHOPPER study, *BMC Public Health*, vol. 19, art. 231 ([link](#)).
- Frijters, P., J.P. Haisken-DeNew en M.A. Shields, 2005, The causal effect of income on health: Evidence from German reunification, *Journal of Health Economics*, vol. 24, nr. 5, pag. 997–1017 ([link](#)).
- Gibson, M., W. Hearty en P. Craig, 2020, The public health effects of interventions similar to basic income: a scoping review, *The Lancet Public Health*, vol. 5, nr. 3, pag. e165–e176 ([link](#)).
- Gramberg, P. en J. De Swart, 2020, *Wat werkt op weg naar werk? Eindrapport experiment Participatiewet gemeente Deventer*, Enschede: Saxion Hogeschool, Lectoraat Social Work ([link](#)).
- Haushofer, J. en E. Fehr, 2014, On the psychology of poverty, *Science*, vol. 344, nr. 6186, pag. 862–867 ([link](#)).
- Houwing, H. en M. Guiaux, 2015, Schuldenproblematiek onder uitkeringsgerechtigden, *UWV Kennisverslag*, vol. 2015, nr. 1, pag. 27–32 ([link](#)).
- Huber, M., M. Lechner en C. Wunsch, 2011, Does leaving welfare improve health? Evidence for Germany, *Health Economics*, vol. 20, nr. 4, pag. 484–504 ([link](#)).
- Jungmann, N. en M. Werksma, 2012, *Wanbetaling zorgkosten: voorkomen is beter dan genezen*, Zwolle: BS&F en Social Force ([link](#)).
- Klaauw, B. van der, 2014, Sociale Experimenten, ESB-blogbericht 11-3-14 ([link](#)).
- Koning, P., 2011, Experimenten in de sociale zekerheid, *ESB*, vol. 96, nr. 4605, pag. 150–153 ([link](#)).
- Machin, M.A. en P.A. Creed, 2003, Understanding the differential benefits of training for the unemployed, *Australian Journal of Psychology*, vol. 55, nr. 2, pag. 104–113 ([link](#)).
- Marcatto, F., L. Colautti, F.L. Filon, O. Luis, L. Di Blas, C. Cavallero en D. Ferrante, 2016, Work-related stress risk factors and health outcomes in public sector employees., *Safety Science*, vol. 89, nr. november, pag. 274–278 ([link](#)).
- McKenzie, D., 2012, Beyond baseline and follow-up: The case for more T in experiments, *Journal of Development Economics*, vol. 99, nr. 2, pag. 210–221 ([link](#)).
- Muffels, R., K. Blom-Stam en S. van Wanrooij, 2020a, *Vertrouwensexperiment Tilburg: werkt het en waarom wel of niet?*, Tilburg: Tilburg University/Reflect/Tranzo ([link](#)).
- Muffels, R., K. Blom-Stam en S. van Wanrooij, 2020b, *Vertrouwensexperiment Wageningen: werkt het en waarom wel of niet?*, Tilburg: Tilburg University/Reflect/Tranzo ([link](#)).
- Mullainathan, S. en E. Shafir, 2013, *Scarcity – The True Cost of Not Having Enough*, Londen: Penguin Books.
- Payne, N., F. Jones en P. Harris, 2002, The impact of working life on health behavior: the effect of job strain on the cognitive predictors of exercise., *Journal of Occupational Health Psychology*, vol. 7, nr. 4, pag. 342–353 ([link](#)).

Posthumus, H.S., S. Doove, W. Strofer en W. van Anandel, 2019, *Verkenning geregistreerde problematische schulden*, Den Haag: Centraal Bureau voor de Statistiek ([link](#)).

Puig-Barrachina, V., P. Giró, L. Artazcoz, X. Bartoll, I. Cortés-Franch, A. Fernández, P. González-Marín en C. Borrell, 2019, The impact of Active Labour Market Policies on health outcomes: a Scoping review, *European Journal of Public Health*, vol. 30, nr. 1, pag. 36–42 ([link](#)).

Ravesteijn, B., H. van Kippersluis en E. van Doorslaer, 2018, The wear and tear on health: What is the role of occupation?, *Health Economics*, vol. 27, nr. 2, pag. e69–e86 ([link](#)).

Roos, A.-F., M. Diepstraten en R. Douven, 2021, *When financials get tough, life gets rough? Problematic debts and ill health*, Den Haag: Centraal Planbureau ([link](#)).

Ruhm, C.J., 2005, Healthy living in hard times, *Journal of Health Economics*, vol. 24, nr. 2, pag. 341–363 ([link](#)).

Ruhm, C.J. en W.E. Black, 2002, Does drinking really decrease in bad times?, *Journal of Health Economics*, vol. 21, nr. 4, pag. 659–678 ([link](#)).

Saffer, H., D. Dave en M. Grossman, 2016, A Behavioral Economic Model of Alcohol Advertising and Price, *Health Economics*, vol. 25, nr. 7, pag. 816–828 ([link](#)).

Shukri, M., F. Jones en M. Conner, 2016, Work Factors, Work–Family Conflict, the Theory of Planned Behaviour and Healthy Intentions: A Cross-Cultural Study, *Stress and Health*, vol. 32, nr. 5, pag. 559–568 ([link](#)).

Verlaat, T., M. de Kruijk, S. Rosenkranz, S. Groot en M. Sanders, 2020, *Eindrapport Weten wat werkt: samen werken aan een betere bijstand*, Utrecht: Universiteit Utrecht ([link](#)).

Verlaat, T. en A. Zulkarnain, 2022, *Vervolgonderzoek experimenten Participatiewet*, Den Haag: Centraal Planbureau ([link](#)).

Westfall, P.H. en S.S. Young, 1993, *Resampling-Based Multiple Testing: Examples and Methods for p-Value Adjustment*, Hoboken: John Wiley & Sons.

Wickham, S., L. Bentley, T. Rose, M. Whitehead, D. Taylor-Robinson en B. Barr, 2020, Effects on mental health of a UK welfare reform, Universal Credit: a longitudinal controlled study, *The Lancet Public Health*, vol. 5, nr. 3, pag. e157–e164 ([link](#)).

Woolf, S.H. en P. Braveman, 2011, Where Health Disparities Begin: The Role Of Social And Economic Determinants—And Why Current Policies May Make Matters Worse, *Health Affairs*, vol. 30, nr. 10, pag. 1852–1859 ([link](#)).

Yaribeygi, H., Y. Panahi, H. Sahraei, T.P. Johnston en A. Sahebkar, 2017, The impact of stress on body function: A review, *EXCLI Journal*, vol. 16, pag. 1057–1072 ([link](#)).

Yildiz, B., M. Schuring, M. Knoef en A. Burdorf, 2020, Chronic diseases and multimorbidity among unemployed and employed persons in the Netherlands: a register-based cross-sectional study, *BMJ Open*, vol. 10, art. e035037 ([link](#)).