

# Sancties voor burgers

Een bruikbaar instrument voor gedragsverandering?

Timo Verlaat  
Erik Bijleveld  
Arne Meeldijk  
Peter Berkhout

Boomcriminologie



Wat moet een overheidsinstantie doen als een burger een regel of wet overtreedt? Is het verstandig een straf op te leggen? Of is het effectiever om – zeker bij een eerste overtreding – een waarschuwing te geven?

In dit rapport zijn 51.848 WW-ontvangers gevolgd die niet hadden voldaan aan hun verplichtingen jegens UWV. Deze mensen kregen meestal een sanctie óf een waarschuwing. Maar wat voor effecten hebben deze maatregelen eigenlijk? Dit rapport laat zien dat men niet veel hoeft te verwachten van de corrigerende effecten van uitkeringssancties. Zo stimuleerden sancties mensen weliswaar om in het vervolg beter te voldoen aan de regels van UWV, maar niet om zich extra in te spannen bij het zoeken naar werk. Waarschuwingen zorgden wél voor extra inspanning bij het zoeken naar werk.

Overheidsinstanties doen er goed aan om te reflecteren op hun gebruik van waarschuwingen en sancties. Dit rapport biedt hierbij houvast.

**Dit is een uitgave in de reeks 'Handhaving en Gedrag'.**

Handhaving en Gedrag is een interdepartementaal samenwerkingsprogramma en is opdrachtgever voor gedragswetenschappelijk onderzoek relevant voor handhaving en naleving van regelgeving. Het programma bouwt aan wetenschappelijke kennis over mechanismen die ten grondslag liggen aan naleving of overtreding van regelgeving en de wijze waarop de overheid dit gedrag kan beïnvloeden. De centrale vraag is hoe het nalevingsgedrag van burgers, bedrijven en instellingen kan worden verklaard vanuit de kenmerken en motieven van de doelgroep en de omgevingsfactoren, waaronder de handhaving. Daarmee draagt het programma bij aan de kwaliteit van de handhaving.

ISBN 978-94-6236-242-0



9 789462 362420

## Sancties voor burgers



# SANCTIES VOOR BURGERS

*Een bruikbaar instrument voor gedragsverandering?*

TIMO VERLAAT  
DR. ERIK BIJLEVELD  
ARNE MEELDIJK  
DR. PETER BERKHOUT

Op verzoek van:  
Onderzoeksprogramma Handhaving en Gedrag



Rijksoverheid

PROGRAMMA  
HANDHAVING  
EN GEDRAG

CCV

centrum voor  
criminaliteitspreventie en  
veiligheid

Boom criminologie  
Den Haag  
2021

Handhaving en Gedrag is een interdepartementaal samenwerkingsprogramma en is opdrachtgever voor gedragswetenschappelijk onderzoek relevant voor handhaving en naleving van regelgeving. Het programma bouwt aan wetenschappelijke kennis over mechanismen die ten grondslag liggen aan naleving of overtreding van regelgeving en de wijze waarop de overheid dit gedrag kan beïnvloeden. De centrale vraag is hoe het nalevingsgedrag van burgers, bedrijven en instellingen kan worden verklaard vanuit de kenmerken en motieven van de doelgroep en de omgevingsfactoren, waaronder de handhaving. Daarmee draagt het programma bij aan de kwaliteit van de handhaving.

Handhaving en Gedrag is een samenwerkingsprogramma van:

- Belastingdienst
- Inspectie Gezondheidszorg en Jeugd
- Inspectie Leefomgeving en Transport
- Inspectie Sociale Zaken en Werkgelegenheid
- Nederlandse Voedsel- en Warenautoriteit
- Centrum voor Criminaliteitspreventie en Veiligheid

Begeleidingscommissie: dhr. S. Goslinga (Belastingdienst), mw. J. Vliegen (Belastingdienst), dhr. T. Niehof (Inspectie Leefomgeving en Transport), mw. A. Mustafić (Ministerie SZW), dhr. R. Dur (Erasmus Universiteit), mw. M. Adriaanse (Centrum voor Criminaliteitspreventie en Veiligheid)

Omslagontwerp: Haagsblauw, Den Haag

Opmaak binnenwerk: Textcetera, Den Haag

© 2021 Timo Verlaat, Erik Bijleveld, Arne Meeldijk en Peter Berkhout | Boom criminologie

*Behoudens de in of krachtens de Auteurswet gestelde uitzonderingen mag niets uit deze uitgave worden vervoelvoudigd, opgeslagen in een geautomatiseerd gegevensbestand, of openbaar gemaakt, in enige vorm of op enige wijze, hetzij elektronisch, mechanisch, door fotokopieën, opnamen of enige andere manier, zonder voorafgaande schriftelijke toestemming van de uitgever.*

*Voor zover het maken van vervoelvoudingen uit deze uitgave is toegestaan op grond van artikel 16h Auteurswet of de reprorechtregeling van Stichting Reprorecht dient daarvoor een billijke vergoeding te worden voldaan aan Stichting Reprorecht (Postbus 3051, 2130 KB Hoofddorp, [www.reprorecht.nl](http://www.reprorecht.nl)). Voor het vervoelvoudigen en openbaar maken van (een) gedeelte(n) uit deze uitgave als toelichting bij het onderwijs, bijvoorbeeld in een (digitale) leeromgeving of een reader (art. 16 Auteurswet), dient een regeling te worden getroffen met Stichting Uitgeversorganisatie voor Onderwijslicenties (Postbus 3060, 2130 KB Hoofddorp, [www.stichting-uvo.nl](http://www.stichting-uvo.nl)).*

*No part of this book may be reproduced in any form, by print, photoprint, microfilm or any other means without written permission from the publisher.*

ISBN 978-94-6236-242-0

ISBN 978-90-5189-177-5 (e-book)

NUR 741

[www.boomcriminologie.nl](http://www.boomcriminologie.nl)

## DANKWOORD

We bedanken het Programma Handhaving en Gedrag voor het financieren van dit onderzoek. Daarnaast is er een aantal mensen zonder wier hulp en medewerking deze publicatie niet tot stand was gekomen. We bedanken de begeleidingscommissie hartelijk voor hun grondige feedback en hun talloze goede suggesties. De commissie bestond uit Sjoerd Goslinga (voorzitter, Belastingdienst), Miriam Adriaanse (secretaris, Centrum voor Criminaliteitspreventie en Veiligheid), Amra Mustafić (ministerie van Sociale Zaken en Werkgelegenheid), Robert Dur (Erasmus Universiteit), Jacques Niehof (Inspectie Leefomgeving en Transport) en Joyce Vliegen (Belastingdienst). Verder bedanken we het Uitvoeringsinstituut Werknemersverzekeringen (UWV) voor het faciliteren van dit onderzoek. Ook bedanken we Louise David (Radboud Universiteit) voor haar bijdrage aan het literatuuronderzoek. We hebben veel gehad aan discussies met onze directe collega's aan de Universiteit Utrecht (School of Economics) en de Radboud Universiteit (Behavioural Science Institute), waarvoor dank. Ten slotte bedanken we verschillende medewerkers van UWV die tijd hebben vrijgemaakt om te reageren op al onze vragen en verzoeken.





# INHOUD

DANKWOORD	5
SAMENVATTING	9
1 EEN GEDRAGSWETENSCHAPPELIJKE KIJK OP SANCTIES	13
1.1 Zijn sancties effectief?	14
1.2 Wat zijn de nadelen van sancties?	16
1.3 Sancties voor uitkeringsgerechtigden	22
1.4 De bijdrage van het huidige onderzoek	24
1.5 Leeswijzer	25
2 SANCTIES EN WAARSCHUWINGEN BIJ UWV	27
2.1 In welke situaties legt UWV een sanctie op?	27
2.2 Van welke sancties onderzoeken we de effecten?	28
2.3 Hoe verloopt deze procedure voor burgers?	29
3 INSTRUMENTELE-VARIABELE-ANALYSE: EEN INTRODUCTIE	33
3.1 Wat is instrumentele-variabele-analyse?	33
3.2 Strengheid als instrumentele variabele	37
3.3 Welke data zijn er nodig voor dit onderzoek?	46
4 METHODE EN DATA	49
4.1 Aanpak	49
4.2 Berekenen van de instrumentele variabelen	50
4.3 Dataset en steekproef	50
4.4 Beschrijving van de dataset	51
4.5 Uitkomstvariabelen	54
4.6 Aannames	54
5 RESULTATEN	59
5.1 Effecten na twaalf maanden	59
5.2 Effecten over de tijd	64
5.3 Effecten per opleidingsniveau	68

6	CONCLUSIES EN AANBEVELINGEN	71
6.1	Samenvatting van de belangrijkste resultaten	71
6.2	Inhoudelijke interpretatie van de resultaten	72
6.3	Reflectie op de methode	76
6.4	Aanbevelingen voor beleidsmakers	77
6.5	Conclusie	80
	LITERATUURLIJST	81
	BIJLAGE A: MODELSPECIFICATIE	85
	BIJLAGE B: AANVULLENDE FIGUREN	87
	BIJLAGE C: AANVULLENDE TABELLEN	89
	BIJLAGE D: BERICHTEN UWV	95

## SAMENVATTING

*Wat kan een overheidsinstantie het beste doen als een burger een regel of wet overtreedt? Moet de instantie een straf opleggen? Of is het effectiever om – zeker bij een eerste overtreding – een waarschuwing te geven?*

Vragen over de effectiviteit van sancties en waarschuwingen spelen veel bij beleidsmakers en overheidsinstanties. In dit rapport beschrijven we een studie naar de effecten van sancties en waarschuwingen in de context van werkloosheidsuitkeringen. We hopen hiermee een bijdrage te leveren aan de ontwikkeling van evidencebased sanctie- en waarschuwingsbeleid.

Mensen die een WW-uitkering ontvangen, zijn verplicht om vier keer per vier weken een sollicitatieactiviteit te ondernemen en te registreren. Wanneer ze niet aan deze verplichting voldoen, krijgen ze meestal een sanctie of een waarschuwing. De causale effecten van deze sancties en waarschuwingen zijn echter nog onvoldoende bekend.

In ons onderzoek gebruikten we een grote dataset van het Uitvoeringsinstituut Werknemersverzekeringen (UWV) met gegevens over WW-gerechtigden die niet voldaan hadden aan hun sollicitatieplicht ( $N = 51.848$ ). We volgden deze WW-gerechtigden nadat ze hiervoor een sanctie, dan wel een waarschuwing, hadden gekregen.

We keken naar vier potentiële gevolgen van sancties en waarschuwingen die bij uitstek belangrijk zijn in de context van werkloosheidsuitkeringen. We onderzochten of mensen (na het krijgen van een sanctie of waarschuwing):

- geneigd waren zich beter aan de regels te houden;
- intensiever op zoek gingen naar nieuw werk;
- succesvoller waren in het vinden van nieuw werk;
- minder uitkeringsafhankelijk werden.

### *Belangrijkste bevindingen*

#### **Waren mensen geneigd zich beter aan de regels te houden?**

- Sancties zorgden ervoor dat mensen zich in het vervolg beter aan de regels hielden. Dat wil zeggen: na het krijgen van een sanctie gebeurde het minder vaak dat mensen nogmaals niet aan hun sollicitatieplicht voldeden.
- Er was geen bewijs dat waarschuwingen ervoor zorgden dat mensen zich in het vervolg beter aan de regels hielden.

**Gingen mensen intensiever op zoek naar nieuw werk?**

- Er was geen bewijs dat sancties mensen motiveerden om intensiever op zoek te gaan naar nieuw werk.
- Waarschuwingen zorgden ervoor dat mensen intensiever op zoek gingen naar nieuw werk.

**Waren mensen succesvoller in het vinden van nieuw werk?**

- Er waren aanwijzingen dat sancties het inkomen uit loon verminderden. Na het ontvangen van een sanctie verdienden mensen minder door te werken, vooral in de periode tussen zes en twaalf maanden na de sanctie. De onzekerheidsmarge rond dit effect was groot. Desondanks is het zeer aannemelijk dat sancties mensen in ieder geval niet succesvoller maakten op de arbeidsmarkt.
- Er was geen bewijs dat waarschuwingen mensen meer of minder succesvol maakten op de arbeidsmarkt.

**Waren mensen minder afhankelijk van een uitkering van UWV?**

- Er was geen bewijs dat sancties en waarschuwingen effect hadden op uitkeringsafhankelijkheid.

*Conclusie*

Over het geheel genomen lijken sancties weinig effectief. Na het krijgen van een sanctie voldeden mensen weliswaar beter aan hun verplichtingen, maar er was geen bewijs dat dit kwam doordat mensen daadwerkelijk intensiever gingen solliciteren. Na het krijgen van een sanctie hadden mensen bovendien minder succes op de arbeidsmarkt (al was de onzekerheidsmarge rondom dit effect groot).

Een belangrijke kanttekening: het huidige onderzoek sluit niet uit dat sancties een ontmoedigende of afschrikwekkende werking kunnen hebben. Het is bijvoorbeeld aannemelijk dat veel mensen zich goed aan de regels houden omdat ze bang zijn dat ze anders een sanctie krijgen. De conclusie – sancties lijken weinig effectief – gaat over het effect van het *opgelegd krijgen* van een sanctie.

Waarschuwingen lijken een bruikbaar alternatief te zijn voor sancties. In tegenstelling tot sancties, zorgden waarschuwingen ervoor dat mensen meer gingen solliciteren. We vonden bovendien geen aanwijzingen dat waarschuwingen ongewenste bijeffecten hadden.

*Relatie tot economische en psychologische modellen*

Economische modellen voorspellen dat sancties mensen zouden moeten motiveren om intensiever te zoeken naar nieuw werk. Bovendien zouden sancties mensen moeten stimuleren om sneller werk te accepteren, ook als dat werk relatief weinig loon oplevert. We vonden weinig bewijs voor deze voorspellingen, al vonden we wel – zij het met een grote onzekerheidsmarge – dat sancties leidden tot een daling in inkomsten uit loon. Het zou dus kunnen dat sancties mensen

inderdaad motiveren om werk aan te nemen met minder goede arbeidsvoorwaarden.

Sommige psychologische modellen suggereren dat sancties mensen stimuleren om zich berekenend te gedragen. Ook suggereren deze modellen dat sancties de vertrouwensrelatie (tussen geveer en ontvanger van de sanctie) schaden. In zijn geheel beschouwd passen onze bevindingen goed bij deze suggesties; we vonden aanwijzingen dat sancties mensen stimuleren om te voldoen aan de regels, maar met minimale extra inspanning. We beschrijven de meest relevante economische en psychologische perspectieven in detail in hoofdstuk 1.

### *Praktische aanbevelingen*

We doen praktische aanbevelingen, die vooral relevant zijn voor overheidsinstanties die onder voorwaarden werkloosheidsuitkeringen verstrekken. Met grote voorzichtigheid kunnen onze aanbevelingen ook vertaald worden naar andere situaties waarin een overheidsinstantie onder voorwaarden uitkeringen of toeslagen verstrekt aan individuele burgers:

- Ons onderzoek laat zien dat men niet veel hoeft te verwachten van de effecten van sancties op individueel niveau. Dat wil zeggen: we vonden niet dat mensen gemotiveerder of effectiever worden nadat ze een sanctie hebben gekregen. Overheidsinstanties die desondanks sancties gebruiken, doen er daarom goed aan om te reflecteren op wat ze willen bereiken met hun sancties. Gaat het de instantie bijvoorbeeld alleen om de afschrikwekkende werking die uitgaat van sancties? Of gebruikt de instantie de sancties ook om het gedrag van mensen te corrigeren?
- Er zijn diverse ongewenste bijeffecten van sancties uit de eerdere literatuur bekend (zie hoofdstuk 1). Onze bevindingen suggereren dat deze bijeffecten redelijkerwijs te verwachten vallen in de context van sancties voor burgers. Overheidsinstanties die sanctiestructuren gebruiken, doen er goed aan om hier rekening mee te houden.
- Onze bevindingen suggereren dat WW-gerechtigden die een sanctie hadden gekregen, feitelijk twee keer te maken kregen met een verlies van inkomsten. Eerst werd er gesneden in hun WW-uitkering; vervolgens hadden ze een lager inkomen uit werk. Sancties zouden dus – op de langere termijn – mensen financieel harder kunnen raken dan dat men intuïtief zou verwachten. Overheidsinstanties die sanctiestructuren gebruiken, doen er goed aan om hierop te reflecteren.
- Waarschuwingen hebben waarschijnlijk minder negatieve bijeffecten dan sancties. Ook liet ons onderzoek zien dat waarschuwingen, in tegenstelling tot sancties, mensen stimuleerden om in het vervolg meer sollicitatieactiviteiten door te geven. Een beleid dat vooral gebruikmaakt van waarschuwingen (en dat terughoudend is met sancties) is het overwegen waard.
- Het is waarschijnlijk beter om eisen te stellen aan de zorgvuldigheid en kwaliteit van sollicitaties dan om eisen te stellen aan het aantal sollicitaties.

**Een nieuwe statistische benadering**

Een innovatief aspect van dit onderzoek is de toepassing van een *instrumentele-variabele-analyse* (IVA) in de context van uitkerings sancties. IVA is een techniek waarbij onderzoekers gebruikmaken van een reeds bestaand willekeurig proces, met het doel om – zelfs als een experimentele benadering niet mogelijk of niet opportuun is – causale uitspraken te kunnen doen ('A veroorzaakt B'). In het huidige onderzoek maken we gebruik van het feit dat WW-gerechtigden willekeurig worden toebedeeld aan UWV-medewerkers, die variëren in hun neiging om sancties en waarschuwingen op te leggen.

We concluderen dat IVA nuttig kan zijn voor analisten bij overheidsinstanties om, met behulp van bestaande registratiedata, de causale effecten van beleid te onderzoeken. We beschrijven deze techniek in detail in hoofdstuk 3; we doen methodologische aanbevelingen in hoofdstuk 6.

Veel instituties maken gebruik van sancties om menselijk gedrag te sturen. Simpel gezegd zijn sancties alle onplezierige gevolgen die (a) voortkomen uit het niet naleven van regels of wetten en (b) zijn opgelegd door derden. Denk bijvoorbeeld aan celstraffen (opgelegd door de rechtbank), schorsingen (opgelegd door de schooldirecteur), ontslagen op staande voet (opgelegd door de werkgever), boetes (opgelegd door de bibliotheek), en kortingen op uitkeringen (opgelegd door de uitkeringsinstantie). Deze voorbeelden laten zien dat maatschappelijke instituties er regelmatig voor kiezen om een sanctiestructuur te hanteren (box 1.1).

In dit hoofdstuk bieden we een gedragswetenschappelijke kijk op de effecten van sancties. Eerst beargumenteren we waarom sancties een effectief middel zijn – of beter gezegd: kunnen zijn – om menselijk gedrag te sturen. Vervolgens bespreken we hoe sancties ook minder goed kunnen werken dan gepland of gehoopt, en hoe ze zelfs onbedoelde en ongewenste effecten kunnen hebben. Ten slotte gaan we dieper in op onderzoek dat zich specifiek heeft gericht op sancties in de context van werkloosheidsuitkeringen.

### **Box 1.1 Waarom gebruiken overheden sanctiestructuren?**

Rechtsfilosofen en criminologen suggereren dat sancties, wanneer ze gebruikt worden door overheden, drie doelen kunnen dienen (Hoekendijk & Kommer, 2011):

- Overheden gebruiken sanctiestructuren om overtredingen en misdrijven te vergelden. Deze vergelding is, volgens sommige rechtsfilosofen, nodig om de balans in de rechtsorde terug te brengen na een overtreding of misdrijf. Verder geeft vergelding een signaal naar de maatschappij, namelijk het signaal dat een bepaald gedrag niet acceptabel is.
- Overheden gebruiken sanctiestructuren om potentiële daders af te schrikken. Het mes snijdt hier aan twee kanten: Enerzijds werkt het vooruitzicht van een zware straf wellicht afschrikwekkend voor de hele populatie (dit heet ook wel *algemene afschrikking*). Anderzijds zullen mensen die al een keer bestraft zijn, door het krijgen van de straf ontmoedigd worden om nog een keer de fout in te gaan (*individuele afschrikking*).
- Overheden gebruiken sanctiestructuren om daders te dwingen de veroorzaakte schade, materieel of immaterieel, te herstellen.

## 1.1 ZIJN SANCTIES EFFECTIEF?

Gezien de huidige staat van de economie en de psychologie is het niet zo gek dat zoveel instituties gebruikmaken van een sanctiestructuur. Uit deze wetenschapsgebieden kun je namelijk concluderen dat sancties – en de dreiging die uitgaat van die sancties – een sterke motiverende kracht hebben.

### **Inzichten uit de economie**

In de economie wordt aangenomen dat mensen, voordat ze een keuze maken (Zal ik koffie of thee bestellen? Zal ik de snelheidslimiet overtreden of te laat komen op mijn afspraak? Zal ik solliciteren op deze baan of nog even afwachten? Zal ik dit baanaanbod accepteren of niet?), een afweging maken van de kosten en baten van de verwachte uitkomsten van die keuzes. Het idee van de afweging van kosten en baten is ook bepalend voor hoe door economen wordt gedacht over de werking van sancties.

Een vroege en invloedrijke theorie op dat gebied komt van de Amerikaanse econoom Gary Becker (1968). Deze theorie voorspelt dat mensen drie factoren meewegen bij hun keuze om zich al dan niet onwettelijk te gedragen. Ten eerste: wat levert de overtreding op? Ten tweede: hoe hoog (of hoe kostbaar) is de straf? En ten derde: hoe hoog is de pakkans?

Zoals gebruikelijk in de economie werkte Becker vanuit de aanname dat mensen nut-maximaliserende en rationele wezens zijn. Elke wets- of regelovertreding is zodoende optimaal, volgens Becker, als de verwachte opbrengsten (van de overtreding) de kosten (van dezelfde overtreding) overstijgen. Om een simpel voorbeeld te geven: als mensen overwegen om te snel te gaan rijden, dan zullen zij (a) de tijdwinst afwegen tegen (b) de mogelijke boete en (c) de pakkans. Het samenspel van deze drie factoren zal, volgens Becker, bepalend zijn of mensen ervoor kiezen om de snelheidslimiet te overtreden.

Overheden, en andere regulerende instanties, hebben volgens het model van Becker twee knoppen waar ze aan kunnen draaien om onwettelijk gedrag te ontmoedigen: Overheden kunnen zwaardere straffen gebruiken en de pakkans proberen te verhogen. De vraag aan welke knop (vooral) gedraaid moet worden om gedrag te sturen, is onderwerp van veel (toegepaste) economische modellen en studies – en, uiteraard, van levendige politieke discussies. Onderzoek suggereert dat beide keuzes in principe effectief kunnen zijn. Een grote observati-  
onele studie uit Zwitserland laat bijvoorbeeld zien dat mensen minder verkeers-  
overtredingen begingen nadat de boetes voor dit soort overtredingen flink  
waren verhoogd (Killias, Villettaz, & Nunweiler-Hardegger, 2016). Een groot  
experiment uit Denemarken laat verder zien dat de (waargenomen) kans op een  
controle een positief effect heeft op hoeveel inkomen belastingbetalers opgeven  
bij hun belastingaangifte (Kleven, Knudsen, Kreiner, Pedersen, & Saez, 2011),  
wat suggereert dat de (waargenomen) pakkans mensen inderdaad ontmoedigt  
om belasting te ontduiken.

Het model van Becker is zodoende een bruikbaar instrument om casussen en anekdotes uit de praktijk te duiden. Een voorbeeld van zo'n casus is de bevinding (van de toenmalige Inspectie Verkeer en Waterstaat, in 2009) dat sommige



ondernemers in het goedertransport calculerend omgaan met boetes voor overtredingen van de wetgeving omtrent rij- en rusttijden. Dat wil zeggen: sommige ondernemers accepteerden de kans op boetes als een soort bedrijfsrisico en overtraden veelvuldig de wet, bijvoorbeeld door te weinig rusttijd te nemen (Inspectie Verkeer en Waterstaat, 2009). Deze praktijken, die goed verklaarbaar zijn vanuit het model van Becker, gaven aanleiding tot een verhoging van de boetes en een herziening van de inspectiemethoden.

Kortom, economische modellen geven heldere voorspellingen over hoe mensen zich zullen gedragen als ze blootgesteld worden aan sanctiestructuren. De helderheid van de voorspellingen zijn een sterke kant van deze economische benadering. Aan de andere kant is deze benadering ook enigszins beperkt, omdat economische modellen doorgaans aannemen dat mensen altijd rationeel denkende en handelende wezens zijn.

### Inzichten uit de psychologie

Net als in de economie wordt in de psychologie vaak aangenomen dat de verwachte kosten en baten van keuzes belangrijk zijn bij het begrijpen en voorspellen van beslissingen. Echter, er zijn ook belangrijke verschillen tussen de twee vakgebieden. In de psychologie ligt de nadruk op het begrijpen van de individuele mentale processen die gedrag veroorzaken. De vraag of die mentale processen rationeel zijn of niet, staat meestal wat meer op de achtergrond. Twee klassieke motivatietheorieën uit de psychologie, die allebei wat zeggen over sancties, zijn *expectancy theory* (Atkinson, 1964) en *goal setting theory* (Locke & Latham, 1990, 2019).

*Expectancy theory* veronderstelt dat de totale motivatie voor een bepaald gedrag (bijvoorbeeld: zoeken naar werk) afhangt van het product van drie factoren: *verwachting* (leidt inspanning tot goede prestatie?), *instrumentaliteit* (leidt goede prestatie tot gewenste uitkomst?) en *waarde* (is de gewenste uitkomst prettig?). Binnen dit model is het duidelijk dat een sanctie een onprettige uitkomst is. Dat wil zeggen: de *waarde* van een sanctie is heel laag. De dreiging van een sanctie voor een bepaald gedrag (bijvoorbeeld: te hard rijden) zou volgens *expectancy theory* dus moeten leiden tot een afname van de motivatie voor dat gedrag (Atkinson, 1964).

*Goal setting theory* is een model dat stelt dat de doelen die mensen voor zichzelf stellen – of de doelen die anderen hun opleggen – bepalend zijn voor prestatie. *Goal setting theory* voorspelt onder andere dat specifieke en uitdagende maar haalbare doelen het sterkste effect hebben op motivatie, vooral als mensen feedback krijgen op hun prestatie en als ze gecommitteerd zijn aan het doel. Aangezien het vooruitzicht op een sanctie (bij het niet behalen van het doel) werkt als een bron van *commitment*, suggereert *goal setting theory* dat sancties – wanneer ze gekoppeld zijn aan concrete, haalbare doelen – helpen om mensen te motiveren (Locke & Latham, 1990, 2019).

Op het grensvlak tussen de economie en de psychologie wordt de motiverende potentie van sancties ook vaak aangestipt. Een belangrijk verschijnsel dat vaak bestudeerd wordt in dit gebied is *loss aversion* (Kahneman & Tversky, 1984). Dit verschijnsel suggereert dat *losses loom larger than gains*, oftewel: potentiële

verliezen wegen zwaarder dan potentiële winsten. Met andere woorden: mensen zullen waarschijnlijk harder werken om een sanctie van € 100 te voorkomen dan om een beloning van € 100 te verkrijgen.

Kortom, er is zowel in de economie als in de psychologie een duidelijke theoretische basis voor het idee dat sancties motiverende eigenschappen hebben. De nadruk ligt in de psychologie vaak (maar niet altijd) op de afschrikkende werking van sancties: het vooruitzicht van een sanctie voor een bepaald gedrag ontmoedigt mensen om dat gedrag te vertonen.

Het basisidee dat sancties een afschrikkende werking hebben, wordt ondersteund door talloze studies die illustreren dat sancties gedrag in de beoogde, geplande richting kunnen sturen. Recente studies uit de psychologie laten bijvoorbeeld zien dat sancties de prestaties op moeilijke cognitieve taken verhogen (Choi & Cho, 2020). Ook kunnen sancties mensen helpen met het volhouden van een gezonde leefstijl (Sykes-Muskett, Prestwich, Lawton, & Armitrage, 2015). Bovendien zorgen sanctiestructuren ervoor dat mensen meer coöpereren in sociale dilemma's (Balliet, Mulder, & Van Lange, 2011).

### **Tussentijdse conclusie**

Op basis van het onderzoek dat we hebben beschreven in deze paragraaf worden beleidsmakers vaak blootgesteld aan het idee dat sancties goed werken – of in ieder geval goed zouden moeten werken. In een eerder rapport in de serie Handhaving en Gedrag valt bijvoorbeeld te lezen dat ‘een combinatie van een milde sanctie en een beloning voor het “correcte” gedrag de meest effectieve oplossing is om gedragsverandering te bewerkstelligen die beklijft’ (Van der Pligt, Koomen, & Van Harreveld, 2007, p. 38). Er lijkt dus draagvlak te zijn voor het idee dat sancties een bruikbaar middel zijn om gedragsverandering te bewerkstelligen – ook in de praktijk.

## 1.2 WAT ZIJN DE NADELEN VAN SANCTIES?

Hoewel de hierboven beschreven redenering suggereert dat sancties krachtige middelen zijn om gedrag te sturen, is er ook veel psychologisch onderzoek dat laat zien dat sancties ook ongewenste – en meestal dus ook ongeplande – neveneffecten hebben. Hier bespreken we vier typen neveneffecten: (a) sancties kunnen leiden tot *countercontrol*, (b) sancties schaden de relatie tussen de ontvanger en de oplegger van de sanctie, (c) sancties zorgen ervoor dat gedrag meer gedreven wordt door economische afwegingen en minder door sociale normen, en (d) sancties kunnen – net als beloningen – leiden tot het *criteriumprobleem*.

### **Sancties en countercontrol**

Het *behaviorisme* is een stroming binnen de psychologie die zich bezighoudt met de vraag hoe de omgeving gedrag stuurt. Hoewel er nog steeds veel onderzoek gedaan wordt binnen deze stroming, waren de hoogtijdagen van het *behaviorisme* in de periode omstreeks 1950. In die tijd deden veel psychologen onderzoek met dieren, zoals duiven en ratten, met het doel om uiteindelijk ook

menselijk gedrag beter te begrijpen. Dit onderzoek heeft veel opgeleverd: het is in die periode duidelijk geworden hoe dieren (en in mindere mate mensen) leren van beloning en straf (O'Donohue & Ferguson, 2001).

Een van de observaties die vaak gedaan werd in deze periode, is dat dieren gedrag snel aan- of afleren als er gebruik wordt gemaakt van onprettige stimuli, zoals pijnlijke elektrische schokken. Echter, een andere bevinding was dat het gebruik van dit soort onprettige stimuli leidde tot *countercontrol* (Delprato, 2002; Skinner, 1971). Dat wil zeggen: zodra er een sanctiestructuur wordt gebruikt bij dieren, laten deze dieren vaak sterk emotionele reacties zien, zoals woede, die zich vaak ook richten op de verzorger of proefleider. Ook proberen dieren relatief vaak te ontsnappen wanneer ze blootgesteld worden aan een sanctiestructuur. Kortom, het gebruik van onprettige stimuli zorgt bij dieren weliswaar voor een snel leerproces, maar ook voor ongepland en ongewenst gedrag. Er zijn duidelijke indicaties dat dit soort *countercontrol*-reacties ook voorkomt bij mensen. Een voorbeeld daarvan is een experiment uit de jaren zeventig, uitgevoerd op een psychiatrische afdeling voor militairen, die daar zaten omdat ze eerder delinquent gedrag hadden vertoond (Boren & Colman, 1970). Op deze afdeling werd een puntensysteem gehanteerd. Dat wil zeggen: bewoners van de afdeling konden punten verdienen met allerlei soorten gewenst gedrag, die ze konden inwisselen voor bepaalde privileges, zoals toegang tot de recreatieruimte, gratis koffie en weekendverlof.

Een van de gedragingen waarmee de bewoners punten konden verdienen, was door te komen opdagen bij groepsbijeenkomsten, die elke dag om 8 uur 's ochtends werden gehouden. Ondanks deze potentiële beloning was de gemiddelde aanwezigheid bij de groepsbijeenkomsten niet perfect; soms was de aanwezigheid namelijk slechts 70% of 80%. Daarom besloten de onderzoekers om ook sancties te introduceren: bewoners die niet aanwezig waren bij de groepsbijeenkomst, kregen onder de nieuwe regels puntenaftrek.

Het werd al snel duidelijk dat deze interventie niet het beoogde effect had. Ten eerste ging de aanwezigheid verder omlaag. De eerste dag was 50% van de bewoners nog aanwezig, de tweede dag kwam er helemaal niemand meer. Ten tweede werden er ook andere zorgelijke gedragingen geobserveerd door de verzorgers. Sommige bewoners spraken onder elkaar van 'rebellie' en instrueerden andere bewoners 's ochtends om in bed te blijven liggen. Ook waren er tijdens deze interventie meer gevechten en opstootjes tussen bewoners, en overtraden bewoners ook vaker andere regels. Al deze gedragingen, die goed passen bij de eerdere observaties van *countercontrol* bij dieren, gaven aanleiding om de interventie voortijdig te staken (Boren & Colman, 1970).

Kortom, hoewel de grootste successen van het *behaviorisme* behaald zijn met dieronderzoek, zijn er binnen deze stroming zeker verschijnselen ontdekt die ook betrekking hebben op mensen. Een van die verschijnselen zijn *countercontrol*-reacties. Bij sanctiestructuren bestaat het risico dat de ontvangers van de sancties zich afzetten tegen de gevers van de sancties, en zich proberen te onttrekken aan het sanctiesysteem.

### Sancties en de relatie tussen geveer en ontvanger

Een modernere onderzoeksstroming, die ligt op het grensvlak tussen de psychologie en de economie, houdt zich bezig met het bestuderen van hoe mensen in sociale dilemma's beslissingen nemen. Sociale dilemma's zijn situaties waarin mensen worden geconfronteerd met een conflict tussen hun eigenbelang en het belang van anderen. Deze onderzoeksstroming wordt gekenmerkt door experimenten, waarin onderzoekers (kleine) variaties aanbrengen in de structuur van een sociaal dilemma, en dan observeren hoe proefpersonen zich gedragen: kiezen proefpersonen voor het belang van anderen (coöperatie) of juist voor hun eigenbelang (defectie)? Binnen deze benadering is het mogelijk om de effecten van belonings- en sanctiestructuren systematisch te onderzoeken (Van Dijk & De Dreu, 2021).

Een duidelijke bevinding uit deze literatuur is dat sanctiestructuren in sociale dilemma's coöperatie in principe bevorderen. Dat wil zeggen: als mensen gestraft kunnen worden voor het overmatig prioriteren van hun eigenbelang (bijvoorbeeld door een computer, door een instantie of door andere deelnemers aan het sociale dilemma; Fehr & Gächter, 2002), dan zijn ze meer geneigd om te coöpereren. Deze conclusie past goed bij het breed gedragen idee dat de dreiging van sancties een sterke motiverende kracht heeft. Het is belangrijk om te vermelden dat deze conclusie ook wordt ondersteund door een meta-analyse (Balliet et al., 2011); dat wil zeggen, deze conclusie houdt stand als veel verschillende experimenten uit dit onderzoeksgebied (in deze analyse: 126 experimenten) samen worden geanalyseerd.

Toch blijkt ook uit onderzoek dat sanctiestructuren niet alleen maar positieve effecten hebben. Er zijn namelijk experimenten die suggereren dat sancties de relatie tussen de geveer en de ontvanger van de sanctie kunnen schaden. Vanuit het perspectief van de ontvanger van een sanctie worden sancties namelijk vaak ervaren als vijandig en oneerlijk. Zodoende leidt het opleggen van een sanctie tot schade aan het vertrouwen tussen geveer en ontvanger (Fehr & Rockenbach, 2003).

Deze schade aan het vertrouwen kan goed worden blootgelegd in het zogenoemde *removing-the-sanction*-paradigma. In dit paradigma worden proefpersonen eerst blootgesteld aan een sanctiestructuur in een sociaal dilemma (dat wil zeggen: proefpersonen krijgen een boete als ze defecteren). Deze sanctiestructuur zorgt doorgaans voor meer coöperatie (Mulder, Van Dijk, De Cremer, & Wilke, 2006). Vervolgens wordt de sanctiestructuur weer verwijderd (dat wil zeggen: proefpersonen krijgen geen boete meer als ze defecteren). Het gedrag dat proefpersonen dan vertonen is heel informatief. Immers, de sanctiestructuur zelf is weggenomen en heeft dan niet meer haar sterke motiverende prikkel; wat overblijft, zijn de neveneffecten van de sanctie. Onderzoek met dit paradigma laat zien dat sancties – nadat ze zijn weggenomen – leiden tot minder onderling vertrouwen en minder coöperatie (Mulder et al., 2006). Dit effect lijkt vooral aanwezig te zijn bij financiële sancties (zoals boetes), en minder of niet bij sociale sancties (zoals de afkeuring van andere mensen; Nelissen & Mulder, 2013). Het blijkt ook vooral aanwezig te zijn bij sanctiestructuren die gezien worden als oneerlijk (Van Prooijen, Gallucci, & Toeset, 2008). Dat laatste principe

– dat de waargenomen *fairness* van sanctiestructuren bepalend is – wordt ook ondersteund door veldonderzoek (Baldassarri & Grossman, 2011; Verboon & Van Dijke, 2011).

Kortom, de potentie voor negatieve effecten van sancties wordt ook duidelijk ondersteund door de hedendaagse literatuur. Onderzoek naar sociale dilemma's suggereert dat sanctiestructuren – en vooral als de sancties binnen die sanctiestructuren daadwerkelijk gebruikt worden (Fehr & Rockenbach, 2003) – niet goed zijn voor de vertrouwensrelatie die bestaat tussen de betrokkenen. Ook voor overheden is deze vertrouwensrelatie heel belangrijk. Mensen wier vertrouwen in de overheid geschaad is, zullen bijvoorbeeld minder eerlijk zijn bij hun belastingaangiften (Kucher & Götte, 1998) en meer weerstand hebben tegen opgelegde verplichtingen (bijvoorbeeld: een mondkapje moeten dragen in de supermarkt; Schmelz, 2021).

### **Sancties, intrinsieke motivatie en sociaal gedrag**

In de motivatiepsychologie is er een oude en invloedrijke humanistische stroming. Deze stroming streeft ernaar de omstandigheden in kaart te brengen waarin mensen opbloeien als mens. Een bekend voorbeeld van een idee uit deze traditie is de *behoeftehiërarchie* van Maslow (Maslow, 1943), die vaak als een piramide wordt weergegeven. Hoewel dit model als achterhaald (en ietwat esoterisch) wordt gezien in de moderne psychologie, zijn er wel onderzoekers die voortbouwen op Maslows werk. Een moderner humanistisch model is bijvoorbeeld *self-determination theory* (Ryan & Deci, 2000), een model dat het onderscheid tussen *intrinsieke motivatie* en *extrinsieke motivatie* op de voorgrond zet.<sup>1</sup> Humanistische psychologen waarschuwen vaak voor de negatieve effecten van extrinsieke motivatoren, zoals die gebruikt worden in belonings- en sanctiestructuren. Dat wil zeggen: deze psychologen stellen dat (a) extrinsieke motivatie veelal van een lagere kwaliteit is dan intrinsieke motivatie (Ryan & Deci, 2000), en dat (b) het gebruik van extrinsieke motivatoren intrinsieke motivatie vermindert (Deci, Koestner, & Ryan, 1999). Om deze stellingen te onderbouwen wordt onder andere vaak gewezen op het klassieke experiment van Edward Deci (1971), dat suggereerde dat een groep studenten minder plezier kreeg in een taak nadat ze voor diezelfde taak betaald hadden gekregen. Door dit experiment, en door soortgelijk onderzoek dat daarop volgde (Deci et al., 1999), ontstond het algemene idee dat extrinsieke motivatoren – beloningen en sancties – niet goed zijn voor de kwaliteit van de motivatie. Deze generalisatie werd ook heel bekend in de praktijk, onder andere door twijfelachtige (maar populaire) boeken zoals *Punished by rewards: The trouble with gold stars, incentive plans, A's*,

---

1 Als de buurjongen mijn gras maait omdat hij grasmaaien een fijne activiteit vindt, is er sprake van intrinsieke motivatie. Als de buurjongen mijn gras maait omdat ik hem daarvoor betaal, is er sprake van extrinsieke motivatie. Kortom, bij intrinsieke motivatie komt de beloning voort uit de activiteit zelf; bij extrinsieke motivatie staat de beloning los van de activiteit en komt de beloning voort uit een externe bron.

*praise, and other bribes* (Kohn, 1993) en *Drive: The surprising truth about what motivates us* (Pink, 2011).

Onderzoekers zijn de laatste jaren sterk teruggekomen van het idee dat extrinsieke motivatoren altijd onwenselijke effecten hebben. Ten eerste lijkt het zo te zijn dat beloningen niet altijd in de weg staan van psychologische behoeften, maar dat ze die behoeften juist ook kunnen bevredigen (bijvoorbeeld: beloningen kunnen dienen als een zelfbevestiging, waardoor mensen zich competentier zullen voelen; Moller & Deci, 2014). Ten tweede laat onderzoek zien dat beloningen de prestaties op complexe mentale taken verbeteren (Garbers & Konradt, 2014), inclusief taken die een bepaalde creativiteit vereisen (Byron & Khazanchi, 2012). Ten derde laat onderzoek binnen organisaties zien dat extrinsieke motivatie en intrinsieke motivatie positief samenhangen (Gerhart & Fang, 2015) – als extrinsieke motivatoren schadelijk zouden zijn voor intrinsieke motivatie, dan zou deze samenhang negatief moeten zijn. Ten vierde zijn er vormen van extrinsieke motivatie die nauwelijks te onderscheiden zijn van intrinsieke motivatie (bijvoorbeeld: de motivatie van zeer bevoegen werknemers; Gerhart & Fang, 2015).

Hoewel extrinsieke motivatie nu dus niet meer zo'n slechte naam heeft als in de afgelopen decennia, zijn er bevindingen die erop duiden dat belonings- en sanctiestructuren ervoor zorgen dat de aard van de motivatie verandert. Een aanwijzing hiervoor komt van een bekende studie die is gedaan bij een kinderdagverblijf in Israël (Gneezy & Rustichini, 2000). Bij dit kinderdagverblijf speelde het probleem dat ouders hun kind vaak te laat ophaalden aan het einde van de dag, waardoor de leiders en leidsters moesten overwerken. Om dit probleem te verminderen werd er een sanctiestructuur ingevoerd waarbij ouders elke keer dat ze te laat waren een kleine boete moesten betalen. Tegen de verwachting in bleek dat de ouders niet minder vaak, maar juist vaker te laat kwamen na invoering van de sanctiestructuur.

Een gangbare interpretatie van deze Israëliëse bevinding is als volgt: Op tijd komen was aanvankelijk iets wat de ouders gewoon deden, gedreven door sociale normen en uit respect naar de leiders en leidsters. Door de invoering van de sanctiestructuur werd de beslissing om op tijd te komen echter onderdeel van een koude economische afweging: het betalen van de boete werd een manier om meer tijd te kunnen kopen. De eerdere fatsoensnormen raakten daardoor op de achtergrond (Gneezy, Meier, & Rey-Biel, 2011; zie ook Imada et al., 2021; Schmitz, Bijleveld, & Veling, 2020).

Kortom, hoewel dat vroeger vaak betoogd werd, is het niet zo dat extrinsieke motivatoren altijd schadelijk zijn voor de intrinsieke motivatie van mensen. Toch heeft het invoeren van beloningen en sancties gevolgen voor de drijfveren van gedrag. Als er potentiële beloningen en sancties worden gebruikt, wordt menselijk gedrag schijnbaar meer gedreven door economische afwegingen en minder door sociale normen.

### Sancties en het criteriumprobleem

Elke instantie die overweegt om een sanctiestructuur (of een beloningsstructuur) te gebruiken, krijgt vroeg of laat te maken met het *criteriumprobleem*. Het criteriumprobleem is een klassiek concept uit de personeelspsychologie (Austin & Villanova, 1992).

Het idee is dat organisaties vaak willen proberen de totale bijdrage van werknemers aan de organisatie – oftewel de *prestatie* van werknemers – in kaart te brengen, bijvoorbeeld om de werknemers vervolgens navenant te belonen. Echter, de totale bijdrage van werknemers aan de organisatie is niet goed te kwantificeren. Immers, deze totale bijdrage heeft veel verschillende aspecten. Werknemers voeren bijvoorbeeld vaak verschillende taken uit en hebben allemaal een unieke set competenties. Bovendien dragen werknemers vaak over een langere tijd bij aan een organisatie en variëren hun takenpakket en hun competenties over tijd. Kortom, als organisatie zul je moeten accepteren dat de totale bijdrage aan de organisatie – ook wel het *ultieme criterium* – abstract en ongrijpbaar is (Thorndike, 1949). Het ultieme criterium kan daardoor nooit direct worden gemeten.

Wat organisaties wel kunnen doen, is het ultieme criterium zo goed mogelijk proberen af te dekken door middel van *operationele criteria*. Operationele criteria zijn meetbare aspecten van prestatie (denk aan: hoeveel omzet een verkoper draait, hoeveel declarabele uren een adviseur maakt, hoeveel artikelen een wetenschapper publiceert, hoe vaak een fabrieksarbeider te laat komt, hoe vaak een uitkeringsgerechtigde per maand solliciteert). Echter, omdat het ultieme criterium zo abstract is, terwijl operationele criteria juist heel concreet moeten zijn, is het kiezen van een goede set operationele criteria niet eenvoudig. Sterker nog: dit gaat heel vaak mis. De moeilijkheid van het kiezen van optimale operationele criteria wordt het *criteriumprobleem* genoemd.

De gevolgen van niet-optimale operationele criteria komen vaak achteraf aan het licht, wanneer er ongepland en ongewenst gedrag wordt vastgesteld. Dit ongeplande en ongewenste gedrag kan verschillende vormen aannemen:

- Soms blijkt dat mensen met minimale inspanning een operationeel criterium kunnen behalen. Denk bijvoorbeeld aan een chirurg wiens bonus afhangt van de hoeveelheid succesvolle operaties die hij of zij per jaar uitvoert. Deze chirurg kan zijn of haar bonus verhogen door, waar mogelijk, risicovolle operaties op collega's af te schuiven (Muller, 2018). Denk bijvoorbeeld ook aan een wetenschapper die promotie kan maken door tien artikelen per jaar te publiceren. Deze wetenschapper kan zijn kansen op promotie verhogen door vooral veel flinterdunne artikeltjes te produceren (Edwards & Roy, 2017). In deze gevallen spreekt men ook wel van *creaming* ('afromen'). Het is denkbaar dat *creaming* ook relevant is in de context van de WW-uitkering. Immers, het zou goed kunnen dat sommige WW-gerechtigden met minimale inspanning aan de sollicitatieplicht voldoen.
- Soms blijkt dat mensen een operationeel criterium kunnen behalen door regels of wetten te overtreden. Een voorbeeld: politici in de Verenigde Staten nemen vaak aan dat scholen beter worden als leerkrachten beloond worden voor het geven van goed onderwijs. Als een gevolg van deze aanname zijn er

scholen waar leerkrachten meer betaald krijgen als hun leerlingen hoger scoren op (gestandaardiseerde) examens. Deze beloningsstructuur heeft geleid tot grote fraudeschandalen, waarbij leerkrachten de fouten van hun leerlingen verbeterden voordat de examens beoordeeld worden (Saultz, Murphy, & Aronson, 2016).

Bij al deze voorbeelden geldt dat het operationele criterium niet sterk gerelateerd lijkt aan het ultieme criterium (goede medische zorg ≠ veel succesvolle operaties; goede wetenschap ≠ veel publicaties; kwalitatief goed werkgedrag ≠ aantal sollicitaties die de deur uit gaan; goed onderwijs ≠ goed scorende leerlingen). In de economie en in het publieke debat spreekt men bij dit soort anekdotes vaak van *perverse beloningen*. Dat wil zeggen: de beloning (of de sanctie) stuurt duidelijk het gedrag richting een operationeel criterium; in die zin is de beloning (of de sanctie) dus 'effectief'. Echter, achteraf blijkt ook soms dat het gedrag zeker niet bijdraagt aan het vervullen van het ultieme criterium. Hoewel deze anekdotes heel verschillend zijn, zijn ze allemaal terug te voeren naar het klassieke criteriumprobleem (Kerr, 1975; Muller, 2018)

### 1.3 SANCTIES VOOR UITKERINGSGERECHTIGDEN

Tot zover hebben we gezien dat (de dreiging van) sancties een sterk motiverend effect hebben. Aan de andere kant hebben we ook gezien dat sancties op verschillende manieren ongewenste effecten kunnen hebben. Dit roept de vraag op: Wat moeten uitkeringsinstanties, zoals UWV, doen? Moeten ze überhaupt gebruikmaken van sancties? Of moeten ze sancties gebruiken, maar daarbij proberen de ongewenste effecten zo veel mogelijk te voorkomen?

Een belangrijk economisch argument voor het gebruik van sancties door uitkeringsinstanties is dat mensen simpelweg een prikkel nodig hebben om werk te hervatten. Immers, omdat de lasten voor de WW-uitkering door anderen worden gedragen (namelijk door mensen die nog werken), is er sprake van een zogenoemde *moral hazard*. Dat wil zeggen: voor mensen met een WW-uitkering kan het aantrekkelijk zijn om de overgang terug naar werk te vertragen, om langer van de WW-inkomsten en van de vrije tijd te genieten. Sancties – en de dreiging die van sancties uitgaat – zijn, volgens economen, nodig om de effecten van deze *moral hazard* tegen te gaan.

Economische modellen die toegespitst zijn op uitkeringssancties, geven meer specifieke voorspellingen over hoe sancties voor uitkeringsgerechtigden werken (zie Fredriksson & Holmlund, 2006, voor een overzicht). Ook hier wordt er onderscheid gemaakt tussen twee typen effecten van sanctiesystemen (zie box 1.1):

- Ten eerste zijn er de effecten die komen door de *dreigende werking* van uitkeringssancties (ook wel: *ex ante effecten* of *algemene afschrikking*). Deze dreiging van een sanctie betekent in feite dat er extra *kosten* komen te zitten aan het leveren van weinig inspanning bij het zoeken naar werk. Als een gevolg van deze extra kosten zorgt de dreiging van een sanctie ervoor dat mensen



harder zullen werken om nieuw werk te vinden en, daardoor, sneller weer aan het werk zullen zijn. Empirisch bewijs voor de dreigende werking van uitkeringssancties komt uit Zwitserland; onderzoek laat zien dat uitkeringsgerechtigden sneller weer aan het werk gaan in regio's waar de kans op een sanctie groter is (Arni, Lalive, & Van Ours, 2013; Lalive, Van Ours, & Zweimüller, 2005).

- Ten tweede zijn er de effecten die komen door het daadwerkelijk *opgelegd krijgen* van een sanctie (ook wel: *ex post effecten* of *individuele afschrikking*). Wanneer mensen gekort worden op hun uitkering, verliezen ze een deel van hun inkomen. Ze kunnen dat verlies compenseren door weer aan het werk te gaan. Volgens economische modellen zijn er twee strategieën die mensen hier gebruiken: (a) harder zoeken naar werk, en (b) hun *reserveloön* (dat wil zeggen: het laagste loon waarvoor ze bereid zijn om te werken) naar beneden bijstellen. Met andere woorden: economische modellen voorspellen dat het ontvangen van een sanctie leidt tot harder zoeken naar werk en het eerder accepteren van werk, ook als dat werk een lager loon oplevert. Ook voor de *ex post effecten* van uitkeringssancties is er empirisch bewijs; studies uit verschillende Europese landen laten zien dat mensen die een sanctie hebben gekregen, eerder weer aan het werk gaan en minder gebruikmaken van hun recht op een uitkering (Abbring, Van den Berg, & Van Ours, 2005; Müller & Steiner, 2008; Svarer, 2011; Van den Berg, Van der Klaauw, & Van Ours, 2004; Van der Klaauw & Van Ours, 2013).

Een belangrijk economisch argument tegen het gebruik van sancties door uitkeringsinstanties is dat eerder onderzoek suggereert dat uitkeringssancties ongewenste en ongeplande effecten kunnen hebben. Studies onder Engelse en Zwitserse uitkeringsgerechtigden laten bijvoorbeeld zien dat strenger sanctiebeleid weliswaar samenhangt met een hogere uitstroom naar werk, maar ook met meer criminaliteit (Machin & Marie, 2006), lager loon (Van den Berg & Vikström, 2014) en minder zekere arbeidsovereenkomsten (Arni et al., 2013). Deze bevindingen passen goed bij het eerdergenoemde psychologische onderzoek dat de verschillende ongeplande bijeffecten van sancties heeft gedocumenteerd. Het lijkt dus goed mogelijk dat ongewenste en ongeplande effecten ook een realiteit zijn in de context van uitkeringen.

Een compromis tussen deze voor- en tegenargumenten zou kunnen zijn dat het voor uitkeringsinstanties optimaal is om wél het recht te hebben om te sanctioneren, maar dat recht zo min mogelijk te gebruiken. Zo wordt er gebruikgemaakt van de motiverende kracht van sancties (vooral *ex ante effecten*), terwijl de nadelige neveneffecten van het daadwerkelijk toepassen van de sanctie (vooral *ex post effecten*) worden geminimaliseerd. Dit compromis past goed bij een bekend idee uit de criminologie en bestuurskunde, namelijk het idee dat handhavers effectiever kunnen werken als mensen weten dat deze handhavers de macht hebben om zware sancties op te leggen. Met andere woorden: 'Regulatory agencies will be able to speak more softly when they are carrying big sticks' (Ayres & Braithwaite, 1992, p. 19; zie ook Vlieg, 2017). Je zou zodoende kunnen verwachten dat uitkeringsinstanties – mits ze het recht hebben om

zwaar te sanctioneren – het meest effectief zijn als ze terughoudend zijn met het daadwerkelijk opleggen van sancties. Wellicht zouden ze beter gebruik kunnen maken van formele waarschuwingen.

Eerder onderzoek in de reeks Handhaving en Gedrag laat inderdaad zien dat waarschuwingen een aantal voordelen hebben ten opzichte van sancties. Ouwerkerk, Van Harreveld en Van der Pligt (2009) vonden dat, vergeleken met sancties, waarschuwingen leiden tot minder boosheid, meer acceptatie, en sterkere intenties tot toekomstige naleving. Deze studies waren echter beperkt omdat ze gebruikmaakten van fictieve scenario's en steekproeven van universitaire studenten. Mulder, Van der Vegt en Ponsioen (2014) deden een soortgelijke bevinding in een steekproef van ondernemers die ervaring hadden met sancties van toezichthouders. Echter, deze studie was ook beperkt vanwege het crosssectionele surveydesign; op basis van deze studie valt weinig te concluderen over de causale effecten van sancties en waarschuwingen.

Kortom, het is nog maar de vraag wat voor effecten de sancties (en de waarschuwingen) van UWV hebben op het gedrag van WW-gerechtigden. Zorgt het voor gewenste effecten, of zijn er ongewenste bijeffecten? Welke van deze typen effecten krijgt de overhand? Het huidige onderzoek poogt deze vragen te beantwoorden, waarbij we ons vooral richten op de effecten van het daadwerkelijk gebruiken van sancties (ook wel: *ex post effecten of individuele afschrikking*).

#### 1.4 DE BIJDRAGE VAN HET HUIDIGE ONDERZOEK

De bijdrage van het huidige onderzoek aan de bestaande kennis is als volgt:

- Ten eerste maken wij in dit onderzoek gebruik van een *instrumentele-variabele-analyse*, een statistische techniek die gebruikt kan worden om causale effecten te schatten. Deze techniek is nog niet eerder ingezet in onderzoek naar sancties door uitkeringsinstanties, wat betekent dat onze studie methodologisch gezien een stap voorwaarts zet (zie hoofdstuk 3).
- Ten tweede analyseren we in dit onderzoek de causale effecten van zowel sancties als waarschuwingen. We kunnen zodoende de richting en de sterkte van de effecten van deze twee beleidsopties goed vergelijken.
- Ten derde kijken we in dit onderzoek naar een groot aantal potentiële gevolgen van sancties en waarschuwingen. Zo kijken we naar werkzoekgedrag, uitkeringslasten, werkhervatting, inkomsten uit loon, het aantal gewerkte uren, en naleving. Bovendien verkennen we hoe deze uitkomstmaten met de tijd veranderen. Door deze brede benadering te kiezen krijgen we een breder zicht op eventuele ongeplande en ongewenste effecten van sancties en waarschuwingen.

- Ten vierde maken we gebruik van een dataset die veel groter en veel representatiever is dan de datasets die gebruikt zijn in eerdere studies.

## 1.5 LEESWIJZER

De rest van dit rapport ziet er als volgt uit:

- In hoofdstuk 2 bespreken we de beleidscontext waarin we het huidige onderzoek hebben uitgevoerd. In dit hoofdstuk leggen we uit hoe UWV gebruikmaakt van waarschuwingen en sancties. Dit doen we door te beschrijven welke processen voorafgaan aan de beslissing om een waarschuwing of een sanctie op te leggen.
- In hoofdstuk 3 geven we een uitleg van *instrumentele-variabele-analyse*, de statistische benadering die we in ons onderzoek gebruiken. Deze uitleg is geschreven om ook toegankelijk te zijn voor lezers die niet dagelijks met statistiek in aanraking komen.
- In hoofdstuk 4 beschrijven we onze dataset en hoe we die verkregen hebben. Bovendien lichten we onze methodologische aanpak verder toe.
- In hoofdstuk 5 bespreken we onze belangrijkste bevindingen. In dit hoofdstuk rapporteren we eerst de causale effecten van het krijgen van een waarschuwing of sanctie op verschillende uitkomsten (naleving, werkzoekgedrag, werkhervatting, inkomsten uit loon, het aantal gewerkte uren, uitkeringslasten). We verkennen vervolgens hoe die effecten zich over tijd ontwikkelen in de twaalf maanden die volgen op de waarschuwing of sanctie. Ten slotte verkennen we of (en, zo ja, hoe) de effecten van waarschuwingen en sancties anders zijn voor mensen met een verschillend opleidingsniveau.
- In hoofdstuk 6 beschrijven we onze conclusies en aanbevelingen.



In dit hoofdstuk gaan we in op de casus die ten grondslag ligt aan het uitgevoerde onderzoek. Eerst beschrijven we de situaties waarin UWV sancties oplegt. Dan gaan we in op de specifieke sanctie waarvan we het effect onderzoeken, en beschrijven we de procedures die UWV gebruikt wanneer deze sanctie wordt opgelegd.

### 2.1 IN WELKE SITUATIES LEGT UWV EEN SANCTIE OP?

UWV legt in verschillende situaties financiële maatregelen op. Deze maatregelen kunnen twee vormen aannemen. Ten eerste kunnen WW-gerechtigden een korting op hun uitkering krijgen (een sanctie); ten tweede kunnen WW-gerechtigden een aanvullende betalingsverplichting krijgen (een boete). In de volgende alinea's beschrijven we twee formele regels die UWV het recht geven om sancties of boetes op te leggen: de inlichtingenplicht en de inspanningsplicht.

#### **De inlichtingenplicht**

Elke WW-gerechtigde heeft de verplichting om veranderingen in zijn of haar situatie tijdig en volledig aan UWV door te geven. Het gaat daarbij om wijzigingen die van invloed kunnen zijn op de hoogte van de uitkering, zoals verandering in inkomsten, vrijwilligerswerk, vakantie, verhuizing, huwelijk, echtscheiding of ziekte. Wanneer WW-gerechtigden de wijziging niet tijdig doorgeven, kunnen zij ten onrechte te veel uitkering ontvangen. Wanneer blijkt dat dit is voorgekomen, moeten WW-gerechtigden het te veel ontvangen bedrag terugbetalen. Daarbovenop kan een sanctie of boete worden opgelegd. In 2019 zijn er 6.600 sancties en 4.800 boetes opgelegd voor overtredingen van de inlichtingenplicht. Aangezien er in 2019 in totaal ongeveer 248.410 WW-gerechtigden<sup>2</sup> waren, kreeg 4,5% van de WW-gerechtigden met een sanctie of boete te maken vanwege het niet nakomen van de inlichtingenplicht. Een boete kan oplopen tot 100% van het ten onrechte ontvangen bedrag.

---

2 Bron: CBS StatLine, Personen met WW-uitkering, peildatum: januari 2019.

### De inspanningsplicht

WW-gerechtigden met arbeidsvermogen zijn verplicht zich voldoende in te spannen om zo snel mogelijk weer werk te vinden. Dit houdt bijvoorbeeld in dat zij zich moeten inschrijven als werkzoekende, beschikbaar moeten zijn voor arbeid, voldoende moeten solliciteren, en zich moeten houden aan de afspraken die ze maken met de adviseurs van UWV. Wie zich onvoldoende inspant, kan worden gekort op de uitkering. In 2019 zijn er in totaal 34.000 sancties opgelegd voor overtreding van deze plicht. Op het totale aantal WW-gerechtigden van ongeveer 248.410 illustreren deze cijfers dat gemiddeld 15,5% van de WW-gerechtigden met een sanctie te maken kreeg vanwege het niet nakomen van de inspanningsplicht. Onder overtredingen van de inspanningsplicht vallen:

- te laat aanvragen van de uitkering (8.600 sancties);
- verwijtbaar werkloos zijn (8.300 sancties);
- onvoldoende inspanning om passend werk te vinden (7.200 sancties);
- niet melden van feiten en omstandigheden (3.700 sancties);
- overtredingen van de controlevoorschriften (4.000 sancties);
- en overige (2.200 sancties).

Afhankelijk van het soort overtreding, en de ernst van de overtreding, is de hoogte van de sanctie minimaal 2% van de uitkering voor een maand lang. De hoogte van de sanctie kan oplopen tot 100% van het uitkeringsbedrag over meer dan vier maanden. De sanctie voor WW-gerechtigden die onvoldoende hun best doen om passend werk te vinden bestaat normaliter uit een korting van 25% op de uitkeringshoogte voor een duur van vier maanden.<sup>3</sup>

## 2.2 VAN WELKE SANCTIES ONDERZOEKEN WE DE EFFECTEN?

In dit onderzoek richten we ons op het effect van sancties die worden opgelegd bij een specifieke overtreding van de inspanningsplicht, namelijk bij onvoldoende inspanning om passend werk te vinden, oftewel het niet nakomen van de sollicitatieplicht. Het gaat hierbij strikt genomen om twee vormen van overtreding die gesanctioneerd kunnen worden. Zo kan een sanctie worden opgelegd wanneer WW-gerechtigden *onvoldoende sollicitatieactiviteiten uitvoeren*, en kan een sanctie worden opgelegd wanneer WW-gerechtigden wel voldoende activiteiten uitvoeren, maar deze activiteiten *niet (tijdig) doorgeven*. De hoogte en duur van de bijbehorende korting verschillen, maar de sancties komen door middel van dezelfde procedure tot stand. In dit onderzoek richten we ons op deze twee sancties om de volgende redenen:

- De sollicitatieplicht heeft een motiverende doelstelling, die geëvalueerd kan worden.
- De sollicitatieplicht geldt voor iedereen met een WW-uitkering.

---

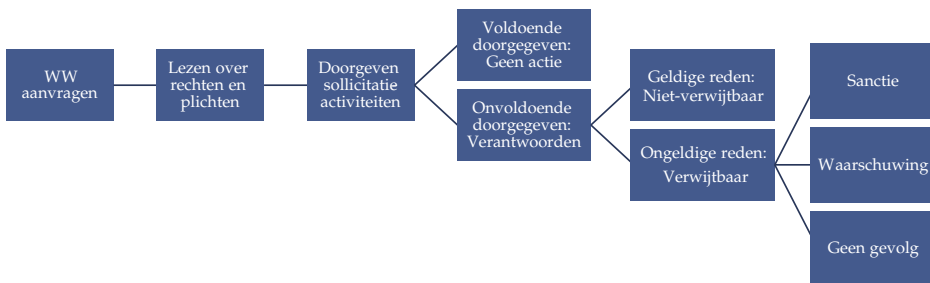
<sup>3</sup> De regels omtrent de hoogte en duur van maatregelen zijn vastgelegd in het Maatregelenbesluit socialezekerheidswetten.

- Er is enig empirisch bewijs uit andere landen en studies dat sancties in deze context deels werken, maar dat er ook ongewenste bijeffecten kunnen optreden.
- WW-gerechtigden worden op willekeurige wijze toegekend aan adviseurs die beoordelen of er sprake is van overtreding. De willekeurige toedeling aan adviseurs is een voorwaarde voor toepassing van de analysetechniek die we willen inzetten voor dit onderzoek (zie hoofdstuk 3).

### 2.3 HOE VERLOOPT DEZE PROCEDURE VOOR BURGERS?

Figuur 2.1 geeft weer hoe het proces van de sollicitatieverplichting verloopt voor burgers. Dit proces begint bij de aanvraag van de WW-uitkering, waar de WW-gerechtigde wordt gewezen op het bestaan van de sollicitatieverplichting. Vervolgens moet de WW-gerechtigde sollicitatieactiviteiten uitvoeren en doorgeven. Deze worden gecontroleerd. Wanneer blijkt dat er geen of onvoldoende sollicitatieactiviteiten zijn doorgegeven, wordt gevraagd om te verantwoorden waarom dat niet is gebeurd. Op basis van die reactie, of het uitblijven daarvan, wordt bepaald of een waarschuwing of sanctie wordt opgelegd.

**Figuur 2.1 Schematische weergave van het proces van de sollicitatieverplichting**



#### **WW aanvragen en lezen over rechten en plichten**

Aan WW-gerechtigden wordt vanaf de toekenning van de WW-uitkering gecommuniceerd dat zij verplicht zijn om iedere vier weken vier sollicitatieactiviteiten door te geven. Dit wordt gecommuniceerd in de brief die naar WW-gerechtigden wordt verstuurd bij de toekenning van de WW-uitkering. Ook op de webpagina van UWV over de rechten en plichten bij een WW-uitkering staat een uitleg van deze plicht. Daarnaast wordt in de online portal van UWV, de zogeheten *Werkmap*, gewezen op de rechten en plichten die horen bij een WW-uitkering.

#### **Doorgeven van sollicitatieactiviteiten**

In de online portal van UWV wordt iedere vier weken een taak klaargezet, met daaraan gekoppeld een deadline, om binnen vier weken vier sollicitatieactiviteiten te registreren. WW-gerechtigden kunnen kiezen uit verschillende

typen activiteiten die UWV erkent als sollicitatieactiviteit, zoals solliciteren op een vacature, een open sollicitatie sturen, een sollicitatiegesprek voeren bij een werkgever, inschrijven bij een uitzendbureau (of detacheringsbureau, arbeidspool, invalpool), bellen met een werkgever voor werk, langsgaan bij een werkgever voor werk, solliciteren via een uitzendbureau of website, een assessment of test doen voor een functie, een netwerkgesprek voeren, een workshop volgen over solliciteren of een cv plaatsen op een vacaturewebsite. Door middel van een dropdownmenu kunnen WW-gerechtigden aangeven welk type activiteit zij wanneer uitgevoerd hebben.

#### **Bij onvoldoende sollicitatieactiviteiten: wat is de reden?**

Wanneer WW-gerechtigden niet aan de plicht voldoen om op tijd voldoende sollicitatieactiviteiten door te geven, neemt UWV (digitaal) contact op met de WW-gerechtigden. Aan hen wordt automatisch gevraagd om alsnog deze activiteiten door te geven en te verantwoorden waarom zij deze activiteiten niet op tijd hebben doorgegeven of niet hebben uitgevoerd. WW-gerechtigden krijgen een week de tijd om hierop te reageren. In bijlage D is een voorbeeld te zien van een bericht waarmee UWV aan WW-gerechtigden vraagt om zich te verantwoorden.

#### **Beoordeling: geldige of ongeldige reden?**

Op basis van de verantwoording van de WW-gerechtigde wordt door een medewerker van UWV (adviseur) vastgesteld of er sprake is van verwijtbaarheid. Hierbij wordt gekeken of de activiteiten wel of niet zijn uitgevoerd en wat hier de reden van is. Deze beoordeling gebeurt volgens een vast protocol, waarbij een overtreding als *niet-verwijtbaar* kan worden aangemerkt wanneer er sprake is van persoonlijke omstandigheden, ontbreken van kennis of vaardigheden, overmacht, onjuiste voorlichting of overige bijzondere omstandigheden. Bij andere redenen stelt de adviseur de mate van verwijtbaarheid vast. Wanneer er wel voldoende activiteiten zijn uitgevoerd, maar deze niet op tijd zijn doorgegeven, wordt ook de reden hiervoor beoordeeld op verwijtbaarheid. Is er sprake van een niet-verwijtbare overtreding, dan wordt aan de WW-gerechtigde gecommuniceerd dat hij of zij een geldige reden heeft voor het niet nakomen van de sollicitatieplicht (zie bijlage D voor een voorbeeld). WW-gerechtigden die niet reageren op het verzoek zich te verantwoorden, worden uitgenodigd voor een hoor-wederhoorgesprek. Als dit gesprek niet tot stand komt, beoordeelt de adviseur de casus op grond van de beschikbare gegevens.

#### **Bij ongeldige reden: waarschuwing, maatregel of geen gevolg**

Wanneer de adviseur heeft besloten dat er sprake is van verwijtbaarheid, stelt hij of zij vervolgens vast of er een waarschuwing of sanctie wordt opgelegd. Meestal wordt er een waarschuwing opgelegd voor het niet tijdig doorgeven, en een sanctie voor het niet uitvoeren van activiteiten. Zijn beide overtredingen aanwezig, dan gaat de zwaardere overtreding (niet uitvoeren van activiteiten) boven de lichtere overtreding (niet tijdig doorgeven van activiteiten). Waarschuwingen en sancties worden dus niet gestapeld en WW-gerechtigden



kunnen als uitkomst van het proces maar één sanctie of waarschuwing ontvangen. In 2018 heeft er een beleidswijziging plaatsgevonden waardoor bij de eerste overtreding van de sollicitatieplicht doorgaans een waarschuwing wordt opgelegd. Alleen wanneer een WW-gerechtigde in de afgelopen twee jaar eerder een waarschuwing heeft gekregen voor dezelfde overtreding, wordt er een sanctie opgelegd.

De WW-gerechtigde ontvangt in zijn of haar Werkmap een bericht dat de niet-nagekomen verplichting wordt beoordeeld en dat de uitkomst per brief zal worden meegedeeld (zie bijlage D voor een voorbeeld). In deze brief wordt medegedeeld of een sanctie of waarschuwing wordt opgelegd en wat de hoogte en duur van de eventuele sanctie is.

De omvang van een sanctie die wordt opgelegd wegens het niet tijdig doorgeven van de activiteiten betreft (meestal) een korting van 5% op de uitkering, voor de duur van één maand. Een sanctie die wordt opgelegd wegens onvoldoende sollicitatie-inspanningen betreft (meestal) een korting van 25% voor de duur van vier maanden. Het komt ook voor dat er sprake is van verwijtbaar gedrag, maar dat er geen waarschuwing of sanctie wordt opgelegd. Dit kan te maken hebben met WW-gerechtigden die al werk hebben, of die tegen de tijd dat de verwijtbaarheid wordt vastgesteld geen lopende uitkering meer hebben.

#### **De rol van de adviseur**

De adviseurs van UWV hebben een rol waarbij enerzijds taken worden uitgevoerd die te maken hebben met controleren of WW-gerechtigden zich aan de verplichtingen houden. Anderzijds voeren zij ook taken uit met betrekking tot het aanbieden van dienstverlening en het coachen van werkzoekenden. De beoordeling van sollicitatieactiviteiten wordt in de meeste situaties gedaan door adviseurs op het kantoor in de regio waar de WW-gerechtigde woont.

Er is geen sprake van een vaste adviseur per WW-gerechtigde. Casussen van overtreding worden normaal gesproken willekeurig toegedeeld aan adviseurs op het betreffende kantoor. De toedeling van casussen aan adviseurs gaat dan als volgt: Adviseurs worden ingedeeld om op een bepaald dagdeel of een bepaalde dag casussen af te handelen. Tijdens deze dienst beoordelen de ingedeelde adviseurs alle casussen die op dat moment openstaan. Dit kunnen casussen zijn waarvoor de dag daarvoor een reactie met een verantwoording binnen is gekomen, of casussen waarvoor de deadline voor de reactie de dag daarvoor is verlopen. Op sommige kantoren kan er sprake zijn van een specialisering van adviseurs op bepaalde groepen WW-gerechtigden, bijvoorbeeld uit een bepaalde sector. Casussen van overtreding worden dan meestal binnen deze teams toegedeeld.



## 3.1 WAT IS INSTRUMENTELE-VARIABELE-ANALYSE?

Het belangrijkste doel van dit project is om de causale effecten van waarschuwingen en sancties op nalevingsgedrag, werkzoekgedrag, arbeidsparticipatie en uitkeringsafhankelijkheid van WW-gerechtigden in kaart te brengen. Met andere woorden: We willen erachter komen of het opleggen van een waarschuwing of sanctie ervoor zorgt dat sollicitatieverplichtingen in de toekomst beter worden nageleefd en of deze maatregelen de uitstroom uit de uitkering door het vinden van werk bevorderen. Als we deze vraagstelling vertalen naar statistische terminologie, dan is het al dan niet opleggen van een sanctie onze onafhankelijke variabele ( $X$ ) en de uitkomsten zijn onze afhankelijke variabelen ( $Y$ ).<sup>4</sup> Figuur 3.1 geeft de relatie tussen de  $X$  en  $Y$  schematisch weer.

**Figuur 3.1** Schematische weergave van het verband tussen sanctie ( $X$ ) en uitkomst ( $Y$ )

**Uitgangspunt: hoe kunnen we causale effecten vaststellen?**

De meest intuïtieve aanpak om onze onderzoeksvraag te beantwoorden zou misschien zijn om WW-gerechtigden die *wel* een sanctie hebben gekregen, te vergelijken met WW-gerechtigden die *geen* sanctie hebben gekregen. We zouden dus twee groepen van uitkeringsgerechtigden kunnen vormen en vervolgens onderzoeken welke groep de regels beter naleeft en sneller uit de uitkering stroomt. Zijn dat degenen die recent een sanctie hebben ontvangen? Of juist degenen die geen sanctie hebben ontvangen? Omdat wij als onderzoekers bij

---

4 Om de statistische uitleg in dit hoofdstuk eenduidig te houden bespreken we alleen de effecten van sancties. Dezelfde uitleg is echter ook van toepassing op de analyse van de effecten van waarschuwingen. Waar *sanctie* staat, kan in dit hoofdstuk dus ook *waarschuwing* worden gelezen.

deze aanpak enkel de bestaande praktijk observeren, maar zelf niet ingrijpen, valt deze aanpak onder wat wij in bredere zin *observationeel onderzoek* noemen. Bij observationeel onderzoek kijken we enkel naar het verband tussen twee bestaande variabelen, bijvoorbeeld sanctie en werkhervatting, maar bepalen we verder niet wie er wordt blootgesteld aan een bepaalde behandeling, zoals dat bijvoorbeeld het geval is bij een klinisch onderzoek naar vaccinaties.

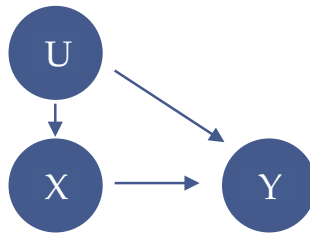
Laten we er voor nu van uitgaan dat we deze aanpak volgen en observeren dat 20% van degenen die een sanctie hebben ontvangen een jaar later een betaalde baan heeft, terwijl dit 10% is van degenen zonder sanctie. We zouden dan kunnen concluderen dat sancties de werkhervatting bevorderen, en kunnen de grootte van het effect waar we naar op zoek zijn becijferen op 10 procentpunt. Echter zit aan deze aanpak een groot methodologisch probleem verbonden, dat we nu nader zullen toelichten.

### **Het probleem van confounding**

Als we de bovenstaande observationele benadering hanteren, is het waarschijnlijk dat WW-gerechtigden die wél een sanctie krijgen opgelegd bepaalde onderscheidende kenmerken of eigenschappen hebben. Een voorbeeld is de motivatie om te zoeken naar werk: waarschijnlijk zitten er onder de groep gesanctioneerden velen die in mindere mate gemotiveerd zijn om werkzoekactiviteiten te verrichten, wat deels verklaart waarom zij een sanctie hebben gekregen. Een ander voorbeeld is parttimewerk: het zou kunnen dat adviseurs van UVW minder sterk geneigd zijn om een sanctie op te leggen aan WW-gerechtigden die naast hun uitkering werken. Kortom, WW-gerechtigden met bepaalde kenmerken of eigenschappen lopen zeer waarschijnlijk een groter risico om gesanctioneerd te worden, terwijl er andere kenmerken of eigenschappen zijn die de kans op een sanctie kunnen verminderen.

Tegelijkertijd zijn deze kenmerken en eigenschappen hoogstwaarschijnlijk ook gerelateerd aan de uitkomstmaten die we binnen dit onderzoek willen bestuderen. Zo is het aannemelijk dat WW-gerechtigden met weinig motivatie om werk te zoeken gemiddeld minder snel de uitkering uitstromen, terwijl dit waarschijnlijk sneller lukt voor degenen die al parttime aan het werk zijn. We zijn dus geconfronteerd met een situatie waarin bepaalde factoren (zoals motivatie en het al dan niet hebben van parttimewerk) zowel een verband hebben met de kans op een sanctie als met de uitkomstmaat waarin we geïnteresseerd zijn (in dit voorbeeld de kans op werkhervatting). In figuur 3.2 zijn deze factoren ( $U$ ) toegevoegd aan onze schematische weergave van het verband tussen  $X$  en  $Y$ .

**Figuur 3.2** Verband tussen sanctie (X) en uitkomst (Y) met verstorende factoren (U)



Doordat deze factoren bestaan, kunnen we niet met zekerheid zeggen of de verschillen die we hebben gevonden tussen de twee groepen – de groepen met en zonder sanctie – verklaard worden doordat de ene groep een sanctie heeft ontvangen en de andere niet, of door andere, reeds bestaande verschillen tussen de groepen. Het kan dus gebeuren dat we effecten die eigenlijk toe te schrijven zijn aan andere factoren, toeschrijven aan het opleggen van een sanctie. Een mogelijk gevolg hiervan is dat we het daadwerkelijke effect van een sanctie onderschatten. Misschien worden vooral WW-gerechtigden gesanctioneerd die weinig gemotiveerd zijn én niet zo snel een baan vinden, waardoor het effect van de sanctie kleiner lijkt dan het in werkelijkheid is. Maar mogelijk overschatten we het daadwerkelijke effect ook, omdat vooral degenen wier kansen op werk hoog zijn een sanctie krijgen. Er bestaat dus een grote kans dat het effect dat we hebben berekend, de 10 procentpunt, niet klopt.

Kortom, er ontstaat een probleem bij het vaststellen van causale effecten als er factoren zijn die zowel samenhangen met de onafhankelijke variabele als met de onderzochte uitkomsten. Omdat deze factoren de onderzochte effecten kunnen verstoren, wordt dit probleem in de statistiek vaak als *confounding* aangeduid.

#### **Mogelijke oplossing 1: controleren voor relevante factoren**

Er zijn meerdere manieren om met dit probleem om te gaan. Een denkbare en voor de hand liggende oplossing voor *confounding* is het vergelijken van WW-gerechtigden met en zonder sanctie die wat betreft de reeds bestaande verschillen zo veel mogelijk op elkaar lijken. We zouden dus eigenlijk alleen WW-gerechtigden met elkaar moeten vergelijken die bijvoorbeeld dezelfde achtergrondkenmerken, dezelfde werkstatus en hetzelfde motivatieniveau hebben. Statistisch is het in principe makkelijk om zulke factoren constant te houden. In een regressiemodel zouden we bijvoorbeeld variabelen kunnen toevoegen die allemaal kenmerken en eigenschappen meten waarvan we denken dat zij ook relevant zijn. Vaak zeggen we dan dat we in ons model voor de variabelen die we opnemen *controleren*. Bij de interpretatie van resultaten kunnen we dan alle variabelen die we hebben opgenomen in ons model als constant veronderstellen en het effect van een sanctie, *ceteris paribus* ('het overige gelijk blijvend'), bepalen.

Deze oplossing werkt echter alleen in een situatie waarin we al die relevante factoren niet alleen kennen, maar ook kunnen meten. Anders kunnen we de factoren immers niet als variabele opnemen in ons model. Nu is het in de praktijk waarschijnlijk dat veel van de factoren die zowel samenhangen met de onafhankelijke als met de afhankelijke variabele buiten ons beeld zijn. En als het ons al lukt om met hulp van theoretische modellen en eerder onderzoek de belangrijkste factoren te identificeren, blijft het nog steeds lastig om al deze factoren goed te meten. In tegenstelling tot werkstatus of sociodemografische gegevens, die makkelijk meetbaar zijn, zal er ook een aantal factoren zijn die niet goed gemeten kunnen worden en waarover dan ook geen data beschikbaar zijn. Voorbeelden hiervan zijn factoren zoals motivatie of bepaalde attitudes en vaardigheden. Als we niet alle relevante factoren kennen en kunnen meten, zal het opnemen van variabelen in ons statistisch model het oorspronkelijke probleem van verstoring hooguit beperken, maar niet verhelpen. Onze resultaten zullen waarschijnlijk nog steeds niet betrouwbaar zijn, juist omdat we niet alle relevante factoren constant kunnen houden.

### **Mogelijke oplossing 2: experimenteren**

Er bestaan echter ook andere statistische methoden die een betere oplossing kunnen bieden voor het probleem van *confounding*. Eén voorbeeld hiervan, dat recent veel aandacht heeft gekregen in de economie, is de experimentele aanpak (waar een zogenoemde *randomized controlled trial* (RCT) een voorbeeld van is).<sup>5</sup> Terwijl we bij een observationeel onderzoek geen invloed hadden op wie er wordt blootgesteld aan een bepaalde behandeling, is dit bij een experiment juist wel het geval. Kenmerkend voor een experiment is namelijk dat er door een lotingmechanisme (*randomisatie*) wordt bepaald wie een bepaalde behandeling ontvangt en wie niet.

In ons geval zou de behandeling het al dan niet opleggen van een sanctie zijn. We zouden dan dus niet meer enkel observeren wie is gesanctioneerd, maar middels randomisatie zelf bepalen wie in aanmerking komt voor een sanctie. Door het aselekt toewijzen van de behandeling zijn de groep die de behandeling wel ontvangt en de groep die de behandeling niet ontvangt (naar verwachting) vergelijkbaar met betrekking tot andere relevante factoren, ongeacht of die bekend en/of meetbaar zijn. Hierdoor laat zich makkelijk berekenen met welke betrouwbaarheid een gevonden verschil tussen de behandelgroep en de controlegroep kan worden toegeschreven aan de andere behandeling.

Een experiment biedt in vele situaties een mooie oplossing voor het probleem van *confounding*, maar brengt helaas ook een aantal belemmeringen met zich. Zo zijn er veelal ethische aspecten, bijvoorbeeld het argument van rechtsongelijkheid tussen de groepen of de vaak grote financiële en organisatorische inspanningen, die de uitvoering van een experiment in de weg staan. Juist bij

---

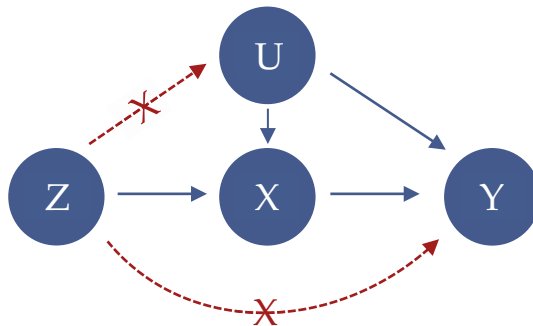
5 De economen Esther Duflo, Abhijit Banerjee en Michael Kremer wonnen in 2019 de Nobelprijs voor Economie voor het inzetten van de experimentele aanpak om wereldwijd armoede te verlichten.

een ingrijpende behandeling zoals het opleggen van sancties zal een experimentele aanpak vanwege ethische overwegingen snel worden uitgesloten.

### Mogelijke oplossing 3: instrumentele-variabele-analyse

Een andere, relatief nieuwe, statistische aanpak die eveneens een oplossing kan bieden voor het probleem van *confounding*, is de zogeheten *instrumentele-variabele-analyse* (ook IVA; Angrist, Imbens, & Rubin, 1996). Deze aanpak brengt niet dezelfde belemmeringen met zich als een experiment, omdat de methode wordt toegepast op bestaande data en onderzoekers dus geen invloed hebben op de toewijzing van een behandeling. IVA is dus ook een vorm van observationeel onderzoek. Dit is de methode die we in ons onderzoek gaan gebruiken. IVA maakt gebruik van een aanvullende variabele in het model, die *instrumentele variabele* wordt genoemd. Een instrumentele variabele is een factor die de behandeling waar het om gaat beïnvloedt ( $X$ ), in ons geval dus een factor die invloed heeft op het opleggen van een sanctie, maar geen direct verband heeft met de afhankelijke variabele in het model ( $Y$ ) of andere relevante factoren ( $U$ ). In figuur 3.3 is de instrumentele variabele ( $Z$ ) toegevoegd aan onze schematische weergave van het verband tussen  $X$  en  $Y$ . Een instrumentele variabele introduceert (in elk geval deels) toeval in het toewijzen van een behandeling binnen observationele data (soms ook *pseudorandomisatie* genoemd) en werkt daarom op een vergelijkbare manier als het lotingmechanisme binnen een experiment. Hierdoor kan het probleem van *confounding* mogelijk worden opgelost. We zullen de werkwijze van deze aanpak nu nader toelichten aan de hand van onze eigen onderzoekssituatie.

**Figuur 3.3** Verband tussen sanctie ( $X$ ) en uitkomst ( $Y$ ) met versturende factoren ( $U$ ) en instrument ( $Z$ )



## 3.2 STRENGHEID ALS INSTRUMENTELE VARIABELE

Om IVA toe te kunnen passen moeten we op zoek gaan naar een factor die invloed heeft op het opleggen van een sanctie, maar geen direct verband heeft met onze gekozen uitkomsten of andere relevante factoren. Een factor die mogelijk hiervoor in aanmerking komt, is de *strengheid* van de UWV-adviseur als het

gaat om het opleggen van een sanctie. Wat strengheid betekent en waarom deze factor een geschikte instrumentele variabele kan zijn, zullen we nu verder uitleggen.

### **Wat is strengheid?**

Zoals eerder toegelicht, beslissen UWV-adviseurs welke consequentie er volgt op het niet nakomen van de sollicitatieplicht. Voor het bepalen van de consequentie bestaan er werkinstructies waarin is vastgelegd welke stappen in het beoordelingsproces ondernomen dienen te worden en op welke manier bepaalde criteria mee moeten worden gewogen. Toch hebben adviseurs bij het nemen van hun beslissing vermoedelijk enige discretionaire ruimte. Als die ruimte er is, dan zullen adviseurs waarschijnlijk verschillen in hoe streng zij zijn bij het bepalen van de consequentie – terwijl adviseur A in een gegeven geval een sanctie zou opleggen, zou adviseur B in dezelfde situatie wellicht voor géén sanctie kiezen.

De strengheid van adviseurs kan worden gemeten. In het makkelijkste geval kan dat door het aantal casussen waarin de adviseur heeft gekozen voor een sanctie te delen door het totale aantal casussen die door de adviseur zijn behandeld. Strengheid wordt zodoende gemeten op een schaal van 0 tot 1, waarbij de meest strenge adviseurs (met waarde 1) in alle gevallen een sanctie opleggen en de minst strenge adviseurs (met waarde 0) in geen van de gevallen. Maar waarom is de strengheid van adviseurs een kansrijke kandidaat voor een instrumentele variabele?

Ter herinnering: een instrumentele variabele dient invloed te hebben op de gekozen behandeling ( $X$ ; in ons geval de sanctie), maar mag geen direct verband hebben met de afhankelijke variabele in het model ( $Y$ ) of met andere relevante factoren ( $U$ ). Strengheid is ten eerste een kansrijke kandidaat omdat de strengheid van een adviseur waarschijnlijk sterk zal samenhangen met de kans om een sanctie opgelegd te krijgen ( $X$ ). De kans op een sanctie is namelijk zeer waarschijnlijk groter als de casus wordt behandeld door een strenge adviseur, en andersom. Ten tweede zal de strengheid van een adviseur zeer waarschijnlijk geen direct verband hebben met onze afhankelijke variabelen ( $Y$ ). Als adviseurs casussen puur beoordelen zonder op een andere manier invloed uit te oefenen op een WW-gerechtigde, dan kan hun strengheid de uitkomsten alleen via het opleggen van een sanctie beïnvloeden. Ten derde worden casussen nagenoeg willekeurig toegedeeld aan een adviseur (zie hoofdstuk 2). Iedere adviseur heeft daarom gemiddeld te maken met dezelfde soort WW-gerechtigden. En dat betekent wederom dat er geen verband kan zitten tussen de strengheid van een adviseur en andere relevante factoren ( $U$ ; zoals motivatie of parttimewerk). Mocht het dus zo zijn dat (a) adviseurs daadwerkelijk verschillen in hoe streng ze zijn bij het beoordelen van een casus, (b) er weinig interactie is tussen adviseur en WW-gerechtigde, en (c) verantwoordingsstaken willekeurig worden toegewezen aan adviseurs, dan is strengheid een kansrijke kandidaat om gebruikt te kunnen worden als instrumentele variabele.



**Hoe kunnen we door middel van strengheid causale effecten schatten?**

Om uit te leggen hoe we strengheid als instrumentele variabele kunnen gebruiken om causale effecten te schatten zullen we nu twee sterk vereenvoudigde voorbeelden gebruiken.<sup>6</sup> Stel dat er bij UWV slechts twee adviseurs werken die casussen beoordelen, en dus beslissen over het wel of niet opleggen van een sanctie. Stel verder dat de ene adviseur (adviseur A) heel streng is en de andere adviseur (adviseur B) niet. Dat wil zeggen dat adviseur A bij elke casus die op zijn tafel komt een sanctie oplegt, terwijl adviseur B in geen enkel geval voor het opleggen van een sanctie kiest. De strengheid van de adviseurs is in dit geval een perfecte voorspeller voor het ontvangen van een sanctie.

WW-gerechtigden die niet aan de sollicitatieplicht hebben voldaan, worden willekeurig toegewezen aan een van de twee adviseurs. Hierdoor ontstaan twee groepen uitkeringsgerechtigden: degenen wier casus door adviseur A (streng) wordt behandeld en degenen wier casus bij adviseur B (niet streng) belandt. Omdat deze twee groepen door toeval zijn ontstaan, kunnen we, als de groepen groot genoeg zijn, ervan uitgaan dat zij vergelijkbaar zijn met betrekking tot alle mogelijke observeerbare en niet-observeerbare achtergrondkenmerken. We gaan er verder van uit dat de twee adviseurs niets anders doen dan een beslissing nemen over de sanctie. Het is bijvoorbeeld niet zo dat een van de adviseurs ook nog een aantal tips geeft over succesvol baanzoekgedrag.

Om te weten te komen wat het effect is van een sanctie hoeven we nu alleen nog maar deze twee groepen uitkeringsgerechtigden met elkaar te vergelijken. We weten immers dat de ene groep altijd een sanctie zal ontvangen en de andere nooit. Laten we ervan uitgaan dat 20% van de groep van adviseur A (streng) een jaar later een betaalde baan heeft, tegenover 10% van de groep van adviseur B (niet streng). Er zit dus een verschil van 10 procentpunt tussen die twee groepen, een verschil dat we redelijk zeker volledig kunnen toeschrijven aan het wel of niet ontvangen van een sanctie. In dit geval zouden we concluderen dat sancties de werkherlevingskans met 10 procentpunt verhogen.

Nu zullen er in de realiteit weinig gevallen zijn waarin adviseurs óf altijd óf nooit een sanctie opleggen. De strengheid van de meeste adviseurs ligt waarschijnlijk tussen de twee extremen 0 (geeft nooit een sanctie) en 1 (geeft altijd een sanctie). Laten we dus voor een realistischer voorbeeld ervan uitgaan dat adviseur A in 80% van de casussen een sanctie oplegt (strengheid van 0,8). Adviseur B is nog steeds minder streng en legt in 40% van de casussen een sanctie op (strengheid van 0,4). Net als bij het vorige voorbeeld wordt nog steeds door toeval bepaald bij welke adviseur een casus belandt. Ook gaan we er nog steeds van uit dat adviseurs alleen sanctioneren en verder geen andere behandeling uitvoeren. In tegenstelling tot het vorige voorbeeld wordt het ontvangen van een sanctie nu niet meer volledig bepaald door bij welke adviseur men terecht komt. Wel verschilt de kans op een sanctie tussen de twee groepen. De groep van adviseur A heeft een 40 procentpunt ( $0,8 - 0,4 = 0,4$ ) hogere kans op

---

6 De voorbeelden zijn ontleend aan Boef, Le Cessie, & Dekkers (2013) en toegepast op onze eigen onderzoekssituatie.

een sanctie vergeleken met de groep van adviseur B. Dit verschil in kans op een sanctie is nog steeds toevallig bepaald en kan daarom ook nog steeds worden gebruikt voor het schatten van een causaal effect.

Om een causaal effect te schatten kunnen we opnieuw kijken naar de twee groepen uitkeringsgerechtigden, die van adviseur A en die van adviseur B. Belangrijk is dat we nog steeds degenen vergelijken die zijn toegewezen aan een bepaalde adviseur, en niet degenen die wel of niet een sanctie hebben gekregen. Stel dat 16% van groep A een jaar later een betaalde baan heeft, tegenover 12% van groep B. Het verschil tussen de twee groepen ligt dus bij 4 procentpunt.

In tegenstelling tot eerder kunnen we dit verschil van 4 procentpunt niet volledig toeschrijven aan het ontvangen van een sanctie. Immers, er zijn WW-gerechtigden in groep A die geen sanctie hebben gekregen (20% van de groep), evenals WW-gerechtigden in groep B die wél zijn gesanctioneerd (60% van de groep). Wat we wel weten is dat het verschil van 4 procentpunt komt doordat de twee groepen een ander risico lopen op een sanctie. Wetende dat de twee groepen toevallig zijn ontstaan, kunnen we er ook nog steeds van uitgaan dat er geen andere factoren zijn die oorzakelijk zijn voor dit verschil. We weten verder dat het verschil in sanctierisico dat de twee groepen lopen 40 procentpunt bedraagt.

Met deze informatie kunnen we berekenen wat het verschil tussen de groepen zou zijn als adviseur A 100% van de groep zou sanctioneren en adviseur B 0%, zoals bij het vorige voorbeeld. Hiervoor delen we het gevonden verschil van 4 procentpunt door het verschil in kans op een sanctie van 40 procentpunt ( $0,04 / 0,4 = 0,1$ ). Bij een strengheid van 100% en 0% zou het verschil tussen de twee groepen dus 10 procentpunt bedragen. Of anders geformuleerd: het verschil van 4 procentpunt gegeven dat één groep 40 procentpunt meer kans maakt op een sanctie kan alleen ontstaan als het pure effect van een sanctie 10 procentpunt bedraagt.

Deze twee voorbeelden illustreren de werkwijze van IVA met strengheid als instrumentele variabele. Box 3.1 beschrijft de twee stappen van IVA in meer detail. Figuur 3.4 verbeeldt de IVA-aanpak en zet deze af tegen een andere, eerder besproken oplossing voor *confounding*, de experimentele aanpak.

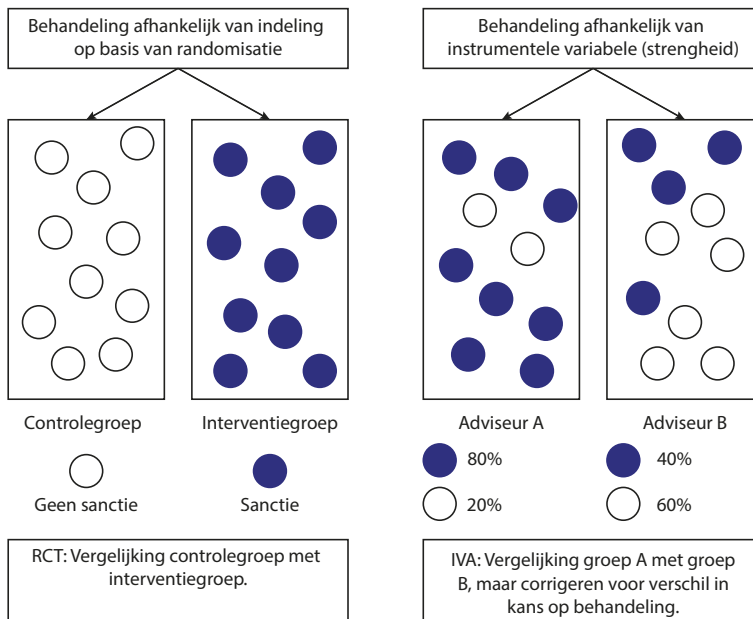
**Box 3.1 De twee stappen van een instrumentele-variabele-analyse (IVA)**

Zoals uit de eerdere voorbeelden blijkt, houdt het schatten van causale effecten middels IVA twee stappen in. In een eerste stap bepalen we wat het verschil in uitkomsten is tussen de groepen die toevallig zijn ontstaan (in ons voorbeeld: de groep van adviseur A en die van adviseur B). In een tweede stap corrigeren we het gevonden verschil tussen de groepen voor het feit dat de kans op behandeling tussen de groepen (toevallig) verschilde. Op deze manier kunnen we het causale effect van de behandeling zelf isoleren.

Omdat we in de realiteit niet te maken hebben met een simpele situatie waarin alleen twee adviseurs bestaan, gebruiken we voor deze twee stappen normaliter een statistische analysetechniek die ook *two-stage least squares* (2SLS) regressie wordt genoemd. Bij deze analysetechniek wordt eerst het verband tussen instrument ( $Z$ ) en behandeling ( $X$ ) geschat en vervolgens het verband tussen behandeling ( $X$ ) en uitkomst ( $Y$ ). In de eerste stap wordt daarbij een nieuwe variabele gecreëerd ( $\hat{X}$ ).  $\hat{X}$  is de voorspelde waarde van  $X$ . Deze voorspelde waarde bevat alleen de variatie in behandeling die door toeval is ontstaan. In de tweede stap (verband tussen  $X$  en  $Y$ ) vervangt deze nieuwe variabele de oude, 'problematische' variabele  $X$  als onafhankelijke variabele.

In de wetenschappelijke literatuur zijn enkele voorbeelden bekend van effectieve toepassingen van IVA met strengheid als instrumentele variabele. Zo is bijvoorbeeld de strengheid van rechters gebruikt om het effect van detentie te schatten (Bhuller, Dahl, Løken, & Mostad, 2020). Ander onderzoek kijkt naar het effect van arbeidsongeschiktheidsuitkeringen en gebruikt hiervoor de strengheid van degenen die de aanvragen voor dit soort uitkeringen beoordelen (Autor, Kostøl, Mogstad, & Setzler, 2019). Een derde voorbeeld is het effect van patenteren, waarvoor de strengheid van octrooigemachtigden is gebruikt (Sampat & Williams, 2019). In de literatuur over sanctiebeleid en regelnaleving is de toepassing van deze aanpak echter nog niet vaak te vinden. Het is belangrijk om te noemen dat er naast strengheid van adviseurs of beoordelaars deels ook andere factoren zijn die mogelijk als instrumentele variabele kunnen worden gebruikt. Zo kan er bijvoorbeeld soms ook gebruik worden gemaakt van behandelverschillen tussen regio's, kantoren of praktijken. In elk geval is te toetsen of de factor geschikt is als instrumentele variabele. De aannames die hierbij belangrijk zijn zullen we in de navolgende paragrafen in meer detail belichten.

**Figuur 3.4 Schematisch overzicht experiment (RCT) versus instrumentele-variabele-analyse (IVA)**



*Noot:* De figuur is geïnspireerd door het figuur in Boef, Le Cessie, & Dekkers (2013).

### Hoe kunnen we IVA-schattingen het beste interpreteren?

Het is belangrijk om te noemen dat het effect dat we met IVA berekenen eigenlijk betrekking heeft op een bepaalde groep WW-gerechtigden binnen onze onderzoekspopulatie. Laten we om dit beter uit te leggen eerst een onderscheid maken tussen drie groepen uitkeringsgerechtigden.

Ten eerste zijn er WW-gerechtigden wier regelovertreding zo duidelijk verwijtbaar is dat zij *altijd* een sanctie opgelegd krijgen, ongeacht of zij bij een strenge of minder strenge adviseur belanden. Omdat deze groep altijd een sanctie krijgt opgelegd, kunnen we voor deze groep niet vaststellen wat het effect van een sanctie is.

Ten tweede zijn er WW-gerechtigden die de regels heel duidelijk voor niet-verwijtbare redenen hebben overtreden en daarom *nooit* een sanctie krijgen opgelegd, ongeacht bij welke adviseur zij belanden. Ook voor deze groep kunnen we het effect van een sanctie niet berekenen, juist omdat deze groep nooit een sanctie krijgt.

Ten derde zijn er WW-gerechtigden die tussen deze twee categorieën in zitten. Dat wil zeggen: deze subgroep bestaat uit WW-gerechtigden die een andere behandeling *hadden kunnen krijgen* als zij terecht waren gekomen bij een andere adviseur. Of in andere woorden: dit zijn WW-gerechtigden die een sanctie hadden gekregen als ze bij een strengere adviseur waren beland, maar niet als ze bij een minder strenge adviseur waren beland. Alleen voor deze derde groep verandert de behandeling afhankelijk van welke adviseur hun casus beoordeeld, en het effect dat wij berekenen heeft dan ook alleen betrekking op deze groep.

De effecten die met IVA worden geschat moet dan ook niet gezien worden als het gemiddelde effect van het krijgen van een sanctie voor de hele onderzoekspopulatie. Het effect dat we berekenen kan het beste geïnterpreteerd worden als een zogenoemd *local average treatment effect* (LATE): het effect van een sanctie voor degenen die een andere behandeling hadden kunnen krijgen als zij terecht zouden zijn gekomen bij een andere adviseur.

### **Samenvatting**

Als we willen weten wat het causale effect is van sancties op verschillende uitkomstmaten, moeten we ervan uitgaan dat er (*confounding*) factoren zijn die het vaststellen van een causaal verband belemmeren omdat zij samenhangen met zowel de kans op een sanctie als met onze uitkomsten. Om dit probleem te omzeilen maakt IVA gebruik van een aanvullende variabele die (deels) toeval introduceert in een observationele dataset. In onze onderzoekssituatie dient de *strengheid* van UWV-adviseurs zich aan als instrumentele variabele. Doordat WW-gerechtigden willekeurig worden toegewezen aan strenge of minder strenge adviseurs, wordt de kans op een sanctie vergroot afhankelijk van wie hun casus beoordeelt. Door deze aselecte toewijzing kan verder worden uitgesloten dat de strengheid van de adviseur samenhangt met andere relevante factoren die de afhankelijke variabelen kunnen beïnvloeden. Uiteindelijk kunnen door de toevallige variatie in de kans op een sanctie causale verbanden worden geschat op de gekozen uitkomstmaten. Wel moeten we vooraf zorgvuldig overwegen of aan een aantal belangrijke aannames voor de toepassing van IVA kan worden voldaan (box 3.2).

### **Box 3.2 Aannames voor het gebruik van een instrumentele-variabele-analyse (IVA)**

Om IVA toe te kunnen passen moeten er vier aannames worden gedaan. In sommige gevallen is het mogelijk om deze aannames op basis van de gebruikte data te onderzoeken. Echter, het kan ook voorkomen dat het voldoen aan de aannames puur theoretisch moet worden onderbouwd. De eerste drie aannames zijn we al eerder tegengekomen bij de zoektocht naar een geschikt instrument. We zullen de aannames nu kort toelichten aan de hand van onze eigen onderzoekssituatie.

#### *Aanname 1: Relevantie*

Deze aanname zegt dat er een (sterk) verband moet zijn tussen de instrumentele variabele ( $Z$ ) en de behandeling ( $X$ ). Dit betekent in ons geval: strengheid moet de kans op een sanctie respectievelijk waarschuwing goed voorspellen. Het verband tussen de instrumentele variabele en de behandeling kan statistisch worden onderzocht, en de sterkte van de samenhang kan worden berekend. Een zwak verband tussen instrumentele variabele en behandeling kan de schattingen van effecten serieus verstoren.

#### *Aanname 2: Onafhankelijkheid*

Deze aanname houdt in dat de instrumentele variabele ( $Z$ ) niet mag samenhangen met andere factoren ( $U$ ) die een uitkomstmaat ( $Y$ ) kunnen beïnvloeden. Anders zouden we namelijk weer terug zijn bij het oorspronkelijke probleem van *confounding*. Dat wil zeggen: we zouden niet meer met zekerheid weten of ons effect is ontstaan door een bepaalde behandeling of ook door andere factoren. In ons geval is het daarom belangrijk dat casussen zo veel mogelijk aselekt worden toegewezen aan adviseurs. De strengheid van adviseurs hangt dan niet samen met andere factoren van invloed.

Om de tweede aanname te toetsen moeten we in eerste instantie goed kijken naar de manier van toedeling. Gebeurt dat altijd aselekt of zijn er bepaalde uitzonderingen? Op welk niveau vindt de toedeling plaats? Per regio, per kantoor, per unit? Kunnen adviseurs casussen ruilen? Zijn er werkinstructies of andere procesdocumenten die vastleggen op welke manier de toedeling plaatsvindt?

In tweede instantie kan deze aanname deels ook statistisch worden onderzocht. Als er data beschikbaar zijn over achtergrondkenmerken van de onderzoekspopulatie, kan namelijk worden gekeken of deze statistisch samenhangen met de instrumentele variabele. Zien we bijvoorbeeld in de data dat strenge adviseurs vooral met jongeren te maken hebben en minder strenge adviseurs met ouderen? Dit zou een aanwijzing kunnen zijn dat de aanname geschonden is. Belangrijk om te noemen is hierbij dat deze statistische test alleen uitsluitsel kan geven over kenmerken die we in onze dataset ook daadwerkelijk observeren. We moeten er dus van uitgaan dat als we geen verband vinden op basis van geobserveerde kenmerken, dit eveneens geldt voor niet geobserveerde kenmerken.

### *Aanname 3: Exclusie*

Volgens deze aanname moet de instrumentele variabele ( $Z$ ) de uitkomsten ( $Y$ ) alleen via de kans op behandeling ( $X$ ) beïnvloeden en niet via een andere weg. Dat betekent in ons geval dat strengheid alleen via de kans op een sanctie respectievelijk waarschuwing invloed mag hebben op een uitkomst zoals werkhervatting. Dit zou niet het geval zijn als bijvoorbeeld strenge adviseurs ook andere behandelingen toepassen die een effect kunnen hebben op werkhervatting, denk aan een extra training of workshop.

Stel dat we niet kunnen uitsluiten dat strengheid gepaard gaat met andere behandelingen die onze uitkomsten mogelijk beïnvloeden. In dat geval zouden effecten net zo goed kunnen worden toegeschreven aan een verschil in gebruik van deze praktijken. We moeten dus proberen om er zo goed mogelijk achter te komen of strengheid gepaard gaat met een andere manier van behandeling. We zullen ons hiervoor vragen stellen zoals: Hebben adviseurs überhaupt de mogelijkheid om andere behandelingen toe te passen? Vindt er persoonlijk contact plaats waarin adviseurs zouden kunnen sturen? Of gaat het om een afhandeling zonder persoonlijk contact? Antwoorden op deze vragen zullen deels in werkinstructies en procesdocumenten te vinden zijn, maar we zullen ook verkennen of we de antwoorden in de beschikbare data kunnen vinden.

### *Aanname 4: Monotoniciteit*

Monotoniciteit betekent dat de instrumentele variabele ( $Z$ ) de kans op behandeling ( $X$ ) alleen in één richting beïnvloedt. In ons geval betekent het dat bijvoorbeeld de kans op een sanctie voor iedereen moet toenemen naarmate zij worden toegewezen aan een strengere adviseur. Of met andere woorden: er bestaan geen casussen waarbij een strenge adviseur géén sanctie zou opleggen, en een minder strenge adviseur wél.

Waarom deze aanname belangrijk is, kan het best worden uitgelegd aan de hand van ons eerdere voorbeeld, waar we groep A (strenge adviseur) en groep B (minder strenge adviseur) met elkaar vergeleken en het gevonden verschil in uitkomsten corrigeerden voor het verschil in kans op sanctie. We wisten toen dat groep A en groep B met elkaar vergelijkbaar waren door de toevallige indeling. Alleen het toeval bepaalde of men blootgesteld zou worden aan een hoge of een lage kans op een sanctie. Juist de uitwisselbaarheid van de twee groepen maakte het mogelijk om het effect van een andere behandeling te identificeren.

Als er nu personen zijn voor wie strengheid de andere kant op werkt, dat wil zeggen personen die bij adviseur B meer kans maken op een sanctie en bij adviseur A minder kans maken op een sanctie, dan is de uitwisselbaarheid van de twee groepen niet meer een gegeven. De schatting van causale effecten wordt dan verstoord. Net als de aanname van *exclusie* wordt de aanname van *monotoniciteit* vaak theoretisch getoetst. Toch is het in sommige gevallen ook mogelijk om met statistische tests aanwijzingen te verzamelen dat er aan de aanname wordt voldaan.

### 3.3 WELKE DATA ZIJN ER NODIG VOOR DIT ONDERZOEK?

Om onze onderzoeksvraag te beantwoorden hebben we data nodig waarmee we:

- kunnen nagaan wie een sanctie respectievelijk waarschuwing krijgt opgelegd;
- kunnen vaststellen hoe streng adviseurs zijn;
- kunnen observeren welke uitkomsten gerealiseerd worden;
- over voldoende informatie beschikken om aannames te testen.

Om tot een dataset te komen die aan deze eisen voldoet koppelen we meerdere bestaande, grote administratieve gegevensbronnen aan elkaar. Omdat we enkel administratieve databronnen gebruiken, lopen we verder weinig gevaar op missende data zoals dit vaak bij vragenlijstenonderzoek voorkomt. Ook kunnen we redelijk zeker zijn dat de gebruikte data van hoge kwaliteit zijn. We zullen de gebruikte informatie nu nader toelichten.

#### **Data over het opleggen van waarschuwingen en sancties**

In de registratiesystemen van UWV is vastgelegd welke WW-gerechtigden hun sollicitatieplicht niet zijn nagekomen en zich hiervoor moeten verantwoorden. Voor elke zogenoemde *verantwoordingstaak* is ook bekend op welke datum deze taak is aangemaakt. Hierdoor kunnen we casussen selecteren waar voor het eerst een overtreding heeft plaatsgevonden. We kiezen ervoor om alleen eerste verantwoordingsstaken mee te nemen, zodat we zeker weten dat de effecten die we observeren niet afhankelijk zijn van eerdere overtredingen. Het zou bijvoorbeeld kunnen voorkomen dat *coulance* bij een eerste overtreding wordt gecompenseerd bij de beoordeling van een vervolgovertreding. Verder staat geregistreerd door welke UWV-adviseur een verantwoordingsstaak is afgehandeld en voor welke consequentie de adviseur heeft gekozen. We kunnen dus identificeren wat de relevante casussen zijn, wat de consequentie van de overtreding was, en door wie die consequentie is bepaald.

#### **Data over strengheid van adviseurs**

We kunnen identificeren (a) hoeveel taken een adviseur in totaal heeft afgehandeld en (b) in hoeveel gevallen een bepaalde consequentie (bijvoorbeeld een sanctie) is opgelegd. Op basis van deze informatie kunnen we voor elke adviseur berekenen wat zijn of haar strengheid is.

#### **Data over uitkomsten**

Omdat verantwoordingsstaken zijn voorzien van een tijdstempel, kunnen we nagaan welke uitkomsten er in de maanden na overtreding zijn gerealiseerd. Zo kunnen we bijvoorbeeld vaststellen of er na de eerste overtreding nog verdere overtredingen hebben plaatsgevonden. Ook kunnen we vaststellen hoeveel sollicitatieactiviteiten iemand in de maanden na de overtreding heeft doorgegeven. Door informatie uit het uitbetalingssysteem van UWV kunnen



we identificeren of mensen een uitkering ontvangen en wat de eventuele hoogte is van het bedrag dat is overgemaakt.

Verder kunnen we data uit de UWV-systemen via een anonieme identificatiesleutel koppelen aan informatie uit de Polisadministratie. In de Polisadministratie verzamelt UWV informatie over lonen en arbeidscontracten. Het gaat hierbij om de gegevens van alle verzekerde werknemers in Nederland. Via deze databron kunnen we arbeidsparticipatie in kaart brengen en vaststellen of men een betaalde baan heeft en, zo ja, voor hoeveel uur, voor welk loon en met welk soort contract.

#### **Data om aannames te testen**

Voor het toepassen van IVA moeten we zo goed mogelijk nagaan of aan bepaalde aannames kan worden voldaan. We willen bijvoorbeeld weten of strengheid samenhangt met bepaalde achtergrondkenmerken van WW-gerechtigden. Hiervoor maken we gebruik van data uit een ander registratiesysteem van UWV, waarin informatie over leeftijd, geslacht en opleidingsniveau wordt opgeslagen. Ook willen we informatie verzamelen over adviseurs en hun beslissingen. Door raadpleging van weer andere UWV-systemen kunnen we erachter komen welke dienstverlening door UWV-adviseurs wordt ingezet.



## 4.1 AANPAK

We zijn geïnteresseerd in de causale effecten van waarschuwingen en sancties op nalevingsgedrag, werkzoekgedrag, arbeidsparticipatie en uitkeringsafhankelijkheid van WW-gerechtigden. We passen een instrumentele-variabeelanalyse toe om dit verband te onderzoeken. Voor de analyse gebruiken we de strengheid van adviseurs van UWV als instrumentele variabele. In deze paragraaf zullen we onze aanpak nader toelichten.

In hoofdstuk 2 lieten we zien dat een verantwoordingstaak verschillende uitkomsten kan hebben, namelijk een beoordeling van de overtreding als *niet-verwijtbaar*, een *waarschuwing*, een *sanctie* en *geen consequentie ondanks verwijtbaar gedrag*. We stellen effecten vast door *gesanctioneerde* en *gewaarschuwde* personen te vergelijken met personen wier gedrag als *niet-verwijtbaar* is beoordeeld. De groep waarvan het gedrag als verwijtbaar is beoordeeld zonder dat er een consequentie is geregistreerd, laten we dus buiten beschouwing.

We schatten causale effecten met een zogeheten *two-stage least squares* (2SLS) regressie (zie box 3.1). In de modellen controleren we voor de achtergrondkenmerken van WW-gerechtigden, indien bekend. Dat betekent dat we deze factoren constant houden voor het geval dat niet alle kenmerken van WW-gerechtigden even goed zijn verdeeld over alle adviseurs (ondanks de aselechte toedeling). Ook nemen we controlevariabelen op voor het kantoor en de maand waarin de casus is afgehandeld. Dit doen we omdat casussen per kantoor worden toegedeeld. Door kantoor en maand constant te houden sluiten we uit dat onze resultaten verklaard kunnen worden door verschillen in beleid over de tijd, of door verschillen in de groep uitkeringsgerechtigden die binnen het werkgebied van een kantoor valt. We vergelijken dus uiteindelijk WW-gerechtigden met dezelfde achtergrondkenmerken die bij hetzelfde kantoor horen en in dezelfde maand wel of niet een sanctie respectievelijk waarschuwing hebben gekregen. Voor de precieze modelspecificaties voor de analyse verwijzen we de lezer naar bijlage A.

## 4.2 BEREKENEN VAN DE INSTRUMENTELE VARIABELEN

Omdat we ons richten op twee verschillende behandelingen – sancties en waarschuwingen – berekenden we twee verschillende instrumentele variabelen. Dat wil zeggen, we berekenden *strengheid* apart voor sancties en voor waarschuwingen. Meer specifiek berekenden we strengheid door het aantal keren dat een adviseur de respectievelijke behandeling heeft toegediend te delen door het totale aantal verantwoordingstaken dat de adviseur heeft afgehandeld. De individuele casus om wie het gaat telden we daarbij niet mee. Met andere woorden: we lieten de casus van persoon A buiten beschouwing als we berekenden met welke strengheid persoon A geconfronteerd werd. In formelere termen: de instrumentele variabele kwantificeert de strengheid van adviseur  $j$ , aan wie de WW-gerechtigde  $i$  werd toegewezen, voor alle casussen, behalve de casus van  $i$  zelf. Deze index wordt ook wel *leave-out mean* genoemd. We gebruikten deze methode om zeker te zijn dat onze instrumentele variabelen onafhankelijk zijn van de individuele casus waarvoor de variabele is berekend.

## 4.3 DATASET EN STEEKPROEF

Onze initiële dataset omvatte ongeveer 150.000 unieke casussen, dat wil zeggen afgehandelde verantwoordingstaken, uit de periode januari 2014 tot en met december 2017. Zoals eerder besproken, gaat het bij deze casussen om eerste verantwoordingstaken; WW-gerechtigden komen dus maar één keer voor in onze dataset. De casussen werden op 37 verschillende UWV-kantoren afgehandeld, door ongeveer 2.900 adviseurs. We lieten verantwoordingstaken die na 2017 zijn afgehandeld buiten beschouwing, omdat er in 2018 een beleidswijziging heeft plaatsgevonden (zie hoofdstuk 2). Deze beleidswijziging heeft ervoor gezorgd dat er bij een eerste overtreding doorgaans een waarschuwing wordt opgelegd. Hierdoor zijn data vanaf 2018 niet geschikt om de effecten van sancties te onderzoeken.

We hebben een aantal casussen uit deze initiële dataset verwijderd om tot een opgeschoonde basis-dataset te komen. Ten eerste sloten we casussen uit waar informatie over het kantoor (0,4% van de casussen) of over de begin- en einddatum van de uitkering ontbrak (6,0%). Ook verwijderden we casussen waar het proces van afhandeling meer dan 31 dagen duurde (1,8%). Verder verwijderden we casussen van WW-gerechtigden jonger dan 18 of ouder dan 64 jaar (4,0%), en casussen van WW-gerechtigden met een laatste baan in de sector Onderwijs en Overheid (1,3%). Voormalige werknemers uit deze sector hebben namelijk te maken met andere regelgeving vanuit UWV. Ten slotte verwijderden we casussen van adviseurs die in totaal minder dan 20 casussen hebben afgehandeld

(6,0%). Deze ondergrens moet ervoor zorgen dat de strengheid van adviseurs op voldoende observaties is gebaseerd.<sup>7</sup>

Na het toepassen van deze selectiecriteria bestond onze dataset uit 124.913 casussen, die door 1.488 adviseurs werden afgehandeld. Deze steekproef vormde de basis-dataset die we gebruikten om de strengheid van adviseurs te berekenen. Omdat data over arbeidsmarkuitkomsten alleen vanaf 2016 beschikbaar waren, moesten we deze dataset voor de analyse verder inperken, namelijk tot casussen die tussen januari 2016 en december 2017 zijn afgehandeld. Voor de analyse maakten we uiteindelijk gebruik van een dataset die bestond uit 51.848 unieke casussen, die zijn afgehandeld door 1.303 adviseurs.

Deze ingekorte dataset koppelden we aan informatie over uitkomsten in de twaalf maanden na overtreding. De informatie over uitkomsten was beschikbaar op maandelijkse basis. Voor elke casus beschikten we dus over informatie over uitkomsten in de eerste tot en met de twaalfde maand na de overtreding. Door de longitudinale structuur van onze dataset konden we het tijdsverloop van de uitkomstmaten onderzoeken.

#### 4.4 BESCHRIJVING VAN DE DATASET

Tabel 4.1 toont de kenmerken van WW-gerechtigden en casussen in onze analyse-dataset. WW-gerechtigden in deze dataset waren gemiddeld 41 jaar oud; ongeveer de helft was vrouw. Voor ongeveer driekwart van de steekproef was een opleidingsniveau bekend: een kwart van de WW-gerechtigden had een laag opleidingsniveau (basisonderwijs, vmbo, havo of vwo); ongeveer 40% had een middelbaar opleidingsniveau (mbo 1-mbo 4); 14% had een hoog opleidingsniveau (hbo of wo). We beschikten verder over informatie aangaande in welke sector iemand voor het laatst had gewerkt. In tabel 4.1 laten we alleen de zes grootste sectoren zien, die bij elkaar rond 75% van onze steekproef beslaan. De grootste sector in onze dataset is de sector Dienstverlening (18%). Op de tweede tot en met vijfde plek volgen de sectoren Handel (16%), Gezondheid (14%), Industrie (13%) en Uitzendbureau (12%). De sectorindeling van UWV volgt niet de Standaard Bedrijfsindeling (SBI) van het Centraal Bureau voor de Statistiek (CBS), maar de sectorindeling van de Belastingdienst. Dat is ook de reden dat er een sector Uitzendbureau te vinden is.

Tabel 4.1 laat verder zien dat de gemiddelde WW-gerechtigde in onze analyse-dataset in de laatste baan ongeveer € 2.100 per maand verdiende, en op het moment van overtreding nog tien maanden recht had op een WW-uitkering.<sup>8</sup> Ongeveer de helft van onze steekproef had werk in de maand voor overtreding. Dit kan ook werk naast de WW-uitkering betreffen. Ongeveer 90%

---

7 We kiezen de ondergrens in lijn met eerder onderzoek op dit gebied. Vaak wordt een ondergrens van 20-50 casussen gehanteerd.

8 De maximale WW-duur bedraagt 24 maanden. De hoge maximale waarde van 283 kan worden verklaard door foutieve data.

van onze steekproef had in de maand voor de overtreding een WW-betaling ontvangen. Er was dus maar een kleine groep die, bijvoorbeeld door verrekening met andere inkomsten, geen gebruik maakte van de WW-uitkering in de maand voor de overtreding.

**Tabel 4.1 Kenmerken van WW-gerechtigden en casussen in de analyse-dataset (N = 51.848)**

Kenmerk	Gemiddelde	SD	Min.	Max.
	(1)	(2)	(3)	(4)
Vrouw	0,53	0,5	0	1
Leeftijd	40,91	11,47	19	64
<i>Opleidingsniveau</i>				
Laag	0,24	0,43	0	1
Middelbaar	0,38	0,49	0	1
Hoog	0,14	0,35	0	1
Onbekend	0,24	0,43	0	1
<i>Sector</i>				
Dienstverlening	0,18	0,39	0	1
Handel	0,16	0,37	0	1
Zorg	0,14	0,34	0	1
Industrie	0,13	0,33	0	1
Uitzendbureau	0,12	0,33	0	1
Transport	0,08	0,27	0	1
Andere	0,20	0,40	0	1
Maandloon laatste baan	2.099	1.057	0	4.718
Resterende WW-maanden	10	9	0	283
Werkzaam ( <i>t</i> -1)	0,48	0,5	0	1
WW-uitkering ( <i>t</i> -1)	0,87	0,34	0	1
<i>Consequentie afhandeling</i>				
Verwijtbaar	0,58	0,49	0	1
Sanctie	0,25	0,43	0	1
Waarschuwing	0,18	0,38	0	1
Geen consequentie	0,15	0,36	0	1
<i>Jaar afhandeling</i>				
2016	0,37	0,48	0	1
2017	0,63	0,48	0	1
<i>Instrumentele variabelen</i>				
Instrument sanctie	0,30	0,23	0	1
Instrument waarschuwing	0,15	0,18	0	1

Tabel 4.1 laat ook zien hoe de casussen in onze analyse-dataset zijn beoordeeld. In ongeveer 60% van de casussen was er volgens de adviseur sprake van verwijtbaar gedrag. Verwijtbaarheid kan een sanctie, een waarschuwing of geen consequentie tot gevolg hebben. Een kwart van alle casussen kreeg een sanctie, 18% kreeg een waarschuwing en bij 15% van de casussen is er geen consequentie geregistreerd. Dat betekent dat bij verwijtbaarheid in 43% van de casussen

een sanctie volgde, in 31% een waarschuwing en in 26% geen consequentie. Tabel 4.1 geeft tevens de verdeling van casussen over de jaren weer. Rond 60% van de casussen kwam uit 2017 en 40% kwam uit 2016.

Ten slotte toont tabel 4.1 beschrijvende informatie over onze instrumentele variabelen. Omdat we strengheid berekenden als toepassing van een behandeling – sanctie of waarschuwing – over alle afgehandelde casussen, liggen de waarden tussen 0 (behandeling nooit toegepast) en 1 (behandeling altijd toegepast). Ons instrument voor sancties heeft een gemiddelde van 0,30 en ons instrument voor waarschuwingen een gemiddelde van 0,15. Dit betekent dat de gemiddelde WW-gerechtigde in onze dataset met een adviseur te maken kreeg die in 30% van de gevallen een sanctie oplegt, en in 15% van de gevallen een waarschuwing. Zoals eerder genoemd controleerden we in onze analyses voor kantoor en maand van afhandeling. Dit betekent dat we alleen gebruikmaakten van variatie in strengheid binnen hetzelfde kantoor en binnen dezelfde maand. Figuur B.1 in bijlage B laat zien hoeveel variatie er in strengheid zit als we corrigeren voor de verschillen tussen kantoren en maanden. Het blijkt dat WW-gerechtigden binnen hetzelfde kantoor en dezelfde maand met verschillende strengheden te maken konden krijgen.

Tabel 4.2 toont kenmerken van de 1.303 UWV-adviseurs in onze analyse-dataset. Helaas was er maar weinig informatie verkrijgbaar over adviseurs. Wel konden we het aantal dienstmaanden van een adviseur binnen de periode januari 2014 tot en met december 2017 berekenen. Dit is de periode die we ook hebben gehanteerd om onze instrumenten te berekenen. We berekenden de dienstmaanden door te kijken naar het eerste en het laatste moment waarop een adviseur verantwoordingsstaken had afgehandeld, en vervolgens het aantal maanden tussen die twee momenten te bepalen. Het gemiddelde lag op 38 maanden; het overgrote deel van de adviseurs was over de hele periode betrokken bij het afhandelen van verantwoordingsstaken.

Tabel 4.2 laat verder zien dat de gemiddelde adviseur in totaal 88 verantwoordingsstaken had afgehandeld. Dit komt neer op gemiddeld 2,5 verantwoordingsstaken per dienstmaand. In lijn met onze selectiecriteria (zie vorige paragraaf) waren er geen adviseurs die in totaal minder dan 20 casussen hadden afgehandeld. Figuur B.2 in bijlage B laat de verdeling van het aantal verantwoordingsstaken zien. De verdeling is rechtsscheef, waarbij de meeste adviseurs minder dan 100 taken in totaal en minder dan 5 taken per maand hadden afgehandeld.

**Tabel 4.2 Kenmerken van UWV-adviseurs in de analyse-dataset (N = 1.303)**

<b>Kenmerk</b>	<b>Gemiddelde</b>	<b>SD</b>	<b>Min.</b>	<b>Max.</b>
	<b>(1)</b>	<b>(2)</b>	<b>(3)</b>	<b>(4)</b>
Dienstmaanden	38	12	2	48
Aantal verantwoordingsstaken totaal	88	114	20	1.639
Aantal verantwoordingsstaken per maand	2,54	2,93	0	36

*Noot:* Dienstmaanden zijn dienstmaanden binnen de periode januari 2014 tot en met december 2017. Per maand betekent gedeeld door het aantal dienstmaanden.

#### 4.5 UITKOMSTVARIABLEN

In deze paragraaf beschrijven we hoe we de uitkomsten nalevingsgedrag, werkzoekgedrag, arbeidsparticipatie en uitkeringsafhankelijkheid naar concrete uitkomstvariabelen hebben vertaald.

Informatie over nalevingsgedrag kwam uit de systemen van UWV. We berekenden vier variabelen. We stelden vast of een persoon zich in de twaalf maanden na overtreding *opnieuw moest verantwoorden* voor te weinig doorgegeven zoekactiviteiten (1/0 variabele). Daarnaast stelden we vast of er in het kader van de eerstvolgende verantwoordingstaak een *sanctie* of een *waarschuwing* is opgelegd, of dat er sprake was van *niet-verwijtbaar gedrag* (allemaal 1/0 variabelen).

Ook de informatie over werkzoekgedrag kwam uit de systemen van UWV. Voor elke maand observeerden we het aantal doorgegeven *sollicitatieactiviteiten*. Doorgeven van sollicitatieactiviteiten hoeft en kan alleen tijdens een lopende uitkering. We deelden daarom het totale aantal sollicitatieactiviteiten in de twaalf maanden na overtreding door het aantal maanden dat iemand in die tijd een lopende uitkering had.

Informatie over arbeidsparticipatie kwam uit de Polisadministratie. We berekenden eerst drie cumulatieve variabelen over de twaalf maanden na overtreding. Ten eerste bepaalden we of een persoon in de twaalf maanden na overtreding *ooit werkzaam* was geweest; dat wil zeggen of een persoon in minstens één maand positieve looninkomsten had gehad (1/0 variabele). Ten tweede bepaalden we het *totale looninkomen*. Ten derde bepaalden we het totale aantal *gewerkte uren* in de twaalf maanden na overtreding. Verder berekenden we vier variabelen op maandelijks niveau. Zo keken we per maand of iemand wel of niet *werkzaam* was (1/0 variabele), en of dit werk onder een *tijdelijk contract* (1/0 variabele) of een *vast contract* plaatsvond (1/0 variabele). Ook het *looninkomen* bepaalden we op maandelijks niveau.

Informatie over uitkeringsafhankelijkheid kwam uit de systemen van UWV. Om uitkeringsafhankelijkheid vast te stellen keken we naar de *ontvangen WW-uitkering*, zowel cumulatief over de twaalf maanden na overtreding als op maandelijks niveau. We keken naar uitkeringsbedragen omdat zowel deeltijds als voltijduitstroom uit de WW daarin terug te zien is. Naast de WW-uitkering beheert UWV ook de ziektewetuitkering (ZW-uitkering). Uitkeringsgerechtigden kunnen aanspraak maken op deze uitkering in het geval van werkloosheid in combinatie met ziekte. Het is denkbaar dat WW-gerechtigden de Ziektewet instromen om handhaving in de toekomst te voorkomen (zie paragraaf 1.2). We keken daarom ook naar de *ontvangen ZW-uitkering*, zowel cumulatief als op maandelijks niveau.

#### 4.6 AANNAMES

Om IVA te kunnen gebruiken moet er worden voldaan aan een aantal aannames (zie box 3.2). In deze paragraaf beschrijven we hoe we deze aannames hebben getoetst. Dit deden we deels met statistische toetsen, en deels door op



andere manieren aanwijzingen te verzamelen. Hierbij volgen we grotendeels de stappen en suggesties uit recent empirisch werk (Arni & Schiprowski, 2019; Bhuller et al., 2020).

### **Aanname 1: Relevantie**

De eerste aanname is dat onze instrumentele variabelen ( $Z$ ) de respectievelijke behandelingen ( $X$ ), in ons geval een sanctie of een waarschuwing, goed voorspellen. We toetsten deze aanname door middel van een regressieanalyse, waarbij we een dummyvariabele voor de behandeling als afhankelijke variabele, en onze instrumentele variabelen als onafhankelijke variabelen gebruikten.<sup>9</sup> In alle regressies namen we controlevariabelen op voor kantoor en maand. Verder controleerden we ook voor achtergrondkenmerken van WW-gerechtigden. Tabel C.1 in bijlage C geeft de resultaten weer.

Uit de resultaten bleek dat onze instrumentele variabelen een sterk en positief verband hadden met de respectievelijke behandeling. Het bleek ook dat de regressiecoëfficiënten bij toevoeging van achtergrondkenmerken als controlevariabelen niet sterk veranderden. Dit is een eerste indicatie dat de toedeling van WW-gerechtigden toevallig plaatsvindt. Om de relevantie van een instrumentele variabele te beoordelen wordt vaak ook gekeken naar de  $F$ -statistiek. In de literatuur wordt hiervoor meestal een grenswaarde van 10 gehanteerd. Alle  $F$ -statistieken lagen ver boven deze grens, wat onze bevinding dat we met relevante instrumenten te maken hebben verder onderbouwt.

### **Aanname 2: Onafhankelijkheid**

De tweede aanname is dat onze instrumentele variabelen ( $Z$ ) niet samenhangen met andere factoren ( $U$ ) die van invloed kunnen zijn op uitkomsten ( $Y$ ). Aan deze aanname wordt voldaan als casussen zo veel mogelijk aselekt worden toegedeeld aan adviseurs. In dat geval kan strengheid namelijk niet (of in ieder geval: niet sterk) samenhangen met bepaalde achtergrondkenmerken van WW-gerechtigden.

Om de tweede aanname te toetsen hebben we als eerste stap contact opgenomen met UWV-kantoren om een beeld te krijgen van de werkwijze omtrent de toedeling van casussen. Het bleek dat managers van de drie kantoren die bij de eerste uitrol van de nieuwe werkwijze in 2014 betrokken waren, alsmede de producteigenaar van dit proces, allen aangaven dat casussen normaal gesproken toevallig verdeeld worden over adviseurs. Zij geven aan dat er veelal met geautomatiseerde lijsten gewerkt wordt, die zonder specifieke volgorde evenredig verdeeld worden over alle op een dag beschikbare medewerkers. Zij geven ook aan dat ruilen van casussen niet plaatsvindt. Op sommige kantoren kan er sprake zijn van een specialisering van adviseurs op bepaalde groepen uitkeringsgerechtigden, bijvoorbeeld uitkeringsgerechtigden uit een bepaalde sector. In dat geval kan de toedeling van casussen ook binnen een

---

9 Dit is gelijk aan de *first stage* van een *two-stage least squares* regressie.

team adviseurs met dezelfde specialisering plaatsvinden. Uit de gesprekken met UWV-medewerkers ontstond dus het beeld dat casussen grotendeels aselekt worden toegedeeld aan adviseurs, en dat adviseurs geen invloed hebben op deze toedeling.

We konden ook onderzoeken of er sprake is van toevallige toedeling door de data verder te verkennen. Dit deden we door te kijken of er statistische verbanden waren tussen onze instrumentele variabelen en achtergrondkenmerken van WW-gerechtigden. Als we geen duidelijke verbanden vinden, is het aannemelijk dat de toedeling toevallig plaatsvindt. Verder keken we ook naar het verband tussen de behandeling zelf (sanctie of waarschuwing) en achtergrondkenmerken. Hier verwachtten we juist wel verbanden te vinden, bijvoorbeeld dat bepaalde groepen een grotere of kleinere kans hebben op een sanctie of waarschuwing.

We onderzochten bovenstaande verbanden door middel van een regressie-analyse, waarbij we óf onze instrumentele variabelen óf een dummyvariabele voor de behandeling als afhankelijke variabele, en een set achtergrondkenmerken van WW-gerechtigden als onafhankelijke variabelen gebruikten. In alle regressies namen we controlevariabelen op voor kantoor en maand. Tabel C.2 in bijlage C geeft de resultaten weer.

Ten eerste vonden we, zoals verwacht, veelal statistisch significante verbanden tussen de behandeling en bepaalde achtergrondkenmerken (kolom 1 en 3). Vrouwen maken bijvoorbeeld meer kans op een sanctie. Ook neemt de kans op een sanctie met stijgende leeftijd af, en stijgt de kans op een waarschuwing bij een hoog opleidingsniveau. Deze effecten hoeven niet te betekenen dat adviseurs van UWV bepaalde groepen bevoordelen of benadelen. Het is waarschijnlijker dat deze observeerbare kenmerken samenhangen met niet-observeerbare karakteristieken, die wederom gepaard gaan met een grotere of kleine kans op een bepaalde behandeling. Deze exercitie laat zien dat gesanctioneerde en gewaarschuwde WW-gerechtigden inderdaad geen aselechte groep zijn, maar bepaalde onderscheidende kenmerken of eigenschappen hebben. Dit was voor ons een van de redenen om te kiezen voor IVA (zie hoofdstuk 3). Deze resultaten onderbouwen onze keuze verder.

De verbanden met achtergrondkenmerken verdwenen nagenoeg compleet als we onze instrumentele variabelen gebruikten als afhankelijke variabelen (kolom 2 en 4). Voor ons instrument voor sancties konden bijvoorbeeld bijna geen statistisch significante verbanden meer worden aangetoond. In de enkele gevallen waar er wel een verband leek te zijn, was de sterkte van de samenhang zwak. Voor ons instrument voor waarschuwingen bleven er meer verbanden statistisch significant, maar ook hier was de sterkte van de samenhang zwak.

Deze resultaten suggereren dat strengere adviseurs binnen een kantoor en maand niet structureel casussen van andere WW-gerechtigden toegewezen krijgen dan minder strenge adviseurs. Dit onderbouwt de kwalitatieve bevindingen uit de gesprekken met deskundigen binnen UWV. Casussen worden dus waarschijnlijk altijd (of in ieder geval meestal) toevallig toegedeeld aan adviseurs.

Dat we hier en daar wel zwakke verbanden vonden tussen onze instrumenten en bepaalde achtergrondkenmerken kan verschillende oorzaken hebben. Zo zou het kunnen dat er af en toe (bijvoorbeeld bij sommige kantoren, in sommige maanden) afgeweken werd van het standaardproces. Ook zou het kunnen dat de teams die deels zijn gespecialiseerd in bepaalde groepen uitkeringsgerechtigden, uit juist strenge of minder strenge adviseurs bestonden. Omdat deze afwijkingen niet zijn vastgelegd in de registratiesystemen van UWV, kunnen we hier echter alleen over speculeren. Zoals eerder vermeld, zullen we in onze statistische modellen voor achtergrondkenmerken van WW-gerechtigden controleren. We gaan ervan uit dat we elementen van mogelijk niet-aselecte toedeling hierdoor kunnen ondervangen.

### Aanname 3: Exclusie

De derde aanname is dat onze instrumentele variabelen ( $Z$ ) alleen de strengheid van een adviseur en niet ook andere keuzes of gedragingen reflecteren die de uitkomsten ( $Y$ ) mogelijk beïnvloeden. Voor deze aanname is het belangrijk dat adviseurs zich, afgezien van hun strengheid, niet op andere manieren sturend gedragen. Het zou bijvoorbeeld problematisch zijn als strengere adviseurs ook relatief sterk geneigd zouden om bepaalde dienstverlening op te leggen of aan te bieden.

Om deze derde aanname te toetsen keken we eerst naar procesdocumenten en werkinstructies. Uit deze documentatie bleek dat verantwoordingsstaken normaliter zonder persoonlijk contact worden afgehandeld. Persoonlijk contact zou alleen moeten plaatsvinden als de WW-gerechtigde niet reageert op het verzoek van UWV om zich te verantwoorden. In deze zogenoemde hoorwederhoorgesprekken bespreekt de adviseur met de uitkeringsgerechtigde het gegeven dat er een overtreding van de sollicitatieplicht heeft plaatsgevonden, en geeft de adviseur de uitkeringsgerechtigde de gelegenheid alsnog te reageren. Volgens de instructies van UWV zijn deze gesprekken primair bedoeld om de eventuele verwijtbaarheid van de overtreding vast te kunnen stellen. Echter, de adviseur heeft volgens de instructies ook de ruimte om samen met de WW-gerechtigde te kijken naar de inzet van verdere dienstverlening.

Het was zodoende nodig om te onderzoeken of onze instrumentele variabelen samenhangen met de extra dienstverlening die vaak door adviseurs wordt ingezet. Er zijn drie soorten dienstverlening: een *workshop* of *training*, een *coachingsgesprek* en *onthefving van de sollicitatieplicht*. Als de toedeling aan een strenge of minder strenge adviseur niet veel uitmaakt voor de kans om in de maanden na overtreding verdere dienstverlening te ontvangen, dan geeft dat een indicatie dat strengheid niet ook andere keuzes van een adviseur reflecteert.

We onderzochten of er statistische verbanden waren door middel van een regressieanalyse, waarbij we onze instrumentele variabelen als afhankelijke variabele gebruikten, en drie dummyvariabelen voor de drie soorten dienstverlening als onafhankelijke variabelen. Deze drie dummyvariabelen gaven aan of de WW-gerechtigde in de zes maanden na overtreding in aanraking was gekomen met de respectievelijke dienstverlening. In alle regressies namen we controlevariabelen op voor kantoor, maand en achtergrondkenmerken. Verder

keken we ook naar de samenhang tussen de behandeling (sanctie of waarschuwing) en de kans op aanvullende dienstverlening. Hier verwachtten we namelijk wel verbanden te vinden. Tabel C.3 in bijlage C geeft de resultaten weer.

Zoals verwacht was de kans op aanvullende dienstverlening gecorreleerd met de behandeling (kolom 3 en 5). Zo werden gewaarschuwde WW-gerechtigden in de zes maanden na overtreding vaker naar een workshop gestuurd en kregen zij vaker een ontheffing van sollicitatieactiviteiten. Gesanctioneerde WW-gerechtigden maakten juist minder kans op een ontheffing. Toen we onze instrumenten als afhankelijke variabelen gebruikten, verdwenen de statistische verbanden (kolom 4 en 6). Dit suggereert dat de strengheid van adviseurs niet aantoonbaar gerelateerd is aan de neiging van die adviseurs om aanvullende dienstverlening op te leggen.

We concluderen dat het aannemelijk is dat onze instrumentele variabelen alleen strengheid reflecteren, en niet ook andere keuzes of gedragingen van een adviseur.

#### **Aanname 4: Monotoniciteit**

Deze aanname gaat ervan uit dat de kans op een behandeling ( $X$ ) voor iedereen toeneemt naarmate de adviseur strenger is. We kunnen deze aanname deels direct testen door nog een keer te kijken naar de analyse voor aanname 1 (relevantie). Als de samenhang tussen de behandeling en onze instrumentele variabelen qua richting niet verandert als we naar verschillende sub-samples kijken, dan hebben we een aanwijzing dat er sprake is van monotoniciteit.

Tabel C.4 in bijlage C toont de resultaten van deze exercitie. Om het overzichtelijk te houden beperken we ons voor deze analyse tot het instrument voor sancties. In de eerste rij zien we ter vergelijking het resultaat uit tabel C.1: een sterk positief verband tussen instrumentele variabele en de bijbehorende behandeling. Tabel C.4 laat in kolom 1 zien dat dit resultaat niet wezenlijk veranderde als we onze steekproef voor de schatting telkens beperken tot verschillende sub-samples. Kolom 2 laat zien dat het resultaat ook niet wezenlijk veranderde wanneer we een zogenoemde *reverse sample* gebruiken om de instrumentele variabele te berekenen (een voorbeeld: we berekenden strengheid op basis van de sub-sample mannen en gebruiken deze strengheid vervolgens als instrumentele variabele voor de schatting met de sub-sample vrouwen).

Op basis van deze resultaten concluderen we dat er sterke aanwijzingen zijn dat de strengheid van adviseurs maar één kant op werkt, en de kans op behandeling dus voor iedereen toeneemt naarmate een adviseur strenger is.

In dit hoofdstuk presenteren we de resultaten van onze statistische analyses. Zoals in hoofdstuk 4 beschreven, stellen we effecten vast door *gesanctioneerde* en *gewaarschuwde* WW-gerechtigden te vergelijken met WW-gerechtigden wier gedrag als *niet-verwijtbaar* is beoordeeld. Als we in dit hoofdstuk spreken over een *effect*, dan gaat het dus om het effect van een sanctie of waarschuwing ten opzichte van een situatie waarin er geen maatregel werd opgelegd. Onze analyse-dataset omvatte 51.848 casussen (zie hoofdstuk 4 voor een gedetailleerde beschrijving, en tabel 4.1 voor een samenvatting). We bespreken de resultaten in drie stappen:

- In paragraaf 5.1 laten we resultaten zien voor cumulatieve uitkomsten, gemeten over de twaalf maanden na de overtreding. Dat betekent bijvoorbeeld dat we alle looninkomsten van deze twaalf maanden bij elkaar optellen en vervolgens toetsen of er een verschil zit tussen de totale inkomsten van gesanctioneerde en niet-gesanctioneerde WW-gerechtigden.
- In paragraaf 5.2 laten we zien hoe effecten zich over de twaalf maanden na de overtreding ontwikkelden.
- In paragraaf 5.3 laten we zien hoe effecten verschillen voor WW-gerechtigden met verschillende opleidingsniveaus. Het zou namelijk kunnen zijn dat een sanctie verschillende effecten heeft voor verschillende groepen mensen.

### 5.1 EFFECTEN NA TWAALF MAANDEN

Zoals we in hoofdstuk 4 beschreven, onderzochten we de effecten van sancties en waarschuwingen op vier typen variabelen:

- Naleving*: waren mensen geneigd zich beter aan de regels te houden (dat wil zeggen: kregen ze minder vaak te maken met een nieuwe verantwoordingszaak)?
- Werkzoekgedrag*: gingen mensen intensiever op zoek naar nieuw werk?
- Arbeidsparticipatie*: waren mensen succesvoller in het vinden van nieuw werk?
- Uitkeringsafhankelijkheid*: werden mensen minder afhankelijk van een uitkering van UWV?

Tabel 5.1 toont de geschatte effecten. De tabel kan als volgt worden gelezen: De linkerkolom toont onze uitkomstmaten. Elke rij bevat alle resultaten voor één uitkomstmaat. Kolom 1 toont het gemiddelde voor de hele steekproef. Door effecten te vergelijken met het steekproefgemiddelde kan men een indruk krijgen van de relatieve grootte van de effecten. Kolom 2 en 3 tonen de effecten van respectievelijk sancties en waarschuwingen. In deze kolommen rapporteren we steeds coëfficiënten, standaardfouten (tussen ronde haakjes) en  $p$ -waarden (tussen vierkante haakjes). Kolom 4 toont het aantal observaties waarop de resultaten in die rij gebaseerd zijn.

**Tabel 5.1 Effecten op cumulatieve uitkomsten in de twaalf maanden na overtreding**

Uitkomstmaten	Gemiddeld	Effect sanctie	Effect waarschuwing	N
	(1)	(2)	(3)	(4)
<b>Paneel A: Naleving</b>				
Nieuwe verantwoordingsstaak (1/0)	0,283	-0,110*** (0,031) [< 0,001]	0,006 (0,030) [0,834]	51.848
<i>Waarvan:</i>				
Niet-verwijtbaar (1/0)	0,159	-0,158*** (0,025) [< 0,001]	-0,066* (0,026) [0,010]	51.848
Waarschuwing (1/0)	0,033	-0,031** (0,011) [0,006]	-0,001 (0,014) [0,959]	51.848
Sanctie (1/0)	0,058	0,076*** (0,016) [< 0,001]	0,074*** (0,015) [< 0,001]	51.848
<b>Paneel B: Werkzoekgedrag</b>				
Gemiddeld aantal activiteiten/maand	2,151	0,104 (0,105) [0,321]	0,878*** (0,105) [< 0,001]	46.751
<b>Paneel C: Arbeidsparticipatie</b>				
Ooit werkzaam (1/0)	0,803	-0,027 (0,025) [0,279]	0,031 (0,025) [0,210]	51.848
Looninkomen (€)	14.575	-1.364 (776) [0,079]	-371 (961) [0,700]	51.848
Uren gewerkt	949	-42 (43) [0,331]	-1 (51) [0,981]	51.848

Uitkomstmaten	Gemiddeld	Effect sanctie	Effect waarschuwing	N
	(1)	(2)	(3)	
<b>Paneel D: Uitkeringsafhankelijkheid</b>				
Ontvangen WW-uitkering (€)	6.373	-498 (365) [0,173]	293 (440) [0,506]	51.848
Ontvangen ZW-uitkering (€)	77	-3 (22) [0,902]	17 (25) [0,496]	51.848

*Noot:* Elke rij toont de resultaten van een *two-stage least squares* (2SLS) regressie. Afhankelijke variabelen staan links. Kolom 1 geeft het steekproefgemiddelde weer. Kolom 2 toont het effect van een sanctie en kolom 3 van een waarschuwing. In het 2SLS-model dient strengheid als instrumentele variabele voor de daadwerkelijke behandeling. Strengheid wordt gemeten als de *leave-out mean* van een adviseur. Alle regressies controleren voor kantoor, maand en achtergrondkenmerken. Standaardfouten worden geclusterd op niveau van de adviseurs ( $N = 1.303$ ) en gerapporteerd tussen ronde haakjes;  $p$ -waarden worden gerapporteerd tussen vierkante haakjes.

\*  $p < 0,05$ .

\*\*  $p < 0,01$ .

\*\*\*  $p < 0,001$ .

### Effecten op naleving

Paneel A in tabel 5.1 toont de resultaten voor naleving. De variabele *nieuwe verantwoordingstaak* meet of de betreffende WW-gerechtigde zich in de twaalf maanden na overtreding opnieuw heeft moeten verantwoorden voor te weinig doorgegeven sollicitatieactiviteiten. Kolom 1 laat zien dat bijna een derde van onze steekproef weer een nieuwe verantwoordingstaak heeft gekregen. Kolom 2 toont het geschatte effect van een sanctie op deze uitkomstmaat. Omdat onze uitkomstmaat gecodeerd is als 0 (nee) en 1 (ja) is de coëfficiënt van  $-0,11$  te interpreteren als een negatief effect van 11 procentpunt. Dit betekent dat het opleggen van een sanctie de kans op een nieuwe verantwoordingstaak met 11 procentpunt verminderde. Aangezien de gemiddelde kans 28% was, komt dit overeen met een vermindering van ongeveer 40% ( $-11,0/28,3$ ). Dit effect was significant ( $p < 0,001$ ). Kolom 3 toont het geschatte effect van een waarschuwing. De coëfficiënt was klein (0,6 procentpunt) en niet significant. We vonden dus geen bewijs dat waarschuwingen invloed hebben op de kans om een nieuwe verantwoordingstaak te krijgen.

Vervolgens onderzochten we wat voor gevolgen deze nieuwe verantwoordingstaken hadden voor WW-gerechtigden. Een nieuwe verantwoordingstaak kan als *niet-verwijtbaar* worden beoordeeld of, in het geval van verwijtbaarheid, leiden tot een *nieuwe sanctie* of een *nieuwe waarschuwing*. We onderzochten vervolgens of een sanctie of waarschuwing invloed heeft op de kans om bij de eerstvolgende verantwoordingstaak weer een sanctie of waarschuwing te krijgen. Kolom 1 liet zien dat gemiddeld 28% van onze streekproef een nieuwe verantwoordingstaak had ontvangen. Ongeveer 60% van deze nieuwe verantwoordingstaken werd als niet-verwijtbaar aangemerkt (15,9/28,3); ongeveer

12% leidde tot een waarschuwing (3,3/28,3); ongeveer 20% leidde tot een sanctie (5,8/28,3).<sup>10</sup>

De effectschattingen in kolom 2 en 3 geven aan in hoeverre het krijgen van een (eerste) sanctie of (eerste) waarschuwing de kans op een (nieuwe) sanctie of (nieuwe) waarschuwing veranderde. Hoewel gesanctioneerde mensen zoals gezegd een kleinere kans hadden op een nieuwe verantwoordingstaak, hadden gesanctioneerde mensen, als ze toch een nieuwe verantwoordingstaak kregen, een grotere kans op een nieuwe sanctie. Ze hadden daarentegen een kleinere kans op een waarschuwing en een kleinere kans om als niet-verwijtbaar beoordeeld te worden. Voor waarschuwingen vonden we een soortgelijk patroon. Als gewaarschuwde mensen een nieuwe verantwoordingstaak kregen, hadden ze een kleinere kans om als niet-verwijtbaar beoordeeld te worden, en een grotere kans op een sanctie. Al deze bevindingen passen goed bij het officiële beleid van UWV: bij vervolgovertredingen wordt bij voorkeur een sanctie opgelegd.

We trekken de volgende tussentijdse conclusie: Sancties hadden een gunstig effect op naleving; de kans op een vervolgovertreding nam af. Waarschuwingen hadden daarentegen geen aantoonbaar effect op naleving. Als mensen een sanctie of waarschuwing hadden gekregen, kregen ze bij vervolgovertredingen eerder een sanctie, wat consistent is met het beleid dat UWV voert. We geven een verdere interpretatie in paragraaf 6.2.

### Effecten op werkzoekgedrag

We observeerden de sollicitatieactiviteiten die WW-gerechtigden registreerden in de *Werkmap*, het online portal van UWV. Dit kunnen verschillende activiteiten zijn, zoals solliciteren op een vacature, een sollicitatiegesprek voeren bij een potentiële werkgever, een netwerkgesprek voeren of een workshop volgen. Werkzoekgedrag is geoperationaliseerd als het aantal sollicitatieactiviteiten per maand (alleen berekend over maanden waarin mensen daadwerkelijk een WW-uitkering hadden; zie paragraaf 2.3 voor details).

Paneel B in tabel 5.1 toont de resultaten voor werkzoekgedrag. De gemiddelde WW-gerechtigde in onze steekproef registreerde, over de hele observatieperiode, ongeveer 2,4 activiteiten per maand. Dit gemiddelde lag dus lager dan de vier activiteiten per vier weken die UWV voorschrijft. Dit valt deels te verklaren doordat sommige WW-gerechtigden te weinig activiteiten doorgeven, maar ook doordat sommige WW-gerechtigden door vrijstellingen tijdelijk minder of helemaal niet hoeven solliciteren.

Sancties hadden geen significant effect op het aantal geregistreerde activiteiten. Er is zodoende geen bewijs dat sancties WW-gerechtigden motiveerden om intensiever te solliciteren. Waarschuwingen leidden daarentegen wel tot een significante toename in werkzoekgedrag ( $p < 0,001$ ). De effectschatting van 0,88 geeft aan dat het opleggen van een waarschuwing zorgde voor gemiddeld bijna

<sup>10</sup> De percentages tellen niet bij elkaar op tot 100% omdat er ook enkele gevallen zijn waarbij geen consequentie is geregistreerd.



één extra activiteit per maand. Dit komt overeen met een positief effect van ongeveer 40% ten opzichte van het steekproefgemiddelde.

We trekken de volgende tussentijdse conclusie: Er is geen bewijs dat sancties mensen stimuleren om harder te zoeken naar werk. Waarschuwingen doen dat wel. We geven verdere duiding in paragraaf 6.2. Daar gaan we ook in op de vraag hoe deze bevindingen (waarschuwingen zorgen voor een toename in sollicitaties) zich verhouden tot de bevindingen over naleving (sancties zorgen voor betere naleving).

#### **Effecten op arbeidsparticipatie**

Paneel C in tabel 5.1 toont de resultaten voor arbeidsparticipatie. We kijken eerst naar de effecten op de uitkomstmaat *ooit werkzaam*. Deze variabele meet of er in de twaalf maanden na overtreding ooit looninkomsten zijn verdiend.<sup>11</sup> Kolom 1 laat zien dat 80% van onze steekproef in twaalf maanden minstens één maand werkzaam is geweest. We vonden geen bewijs dat de kans om ooit werkzaam te zijn groter of kleiner werd door het krijgen van een sanctie of een waarschuwing.

Arbeidsparticipatie uit zich ook in inkomsten en gewerkte uren. Als volgende stap kijken we daarom naar de effecten op de *verworven looninkomsten* en op het *aantal gewerkte uren* in de twaalf maanden na overtreding. Het gaat ook hier om een opsomming van alle maandelijkse bedragen/uren over die twaalf maanden. De gemiddelde WW-gerechtigde in onze steekproef heeft in de observatieperiode ongeveer € 15.000 verdiend en 950 uren gewerkt. De effectschattingen suggereren dat het opleggen van een sanctie leidt tot € 1.400 minder looninkomsten en 42 minder gewerkte uren. Hoewel beide effecten niet verwaarloosbaar zijn, is de foutmarge rond deze schattingen groot en zijn deze effecten niet statistisch significant. De effectschattingen van waarschuwingen waren veel kleiner, en waren ook niet significant.

We trekken de volgende tussentijdse conclusie: We vonden geen bewijs dat sancties en waarschuwingen ervoor zorgden dat mensen (a) vaker weer aan het werk gaan, (b) meer verdienen, en (c) meer uren werken. Het effect van sancties op looninkomsten was weliswaar niet significant, maar wel opmerkelijk: gemiddeld verdienen gesanctioneerde mensen € 1.400 per jaar *minder* dan niet-gesanctioneerde mensen. Dit opmerkelijke patroon zullen we verder verkennen in paragraaf 5.2, waar we kijken naar de tijdsdynamiek van effecten. Daar zal blijken dat er in bepaalde maanden inderdaad negatieve effecten van sancties op arbeidsparticipatie optreden.

#### **Effecten op uitkeringsafhankelijkheid**

Paneel D in tabel 5.1 toont de resultaten voor uitkeringsafhankelijkheid. Omdat zowel deeltijd- als voltijduitstroom uit de uitkering leidt tot lagere uitkeringsbedragen, onderzochten we de bedragen die door UWV zijn overgemaakt in de

---

11 We hebben geen gegevens over inkomsten uit zelfstandig werk.

twaalf maanden na overtreding. We richtten ons niet alleen op de WW-uitkering, maar ook op de ziektewetuitkering (ZW-uitkering). We deden dit omdat eerder onderzoek suggereert dat sanctiestructuren en handhaving mensen kunnen aanzetten tot instroom in andere, minder restrictieve uitkeringen (zie bijvoorbeeld McVicar, 2008).

De gemiddelde WW-gerechtigde in onze steekproef ontving in de observatieperiode ongeveer € 6.400 aan WW-uitkering en ongeveer € 80 aan ZW-uitkering. Na een sanctie ontvingen mensen ongeveer € 500 (of 8%) minder WW-uitkering; na een waarschuwing ontvingen mensen ongeveer € 300 (of 5%) meer WW-uitkering. Beide effecten waren niet statistisch significant. De effecten van sancties en waarschuwingen op ontvangen ZW-uitkeringen waren verwaarloosbaar.

Dat we een negatief (maar niet significant) effect vonden van sancties op de ontvangen WW-uitkering, heeft waarschijnlijk te maken met het financiële gevolg van deze maatregel in de vorm van een korting op de uitkering. In paragraaf 5.2, waar we kijken naar de dynamiek van effecten over de tijd, zal de financiële impact van een sanctie beter zichtbaar worden.

We trekken de volgende tussentijdse conclusie: Sancties noch waarschuwingen hadden significant effecten op uitkeringsafhankelijkheid. Ook waren er geen aanwijzingen dat sancties of waarschuwingen leiden tot een instroom in de Ziektewet.

## 5.2 EFFECTEN OVER DE TIJD

In deze paragraaf richten we ons op de tijdsdynamiek van effecten. In plaats van cumulatieve uitkomsten, zoals in de vorige paragraaf, kijken we nu naar de uitkomstvariabelen per maand. Door te kijken naar maandelijkse uitkomsten kunnen we inzichtelijk maken hoe effecten zich over de maanden ontwikkelen. De uitkomstmaten in deze paragraaf hebben alleen betrekking op de domeinen arbeidsparticipatie en uitkeringsafhankelijkheid.<sup>12</sup> De observatieperiode blijft onveranderd; onze analyse richt zich dus nog steeds op de eerste twaalf maanden na overtreding.

De figuren in deze paragraaf tonen tijd op de horizontale as. De horizontale as beslaat steeds 14 maanden (van maand -1 tot en met maand 12, waarbij maand 0 de maand is waarin de verantwoordingsstaak is geregistreerd).<sup>13</sup> De geschatte effecten (verticale as) van een sanctie zijn in een zwarte, doorgetrokken lijn weergegeven; de geschatte effecten van een waarschuwing zijn weergegeven

12 We bekijken effecten op naleving niet op maandelijks niveau omdat we ook geïnteresseerd zijn in hoeverre het opleggen van een sanctie of waarschuwing de beoordeling van een nieuwe verantwoordingsstaak beïnvloedt. Omdat de beoordeling soms niet in dezelfde maand plaatsvindt als de overtreding, kunnen we uitkomsten beter over een langere periode bekijken. Ook werkzoekgedrag bekijken we niet op maandelijks niveau omdat de registratie per vier weken niet overeenkomt met kalendermaanden.

13 De observatieperiode is dus voor elk individu in onze steekproef dermate aangepast dat maand 0 altijd de maand van de verantwoordingsstaak aangeeft.

in een blauwe, gestippelde lijn. De geplote coëfficiënten geven – net zoals in paragraaf 5.1 – het geschatte effect weer van het opleggen van een sanctie respectievelijk waarschuwing ten opzichte van de beslissing *niet-verwijtbaar*. Naast de coëfficiënten tonen de figuren ook 95% betrouwbaarheidsintervallen (grijze en blauwe vlakken; zie box 5.1).

### Box 5.1 Interpretatie betrouwbaarheidsintervallen

Het 95% betrouwbaarheidsinterval geeft de onzekerheid rond een effectschatting weer en kan als volgt worden geïnterpreteerd: als we het onderzoek 100 keer herhalen, dan verwachten we dat de parameter die de effectschatting weergeeft minstens 95 keer binnen het betrouwbaarheidsinterval ligt.

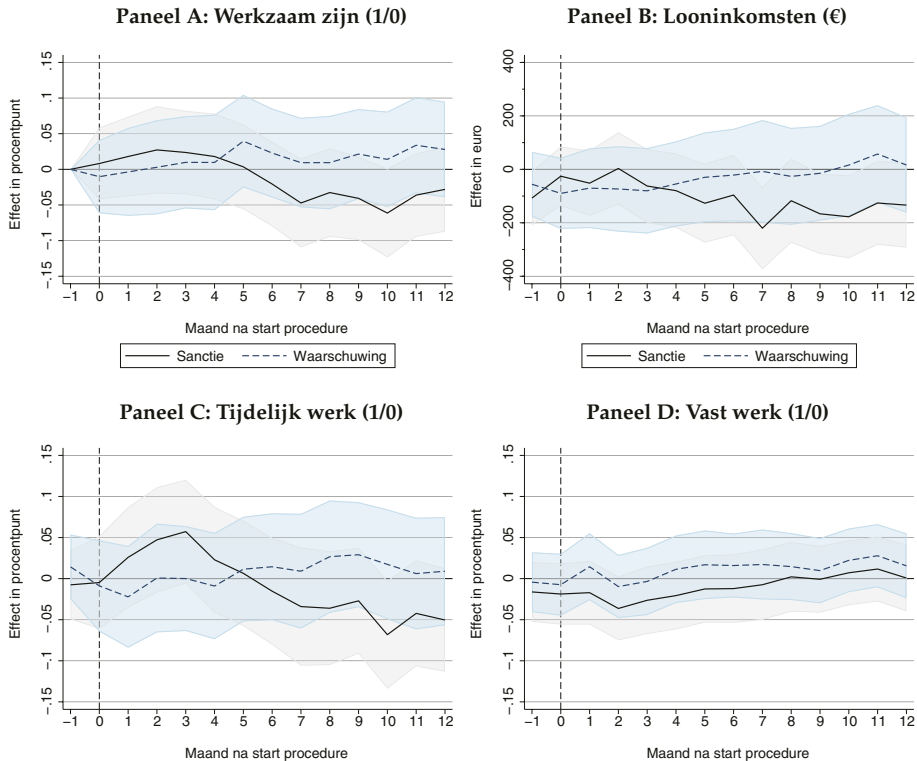
### Effecten op arbeidsparticipatie over de tijd

Figuur 5.1 laat zien wat de maandelijkse effecten zijn op de kans om werkzaam te zijn (paneel A), op looninkomsten (paneel B) en op de kans om tijdelijk (paneel C) of vast werk (paneel D) te hebben.

Paneel A suggereert dat gesanctioneerde WW-gerechtigden in de tweede helft minder kans hadden op werk. De effectschattingen waren voor deze periode negatief, maar niet significant (dat wil zeggen: de waarde nul ligt steeds in het betrouwbaarheidsinterval). Desalniettemin moet worden opgemerkt dat een groot deel van het betrouwbaarheidsinterval voor de laatste zes maanden in het negatieve domein lag. Daardoor is het aannemelijk dat het opleggen van een sanctie in ieder geval geen grote *positieve* effecten op werkkansen heeft. Voor een waarschuwing liggen de effectschattingen steeds dicht bij nul. We vinden dus nog steeds geen aanwijzingen dat waarschuwingen de kans op werk vergroten of verkleinen.

Mensen die werkzaam zijn, kunnen werken onder een contract voor bepaalde of onbepaalde tijd. De contractsoort kan als kwaliteitsindicator dienen voor het werk dat men doet. Immers, een contract voor onbepaalde tijd geeft stabiliteit; onder dit soort contract is de kans om (opnieuw) werkloos te raken kleiner. In de volgende stap splitsten we de effecten op werkkansen uit paneel A op in effecten naar contractsoort. Dit deden we met het doel om iets te kunnen zeggen over mogelijke effecten op de kwaliteit van werkhervatting. Paneel C en D tonen de effecten op de kans om tijdelijk respectievelijk voor onbepaalde tijd te werken. De effecten in paneel C en D tellen op tot het effect op werkkansen in paneel A.

Figuur 5.1 Effecten op arbeidsparticipatie over de tijd



Paneel C suggereert dat sancties ervoor zorgden dat mensen aanvankelijk vaker tijdelijk werk hadden, en na vijf à zes maanden juist minder vaak (de zwarte lijn gaat eerst omhoog, dan omlaag). Dit patroon moet echter zeer voorzichtig geïnterpreteerd worden, aangezien de waarde nul bijna altijd in het betrouwbaarheidsinterval lag (behalve in maand 10,  $p = 0,042$ ). Een dergelijk tijdsverloop was niet te zien voor waarschuwingen (de blauwe stippellijn lijkt vlakker). Paneel D suggereert dat de effecten van sancties en waarschuwingen op de kans op vast werk allebei rond de nul liggen (beide lijnen lijken vlak); er is geen opvallend tijdsverloop.

Ten slotte kijken we naar de maandelijkse effecten op looninkomsten. Paneel B laat zien dat er negatieve effecten van een sanctie lijken op te treden in de zesde tot twaalfde maand na de overtreding (de zwarte lijn zakt daar onder de nullijn). Dit patroon past bij de eerder beschreven negatieve effecten op werkansen in dezelfde periode (zie paneel A). Het negatieve effect van sancties op looninkomsten was het grootst in maand 7. In deze maand verdienden gesanctioneerde mensen € 220 minder dan het steekproefgemiddelde ( $p = 0,005$ ), wat neerkomt op een negatief effect van 18%. Over de hele tweede helft genomen bedroeg het effect gemiddeld -€ 155 per maand ( $p = 0,028$ ), wat 12% minder is

dan het steekproefgemiddelde. De effecten van waarschuwingen hadden geen opmerkelijk tijdsverloop (blauwe stippellijn is vlak en ligt rond de nullijn). We trekken de volgende voorzichtige, tussentijdse conclusie: Er ontstaat een genuanceerd beeld wat betreft het tijdsverloop van de effecten van sancties. Mogelijk leidde een sanctie tot lagere kansen op werk en (mede daardoor) tot lagere looninkomsten, vanaf ongeveer zes maanden na de overtreding. Voor looninkomsten was dit negatieve effect met enige betrouwbaarheid vast te stellen. Voor de kans op werk waren de effectschattingen echter met veel onzekerheid omgeven. Een opsplitsing van effecten naar contractsoort suggereerde dat, als er al effecten waren, deze vooral gedreven werden door tijdelijk werk.

#### **Effecten op uitkeringsafhankelijkheid over de tijd**

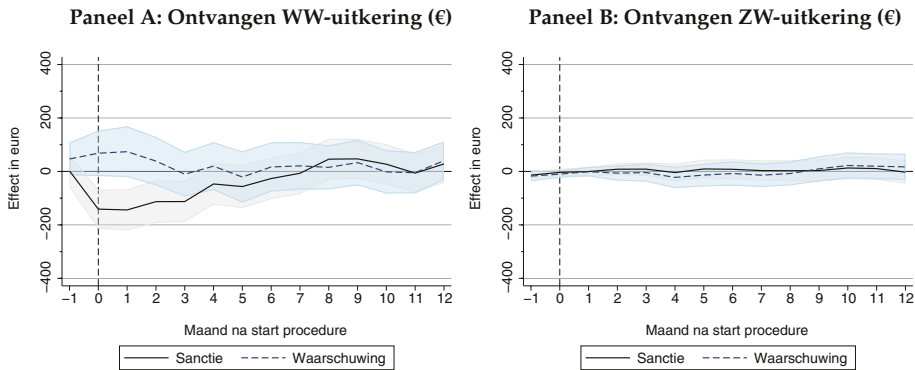
Figuur 5.2 laat zien wat de maandelijkse effecten zijn op de ontvangen WW-uitkering (paneel A) en ZW-uitkering (paneel B), de twee variabelen die we gebruiken om uitkeringsafhankelijkheid vast te stellen.

Sancties nemen vaak de vorm aan van een tijdelijke korting van de WW-uitkering. Deze korting wordt in paneel A duidelijk zichtbaar. In de maand van de verantwoordingstaak (maand 0) en de drie daaropvolgende maanden was het effect van een sanctie op de ontvangen WW-uitkering negatief en statistisch significant. De grootte van het effect was ongeveer –€ 140 in maand 0 ( $p < 0,001$ ) en nam in de maanden daarna af. Dit komt neer op een vermindering van de uitkering van ongeveer 15% ten opzichte van het steekproefgemiddelde.

Deze bevinding is consistent met het sanctiebeleid van UWV, dat voorschrijft dat een sanctie een korting op de uitkering is van maximaal 25%, voor vier maanden. Dat de korting in ons geval gemiddeld ongeveer 15% was, hangt mogelijk samen met het feit dat de genoemde 25% een maximum is; er zijn waarschijnlijk WW-gerechtigden die een mildere sanctie hebben gekregen. Het resultaat suggereert ook dat de korting op de uitkering onmiddellijk wordt doorgevoerd. Het effect trad immers al in de maand van de verantwoordingstaak op.

Na drie maanden lijkt het effect compleet verdwenen te zijn, wat erop wijst dat het opleggen van een sanctie verder geen invloed had op het gebruik van het recht op een WW-uitkering. Ook voor waarschuwingen waren er geen noemenswaardige effecten. Verder hebben zowel sancties als waarschuwingen geen effect op de maandelijks ontvangen ZW-uitkering. Het blijkt dus dat de twee maatregelen – ook wanneer we het tijdsverloop onderzoeken – geen effect hebben op uitkeringsafhankelijkheid.

**Figuur 5.2** Effecten op uitkeringsafhankelijkheid over de tijd



### 5.3 EFFECTEN PER OPLEIDINGSNIVEAU

Tot nu toe hebben we gekeken naar gemiddelde effecten voor onze hele steekproef. Het is echter goed mogelijk dat sancties en waarschuwingen verschillende effecten hebben op verschillende groepen WW-gerechtigden. Dit wordt ook wel *heterogeniteit van effecten* genoemd. Meer specifiek onderzochten we of de effecten van sancties en waarschuwingen verschillend zijn voor WW-gerechtigden met verschillende opleidingsniveaus.<sup>14</sup>

Er zijn verschillende redenen waarom de effecten van sancties en waarschuwingen anders zouden kunnen zijn voor WW-gerechtigden met verschillende opleidingsniveaus.<sup>15</sup> Ten eerste kunnen we ervan uitgaan dat een financiële korting op de uitkering relatief meer impact heeft op WW-gerechtigden met een lager opleidingsniveau. Omdat de hoogte van de WW-uitkering afhangt van het laatstverdiende loon, krijgen lager opgeleiden vaak een lagere WW-uitkering. Lager opgeleiden in onze steekproef kregen in de maand voor overtreding gemiddeld € 908; hoger opgeleiden kregen gemiddeld € 1.324 (46% meer). Ten tweede hebben lager opgeleiden vaak minder spaargeld, waardoor zij een minder grote financiële buffer hebben waarop zij kunnen terugvallen tijdens periodes van werkloosheid (Browning & Lusardi, 1996). Kortom, een sanctie zal lager opgeleiden waarschijnlijk harder treffen omdat de korting op de uitkering hen financieel meer inperkt en dichter bij – of zelfs onder – het bestaansminimum

14 Het is denkbaar dat effecten ook voor andere groepen verschillen, bijvoorbeeld voor personen die al langer een WW-uitkering ontvangen en personen die net de WW zijn ingestroomd. Een opsplitsing van effecten langs alle mogelijke groepen valt echter buiten het bestek van dit rapport. We richten ons op opleidingsniveaus omdat deze indeling met de beschikbare data goed te onderzoeken is, en omdat er vanuit theoretisch perspectief veel te zeggen valt over een verschil in effecten met betrekking tot opleidingsniveau.

15 Ter herinnering: een laag opleidingsniveau was gedefinieerd als basisonderwijs, vmbo, havo of vwo; een gemiddeld opleidingsniveau als mbo 1 tot en met mbo 4; een hoog opleidingsniveau als hbo of wo.

brengt. Zodoende zijn er wellicht sterkere effecten van sancties te verwachten onder lager opgeleiden dan onder hoger opgeleiden.

Aan de andere kant is er ook een reden te bedenken waarom sancties sterker zouden kunnen werken voor hoger opgeleiden, namelijk dat er voor hoger opgeleiden een kleiner aanbod van passend werk is. Een gevolg hiervan is dat hoger opgeleiden wellicht aanvankelijk minder snel een nieuwe baan zullen accepteren dan lager opgeleiden, bijvoorbeeld omdat het loonaanbod te laag ligt of omdat de baan niet direct aansluit bij hun kwalificaties. Een ander gevolg is echter dat het voor hoger opgeleiden relatief eenvoudig is om het bestand aan banen dat zij bereid zijn te accepteren uit te breiden, wat ze – volgens economische modellen – zouden moeten doen na het krijgen van een sanctie. Als deze redenering klopt, zouden sancties juist voor hoger opgeleiden sterkere effecten moeten hebben, met name op werkhervatting en op uitkeringsafhankelijkheid. Tabel 5.2 toont geschatte effecten voor dezelfde uitkomstmaten als tabel 5.1, maar dan apart voor drie subgroepen (laag, gemiddeld, hoog opgeleid).<sup>16</sup> Om de tabel overzichtelijk te houden rapporteren we deze keer geen standaardfouten en *p*-waardes, maar duiden statistisch significante effecten alleen aan door middel van asterisken. Kolom 1 en 4 geven de geschatte effecten weer van respectievelijk sancties en waarschuwingen voor lager opgeleiden. De overige kolommen tonen de verschillen tussen de effecten voor lager en gemiddeld opgeleiden (2 en 5), en de verschillen tussen lager en hoger opgeleiden (3 en 6).<sup>17</sup> De resultaten laten zien dat de effecten van een sanctie en een waarschuwing in het algemeen niet erg verschillen voor groepen met verschillende opleidingsniveaus. Het effect van sancties op naleving zwakt af naarmate het opleidingsniveau toeneemt, maar deze afzwakking is niet significant. Hetzelfde geldt voor de effecten van waarschuwingen. Ook voor werkzoekgedrag liggen de effecten voor de verschillende subgroepen redelijk dicht bij elkaar, en zijn de verschillen tussen subgroepen niet significant.

Voor de effecten op arbeidsparticipatie en uitkeringsafhankelijkheid lijken er iets grotere verschillen te zitten tussen de subgroepen. Zo treden er grotere negatieve effecten op voor gesanctioneerde hoger opgeleiden vergeleken met gesanctioneerde lager opgeleiden. Maar ook hier zijn de effecten veelal niet statistisch significant. Het is belangrijk om te beseffen dat subgroepen minder observaties bevatten dan de hele steekproef; de steekproef bestaat bijvoorbeeld voor slechts 14% uit hoger opgeleiden. Hierdoor zijn de foutmarges groot, waardoor het bewijs – als het er al is – niet erg overtuigend is.

---

16 Deze keer laten we voor naleving alleen de effecten zien op de kans een nieuwe verantwoordingsstaak te krijgen.

17 Met de informatie in de tabel kan ook het effect voor gemiddeld en hoger opgeleiden worden berekend. Het effect van een sanctie op naleving voor hoger opgeleiden, bijvoorbeeld, is de som van de coëfficiënten in kolom 1 en 3 ( $-0,129 + 0,060 = -0,069$ ). Het opleggen van een sanctie verkleint de kans op een nieuwe verantwoordingsstaak dus met 13 procentpunt voor lager opgeleiden en met 7 procentpunt voor hoger opgeleiden.

We trekken de volgende tussentijdse conclusie: In tegenstelling tot onze verwachtingen vinden we geen bewijs dat effecten van sancties (en waarschuwingen) verschillend zijn voor groepen met verschillende opleidingsniveaus. Hoewel er enige aanwijzingen waren dat effecten anders uitpakten voor hoger opgeleiden (met name: er leken grotere negatieve effecten van sancties op arbeidsparticipatie te zijn voor hoger opgeleiden), zijn deze effecten niet overtuigend genoeg om er conclusies aan te verbinden.

**Tabel 5.2 Effecten op cumulatieve uitkomsten per opleidingsniveau**

Uitkomstmaten	Effect sanctie			Effect waarschuwing			N
	Laag	Gemiddeld vs. laag	Hoog vs. laag	Laag	Gemiddeld vs. laag	Hoog vs. laag	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
<b>Paneel A: Naleving</b>							
Nieuwe verantwoordingsstaak (1/0)	-0,129***	0,035	0,060	0,016	-0,007	-0,065	51.848
<b>Paneel B: Werkzoekgedrag</b>							
Gem. aantal activiteiten/maand	0,193	-0,097	-0,203	1,004***	-0,097	-0,147	46.751
<b>Paneel C: Arbeidsparticipatie</b>							
Ooit werkzaam (1/0)	0,016	-0,045	-0,079*	0,049	-0,036	-0,003	51.848
Looninkomen (€)	-318	-1.102	-1.982	-690	-103	2.535	51.848
Uren gewerkt	20	-81	-130	-15	4	60	51.848
<b>Paneel D: Uitkeringsafhankelijkheid</b>							
Ontvangen WW-uitkering (€)	-283	-98	-498	222	-21	-126	51.848
Ontvangen ZW-uitkering (€)	-15	20	3	15	27	-66	51.848

*Noot:* Elke rij toont de resultaten van een *two-stage least squares* (2SLS) regressie. Uitkomstmaten staan links. Kolom 1-3 geven de effecten van een sanctie per opleidingsniveau weer, kolom 4-6 de effecten van een waarschuwing per opleidingsniveau. We schatten effecten in een model met interactietermen. In het 2SLS-model dient strengheid als instrumentele variabele voor de daadwerkelijke behandeling. Strengheid wordt gemeten als de *leave-out mean* van een adviseur. Alle regressies controleren voor kantoor, maand en achtergrondkenmerken. Standaardfouten worden geclusterd op niveau van de adviseurs ( $N = 1.303$ ).

\*  $p < 0,05$ .

\*\*  $p < 0,01$ .

\*\*\*  $p < 0,001$ .



## 6.1 SAMENVATTING VAN DE BELANGRIJKSTE RESULTATEN

### Waren mensen geneigd zich beter aan de regels te houden?

- *Sancties* zorgden ervoor dat mensen zich in het vervolg beter aan de regels, zoals de sollicitatieplicht, hielden. Meer specifiek: na het krijgen van een sanctie kregen WW-gerechtigden 40% minder vaak nogmaals een verantwoordingstaak voor onvoldoende of niet op tijd doorgegeven sollicitaties.
- Er was geen bewijs dat *waarschuwingen* ervoor zorgden dat mensen zich in het vervolg beter aan de regels hielden.

### Gingen mensen intensiever op zoek naar nieuw werk?

- Er was geen bewijs dat *sancties* WW-gerechtigden motiveerden om intensiever op zoek te gaan naar nieuw werk.
- *Waarschuwingen* hadden een duidelijk effect op de intensiteit waarmee WW-gerechtigden op zoek gingen naar nieuw werk. Na het krijgen van een waarschuwing gaven WW-gerechtigden meer sollicitatieactiviteiten door.

### Waren mensen succesvoller in het vinden van nieuw werk?

- Er waren aanwijzingen dat *sancties* het inkomen uit loon verminderden. Nadat WW-gerechtigden een sanctie hadden gekregen, verdienden ze minder (door te werken) dan dat ze zouden hebben verdiend als ze geen sanctie gekregen zouden hebben. Nuance is hier op zijn plaats: Dit inkomstenverlies trad niet meteen op; het werd pas zichtbaar na een paar maanden. Tussen de 6de en de 12de maand na de sanctie verdienden mensen zo'n € 155 per maand minder dan dat ze zonder sanctie zouden hebben verdiend. Het zou dus kunnen dat sancties een vertraagd effect hebben op de bereidheid en/of het vermogen om te werken. Let wel: de onzekerheidsmarges rond onze schattingen van de effecten op inkomsten zijn groot.
- Er was geen bewijs dat *waarschuwingen* effect hadden op inkomen uit loon, noch op andere aspecten van arbeidsparticipatie. Dat wil zeggen: na het ontvangen van een waarschuwing waren WW-gerechtigden niet aanmerkelijk meer of minder succesvol op de arbeidsmarkt. Er was ook geen bewijs voor een vertraagd effect van waarschuwingen, zoals bij sancties wel enigszins het geval was.

### Waren mensen minder afhankelijk van een uitkering van UWV?

- Er was geen bewijs dat *sancties* of *waarschuwingen* invloed hadden op het toekomstige gebruik van WW- of ZW-uitkeringen.

Verder onderzochten we of de effecten van waarschuwingen en sancties verschillend waren voor mensen met verschillende opleidingsniveaus. We vonden geen bewijs voor systematische verschillen.

### We concluderen het volgende over het gebruik van sancties en waarschuwingen in de context van de werkloosheidsuitkering:

- In zijn geheel beschouwd valt de effectiviteit van sancties tegen. Het enige positieve effect dat we vonden is dat sancties de kans verminderen dat mensen nogmaals niet voldoen aan hun verplichtingen. Echter, er was geen bewijs dat deze vermindering veroorzaakt werd doordat mensen daadwerkelijk meer gingen solliciteren. Bovendien lijken sancties de inkomsten uit werk te verminderen.
- Op basis van economische modellen (en wellicht op basis van intuïtie) zou men kunnen verwachten dat het krijgen van een sanctie mensen motiveert om weer aan het werk te gaan. Het huidige onderzoek bevestigt die verwachting niet. Integendeel: de bevindingen passen beter bij de interpretatie dat sancties een *vermindering* van het werkvermogen of de werkbereidheid teweegbrengen.
- Een belangrijke kanttekening bij de twee bovenstaande conclusies: het huidige onderzoek sluit niet uit dat sancties een *dreigende werking* hebben (zie paragraaf 1.3). Het is bijvoorbeeld aannemelijk dat veel mensen zich goed aan de regels houden omdat ze bang zijn dat ze anders een sanctie krijgen (Arni et al., 2013; Lalive et al., 2005). De twee conclusies hierboven (de effectiviteit van sancties valt tegen; sancties leiden wellicht tot een vermindering van werkvermogen of werkbereidheid) gaan alleen over de effecten van het daadwerkelijk *opgelegd krijgen* van een sanctie.
- In tegenstelling tot sancties stimuleren waarschuwingen mensen wél om intensiever te zoeken naar werk. We vonden verder geen effecten die erop zouden kunnen wijzen dat waarschuwingen een demotiverende rol spelen. Vanuit het perspectief van UWV lijkt het gebruik van waarschuwingen dus veiliger te zijn dan het gebruik van sancties.

## 6.2 INHOUDELIJKE INTERPRETATIE VAN DE RESULTATEN

De resultaten roepen twee belangrijke vragen op, die we eerst zullen bespreken. Dan zullen we de resultaten interpreteren aan de hand van eerder onderzoek uit de economie en psychologie.

**Sancties zorgden voor meer naleving van de sollicitatieplicht. Toch gingen mensen door een sanctie niet aanmerkelijk meer solliciteren. Hoe komt dat?**

Dit patroon is verrassend, maar goed verklaarbaar als we in acht nemen dat de sollicitatieplicht niet alleen gaat over de hoeveelheid sollicitatieactiviteiten, maar ook over het tijdig doorgeven ervan. Het zou kunnen dat het krijgen van een sanctie – die uiteraard onprettig is – ervoor zorgt dat mensen op een zeer doelgerichte manier proberen te voorkomen dat ze nog een keer niet voldoen aan de sollicitatieplicht. Met andere woorden: het kan zijn dat de sanctie mensen stimuleert om precies te doen wat vereist is (vier sollicitatieactiviteiten binnen vier weken; deze activiteiten op tijd doorgeven), maar niets meer. Zodoende kan het dat mensen vaker voldoen aan de sollicitatieplicht zonder noemenswaardig meer sollicitatieactiviteiten te ondernemen.

Een soortgelijke redenering is ook bruikbaar bij het interpreteren van de effecten van waarschuwingen. Waarschuwingen zorgden voor een toename van sollicitatieactiviteiten; echter, deze toename ging niet gepaard met een betere naleving van de sollicitatieplicht. Het zou kunnen dat waarschuwingen mensen weliswaar stimuleerden om intensiever te zoeken naar nieuw werk, maar niet om zich beter aan de regels te houden. Hierdoor kan het zijn dat gewaarschuwde mensen alsnog soms – of althans even vaak als niet-gewaarschuwde mensen – tekortschoten (bijvoorbeeld door sollicitatieactiviteiten te laat door te geven of door incidenteel, bij wijze van negatieve uitschieter, onvoldoende te solliciteren).

**Waarschuwingen stimuleerden mensen om intensiever te zoeken naar nieuw werk. Toch leidden waarschuwingen niet tot een snellere uitstroom uit de WW. Hoe komt dat?**

Intuïtief zou je kunnen verwachten dat als mensen meer sollicitatieactiviteiten ontplooiën, ze ook eerder uit de WW zouden moeten uitstromen. Immers, meer sollicitatieactiviteiten zouden moeten leiden tot een grotere kans op nieuw werk. Onze bevinding is zodoende enigszins verrassend, maar kan ook goed verklaard worden als we beseffen dat er geen een-op-een relatie is tussen het aantal sollicitatieactiviteiten en de kans op nieuw werk. Of sollicitatieactiviteiten succesvol zijn, hangt namelijk af van veel verschillende factoren. Denk bijvoorbeeld aan de zorgvuldigheid en de kwaliteit van de sollicitatie, de match met de beoogde functie, en de hoeveelheid concurrentie. Bovendien is het voor sommige groepen mensen moeilijker om werk te vinden door discriminatie op de arbeidsmarkt (Ramos, Thijssen, & Coenders, 2021). Kortom, er zijn veel invloeden op uitstroom uit de WW die onafhankelijk zijn van het aantal uitgevoerde sollicitatieactiviteiten (Van Hoof, Wanberg, & Van Hoya, 2013).

Bovendien is het goed om te memoreren dat WW-gerechtigden – om aan de sollicitatieplicht te voldoen – ook activiteiten kunnen ontplooiën die niet direct hoeven te leiden tot werk. Denk hierbij aan het volgen van sollicitatieworkshops en het voeren van netwerkgesprekken. Dit soort activiteiten heeft niet altijd een meetbaar resultaat. Om al deze redenen moeten we dus niet verbaasd zijn dat waarschuwingen weliswaar de hoeveelheid sollicitatieactiviteiten omhoog

brachten, maar dat deze toename zich niet vertaalde in een zichtbaar hogere uitstroom uit de WW.

### **Interpretatie aan de hand van economische modellen**

Er zijn eerdere economische modellen die voorspellingen maken over de effecten van sancties (zie hoofdstuk 1). In het kort voorspellen deze modellen dat sancties ervoor zorgen dat mensen intensiever op zoek zullen gaan naar nieuw werk. Ook voorspellen deze modellen dat sancties het *reserveloan* – het minimale loon waarvoor mensen bereid zijn te werken – verlagen. Door een verlaging van het reserveloan vergroten sancties de kans op werk en op uitstroom uit de WW, maar ook de kans dat mensen banen met minder goede arbeidsvoorwaarden accepteren.

Deze voorspellingen worden grotendeels niet ondersteund door onze bevindingen. Zo vonden we geen bewijs dat sancties mensen motiveerden om intensiever te zoeken naar nieuw werk. We vonden ook geen bewijs dat sancties ervoor zorgden dat mensen sneller uitstroomden uit de WW.

Er was desondanks ook een bevinding die goed lijkt te passen bij economische modellen, namelijk de bevinding dat mensen die een sanctie hadden gekregen, in de 6de tot de 12de maand na de sanctie minder inkomsten uit loon hadden dan dat ze zonder sanctie zouden hebben gehad. Een verklaring hiervoor, gebaseerd op economische modellen, zou als volgt kunnen zijn: Na een sanctie willen mensen hun verlies van WW-inkomen snel compenseren. Daardoor accepteren ze eerder werk met slechtere arbeidsvoorwaarden, zoals werk met een lager loon of met een kortdurend contract. Als mensen vervolgens werk met slechtere arbeidsvoorwaarden hebben, is de kans wellicht groter dat zij ook weer sneller stoppen – of sneller moeten stoppen – met het werk. Hoewel deze redenering speculatief is, zou ze kunnen verklaren waarom mensen na een sanctie niet direct, maar juist na een aantal maanden in inkomen achterblijven op mensen zonder sanctie.

### **Interpretatie aan de hand van psychologische modellen**

Enerzijds wijzen psychologische modellen op de motiverende (dreigende) werking van sanctiestructuren, anderzijds wijzen psychologische modellen op de nadelige bijeffecten van sancties, zoals *countercontrol*, schade aan de vertrouwensrelatie tussen de gever en ontvanger van de sanctie, en verschillende soorten berekenend gedrag (zie hoofdstuk 1; Delprato, 2002; Gneezy & Rustichini, 2000; Mulder et al., 2006). De modellen die voorspellingen maken over de nadelige bijeffecten van sancties zijn over het algemeen consistent met onze bevindingen.

We zagen dat sancties leidden tot een betere naleving van regels, terwijl er geen bewijs was voor een effect op werkzoekgedrag. Zoals hierboven beschreven, kan dit erop duiden dat WW-gerechtigden met name hebben geprobeerd het criterium – vier sollicitatieactiviteiten per vier weken – te behalen. Dit konden ze wellicht doen door extra aandacht te besteden aan het op tijd doorgeven van sollicitatieactiviteiten, zonder aanmerkelijk meer te solliciteren. Dit gedrag kan goed beschreven worden als *creaming* (paragraaf 1.2; Muller, 2018). Het lijkt erop

dat sancties mensen stimuleren het operationele criterium te behalen, maar met minimale extra inspanning.

Verder zagen we – zij het met een grote onzekerheidsmarge – dat mensen die een sanctie hadden gekregen, in de 6de tot de 12de maand na de sanctie minder inkomsten uit loon hadden dan dat ze zonder sanctie zouden hebben gehad. Hoewel deze bevinding goed past bij economische modellen, zou het ook kunnen dat psychologische mechanismen aan dit effect hebben bijgedragen. Psychologische modellen suggereren immers dat sancties leiden tot sterke emotionele reacties en tot schade aan het vertrouwen in de gever van de sanctie (in dit geval de overheid; zie paragraaf 1.2). Het zou kunnen dat de combinatie van negatieve gebeurtenissen (eerst verlies van werk, vervolgens wellicht moeite met het vinden van werk, daarna een sanctie) mensen ontmoedigt om verder te streven naar werkhervatting. Misschien komen mensen na een sanctie daarom na verloop van tijd vaker in de bijstand, of kiezen ze, als ze die mogelijkheid hebben, voor een grotere rol als huisman of huisvrouw.

Als de bovenstaande speculatie klopt, dan is het belangrijk om op te merken dat deze effecten niet van de ene op de andere dag plaatsvinden. De beslissing om te stoppen met werken zou een gevolg kunnen zijn van langdurig wikken en wegen, wat zou kunnen verklaren waarom het effect pas na zes tot twaalf maanden zichtbaar wordt.

### **Conclusie**

Verschillende effecten die we hebben gevonden, passen goed bij de eerdere economische en psychologische literatuur. Dat gezegd hebbende: er is niet één model waarvan we kunnen zeggen dat dat model echt het beste past bij de bevindingen. Bovendien: er is niet één model dat de effecten van sancties en waarschuwingen van tevoren goed kon voorspellen. Deze kanttekeningen wijzen ook op een belangrijk agendapunt voor de (toegepaste) economie en (toegepaste) psychologie: er is meer onderzoek nodig om de psychologische en economische effecten van sancties en waarschuwingen beter te begrijpen (zie tabel 6.1).

Hoewel eerdere modellen doorgaans geen onderscheid maken tussen de effecten van waarschuwingen en sancties, is het gebruikelijk om waarschuwingen te interpreteren alsof het milde sancties zijn (Mulder et al., 2014). Onze bevindingen passen bij deze interpretatie. Waarschuwingen hadden geen noemenswaardig effect op naleving, maar daar stond tegenover dat de kans op nadelige bijeffecten ook kleiner leek te zijn. Dat wil zeggen: er was geen bewijs dat waarschuwingen leidden tot inkomensverlies (waar dat er wel was voor sancties), maar waarschuwingen stimuleerden mensen wel om meer sollicitatieactiviteiten te ontplooien (waar sancties dat niet deden).

Tabel 6.1 Overzicht van vervolgvragen

Onderwerp	Aanleiding
Effecten van waarschuwingen	Onze bevindingen suggereren dat waarschuwingen veelbelovende instrumenten zijn. Het is waarschijnlijk cruciaal hoe waarschuwingen geformuleerd worden. Welke boodschap werkt het beste en waarom?
Effecten van de hoogte en duur van sancties	Uit onze data bleek dat de meeste WW-gerechtigden een standaard-sanctie opgelegd kregen (25% minder uitkering, voor vier maanden). Welke hoogte en duur zijn optimaal?
Effecten op gezondheid, welbevinden, schulden	Onze bevindingen suggereren dat sancties een dubbel negatief effect hebben op besteedbaar inkomen (eerst komt de sanctie zelf, dan het inkomensverlies). Het is belangrijk om te kijken naar de bredere gevolgen van dit inkomstenverlies. Hebben sancties effect op fysieke en mentale gezondheid? Op de opbouw van schulden?
Effecten op vertrouwen	Onze bevindingen zijn consistent met het idee dat sancties een negatief effect hebben op vertrouwen in overheidsinstanties. Echter, omdat we vertrouwen in de overheid niet direct geobserveerd hebben, is dit slechts speculatie. Speelt vertrouwen inderdaad een rol?
Effecten op verschillende soorten werkzoekgedrag	In het huidige onderzoek keken we naar werkzoekgedrag door te kijken naar het aantal geregistreerde sollicitatieactiviteiten. Een meer specifieke analyse is waarschijnlijk nuttig. Stimuleren sancties en waarschuwingen misschien bepaalde typen werkzoekgedrag? Welk type sollicitatieactiviteiten is het meest kansrijk?

### 6.3 REFLECTIE OP DE METHODE

We gebruikten een *instrumentele-variabele-analyse* (IVA) om, ondanks dat we gebruikmaakten van bestaande, observationele data, toch uitspraken te kunnen doen over causaliteit. We bespraken deze techniek in detail in hoofdstuk 3. Wat nu volgt is onze reflectie op het gebruik van de techniek. We richten ons op de vraag: Is deze analysetechniek nuttig voor analisten bij overheidsinstanties om de causale effecten van beleid te onderzoeken?

Onze ervaringen met het gebruik van IVA zijn over het algemeen positief. Immers, de causale effecten van sancties en waarschuwingen zijn potentieel interessant voor verschillende partijen; echter, deze causale effecten zijn – in ieder geval in deze context – heel moeilijk met een experimentele benadering te onderzoeken. Het zou onethisch zijn om mensen (die niet hebben voldaan aan hun sollicitatieplicht) willekeurig toe te wijzen aan experimentele groepen, waarbij de eerste groep bijvoorbeeld een sanctie en de tweede groep een waarschuwing krijgt. IVA is zodoende in principe een nuttige toevoeging op bestaande experimentele en quasi-experimentele technieken, zoals *regression discontinuity* of *difference-in-differences*-designs.

We bespreken drie aandachtspunten voor analisten die overwegen deze techniek te gebruiken om (bestaande) registratiedata te analyseren:

- Om IVA te gebruiken moet er sprake zijn van een willekeurig proces dat de kans op een bepaalde behandeling vergroot. Dit soort processen is niet altijd makkelijk te vinden. Om deze processen op te sporen is het nuttig om de ogen open te houden voor de volgende twee omstandigheden: (a) Er moet *discretionaire ruimte* zijn voor iemand die verantwoordelijk is voor een bepaalde beslissing. Het kan hier bijvoorbeeld gaan om een adviseur, een handhaver, een zorgmedewerker of een rechter. De discretionaire ruimte kan liggen in hoe streng deze beslisser is in het toekennen van sancties of andere straffen, maar zij kan ook een totaal andere vorm aannemen (bijvoorbeeld: hoe terughoudend een arts is met het voorschrijven van antibiotica). (b) Bovendien moet er sprake zijn van *willekeurige toedeling* van casussen (de mensen waarover de beslissing wordt genomen) aan degenen die de beslissing moeten nemen. Er is bijvoorbeeld vaak sprake van (gedeeltelijk) willekeurige toedeling als er een geautomatiseerd systeem gebruikt wordt om nieuwe casussen te koppelen aan beslissers.
- Om IVA te kunnen uitvoeren is het belangrijk om te beschikken over de juiste data. Zo is het cruciaal om te weten wie welke beslissing heeft genomen over welke casus. Bovendien moet er data beschikbaar zijn over de uitkomsten van de beslissing – dat wil zeggen: het moet mogelijk zijn om te analyseren hoe het de casus is vergaan in de periode na de beslissing. Verder is het belangrijk om de aannames van IVA te kunnen toetsen. Hiervoor zijn bijvoorbeeld data over kenmerken van casussen nodig, maar ook data over kenmerken van de beslissers. Het is zodoende vaak nodig om verschillende databronnen te koppelen. Ook is het nodig om de datakwaliteit voortdurend te controleren. Kortom, bij een project dat gebruikmaakt van IVA moet er veel aandacht zijn voor datamanagement en voorbereiding.
- Omdat de willekeurige toedeling (bijvoorbeeld: van mensen aan adviseurs) cruciaal is bij IVA, is het belangrijk om goed zicht te krijgen op hoe dit toedelingsproces werkt. Vooral bij grote organisaties en complexe procedures zou het kunnen dat er soms wordt afgeweken van wat er op papier staat. Het kan ook zijn dat er verschillen in uitvoering bestaan tussen kantoren, teams of afdelingen, of dat er verschillen in uitvoering bestaan tussen verschillende tijdsperiodes. Het is raadzaam om zowel gebruik te maken van beleidsdocumenten als het gesprek te zoeken met mensen die dicht bij het toedelingsproces staan zoals het wordt uitgevoerd in de praktijk.

#### 6.4

#### AANBEVELINGEN VOOR BELEIDSMAKERS

Op basis van onze bevindingen kunnen we uitspraken doen over beleid van overheidsinstanties die, onder voorwaarden, werkloosheidsuitkeringen verstrekken (in Nederland: UWV en alle gemeenten). Eerst doen we aanbevelingen over het gebruik van sancties en waarschuwingen binnen de huidige sanctiestructuur van UWV, waarin mensen vier keer per vier weken moeten

solliciteren. Vervolgens doen we meer algemene aanbevelingen over het gebruik van sanctiestructuren in de context van uitkeringen. Met grote voorzichtigheid kunnen onze aanbevelingen vertaald worden naar andere situaties waarin een overheidsinstantie onder voorwaarden uitkeringen of toeslagen verstrekt aan individuele burgers.

### **Sancties en waarschuwingen binnen de huidige sanctiestructuur**

Binnen een structuur waarin mensen vier keer per vier weken moeten solliciteren, zijn er verschillende beleidsopties denkbaar. Ten eerste kan een overheidsinstantie het beleid voeren dat er altijd gesanctioneerd wordt, ook wanneer iemand voor het eerst zijn of haar verplichtingen niet nakomt. Ten tweede kan een overheidsinstantie het beleid voeren waarin mensen bij de eerste overtreding een waarschuwing krijgen, en bij elke daaropvolgende overtreding een sanctie. Voor instanties die een van deze beleidsopties overwegen, adviseren we het volgende:

- Onze bevindingen laten zien dat sancties, in zijn geheel beschouwd, weinig effectief zijn. Overheidsinstanties die desondanks sancties gebruiken, doen er goed aan om te reflecteren op wat de reden voor het gebruik van sancties eigenlijk is. Op populatieniveau zijn er redenen te bedenken (bijvoorbeeld: misschien beoogt de instantie te laten zien dat zij bereid is te sanctioneren om de afschrikwekkende werking voor anderen te versterken), maar het is goed om te realiseren dat men niet veel kan verwachten van de effecten van sancties op individueel niveau (*ex post effecten*, zie paragraaf 1.3).
- Er zijn diverse ongewenste bijeffecten van sancties uit de eerdere literatuur bekend. Onze bevindingen suggereren dat deze bijeffecten realistisch zijn en te verwachten vallen in de context van sancties voor burgers. Instanties die het gebruik van sancties overwegen, moeten rekening houden met verschijnselen zoals *creaming* (zie paragraaf 1.2).
- Onze bevindingen suggereren dat WW-gerechtigden die een sanctie hadden gekregen, feitelijk twee keer te maken kregen met een verlies van inkomsten. Eerst werd er gesneden in hun WW-uitkering, vervolgens hadden ze een lager inkomen uit werk. Sancties zouden dus – op de langere termijn – mensen financieel harder kunnen raken dan dat men intuïtief zou verwachten. Instanties die het gebruik van sancties overwegen, doen er goed aan om hierop te reflecteren.

Ten derde kan een overheidsinstantie het beleid voeren om zeer terughoudend om te springen met sancties en vooral waarschuwingen te gebruiken. Ons advies ten aanzien van deze beleidsoptie is als volgt:

- Onze resultaten suggereren dat deze beleidsoptie bruikbaar kan zijn als het doel van het beleid is om de inspanning van burgers te bevorderen. Een voordeel van deze beleidsoptie – ten opzichte van de voorgaande opties – is dat waarschuwingen waarschijnlijk minder negatieve bijeffecten hebben dan sancties. In dit onderzoek vonden we immers dat waarschuwingen, in tegenstelling tot sancties, mensen stimuleerden om in het vervolg meer sollicitatieactiviteiten door te geven. Ook waren er geen aanwijzingen voor



negatieve effecten op inkomen uit loon, waar die aanwijzingen er voor sancties wel waren. Dit patroon is consistent met eerder onderzoek, dat laat zien dat waarschuwingen als minder onrechtvaardig worden gezien dan sancties (Mulder et al., 2014).

### **Sancties en waarschuwingen in het algemeen**

De bovenstaande discussie nodigt wellicht uit tot de conclusie dat overheidsinstanties – in dit geval UWV en gemeenten – misschien wel helemaal moeten stoppen met het gebruik van sancties. We wijzen erop dat deze conclusie niet getrokken kan worden uit het huidige onderzoek. Immers, het huidige onderzoek leent zich er niet voor om conclusies te trekken over de *dreigende werking* die uitgaat van sancties (*ex ante effecten* of *algemene afschrikking*; zie paragraaf 1.3). Zo kan het goed zijn dat de afschrikwekkende werking van sancties grote groepen mensen motiveert om te voldoen aan hun verplichtingen (Arni et al., 2013; Lalive et al., 2005). Het zou zodoende te ver gaan om uit het huidige onderzoek te concluderen dat sancties helemaal geen rol zouden moeten spelen in het beleid van UWV en gemeenten.

Een algemeen advies dat volgt, is dat de *criteria* die gebruikt worden om overtredingen vast te stellen herzien moeten worden (Kerr, 1975; Muller, 2018). De huidige sanctiestructuur stuurt immers alleen op het aantal sollicitaties. Echter, eerder onderzoek liet zien dat de correlatie tussen het aantal sollicitatieactiviteiten en de kans op werkhervatting niet zo groot is (Van Hooft et al., 2013). Het huidige onderzoek ondersteunt deze eerdere bevinding: we vonden geen bewijs dat een toename in sollicitatieactiviteiten (in ons geval: door het krijgen van een waarschuwing) gepaard gaat met een grotere kans op werkhervatting. Het is zodoende nuttig om na te denken over andere, betere operationele criteria (zie hoofdstuk 1).

Een goed startpunt zou kunnen zijn om in kaart te brengen welke gedragingen een positief effect hebben op duurzame werkhervatting. Deze gedragingen zouden vervolgens gestimuleerd kunnen worden, bijvoorbeeld met een sanctie- of beloningsstructuur. Uit eerder onderzoek is bekend dat de zorgvuldigheid en de kwaliteit van sollicitatieactiviteiten belangrijke voorspellers zijn van duurzame werkhervatting. Ook laat eerder onderzoek zien dat *zelfregulatie* onder werkzoekenden een belangrijke voorspeller is: mensen die kunnen omgaan met tegenslagen, die heldere doelen hebben, en die op hun eigen gedrag en gedachten kunnen reflecteren, hebben een grotere kans op het vinden van een passende baan (zie voor een overzicht Van Hooft, Kammeyer-Mueller, Wanberg, Kanfer, & Basburg, 2021). Samen met het huidige rapport suggereert deze eerdere onderzoekslijn twee adviezen die het verkennen waard zijn: (a) het is wellicht beter om eisen te stellen aan de zorgvuldigheid en kwaliteit van sollicitaties in plaats van aan het aantal sollicitaties, en (b) het is opportuun om prioriteit te geven aan trajecten die zich richten op het verbeteren van zelfregulatievaardigheden (zoals UWV al gedeeltelijk doet via zelftesten, webinars en coachingsgesprekken).

## 6.5 CONCLUSIE

Veel instituties maken gebruik maken van sancties om menselijk gedrag te sturen. Het huidige rapport beschrijft hoe sancties en waarschuwingen verschillende effecten kunnen hebben. Deze effecten zijn soms ongewenst en ongepland. Voor uitvoerende overheidsinstanties is het daarom aan te raden om regelmatig te reflecteren op hun gebruik van waarschuwingen en sancties. We hopen dat dit rapport daar bruikbare aanknopingspunten voor biedt.

## LITERATUURLIJST

- Abbring, J.H., Berg, G.J. van den, & Ours, J.C. van (2005). The Effect of Unemployment Insurance Sanctions on the Transition Rate from Unemployment to Employment. *Economic Journal*, 115(505), 602-630. <https://doi.org/10.1111/j.1468-0297.2005.01011.x>.
- Angrist, J.D., Imbens, G.W., & Rubin, D.B. (1996). Identification of Causal Effects Using Instrumental Variables. *Journal of the American Statistical Association*, 91(434), 444-455. [www.jstor.org/stable/2291629](http://www.jstor.org/stable/2291629).
- Arni, P., Lalive, R., & Ours, J.C. van (2013). How effective are unemployment benefit sanctions? Looking beyond unemployment exit. *Journal of Applied Econometrics*, 28(7), 1153-1178. <https://doi.org/10.1002/jae.2289>.
- Arni, P., & Schiprowski, A. (2019). Job search requirements, effort provision and labor market outcomes. *Journal of Public Economics*, 169, 65-88. <https://doi.org/10.1016/j.jpubeco.2018.09.004>.
- Atkinson, J.W. (1964). *An introduction to motivation*. Van Nostrand.
- Austin, J.T., & Villanova, P. (1992). The criterion problem: 1917-1992. *Journal of Applied Psychology*, 77(6), 836-874. <https://doi.org/10.1037/0021-9010.77.6.836>.
- Autor, D., Kostøl, A., Mogstad, M., & Setzler, B. (2019). Disability Benefits, Consumption Insurance, and Household Labor Supply. *American Economic Review*, 109(7), 2613-2654. <https://doi.org/10.1257/aer.20151231>.
- Ayres, I., & Braithwaite, J. (1992). *Responsive regulation: Transcending the deregulation debate*. Oxford University Press.
- Baldassarri, D., & Grossman, G. (2011). Centralized sanctioning and legitimate authority promote cooperation in humans. *Proceedings of the National Academy of Sciences of the United States of America*, 108(27), 11023-11027. <https://doi.org/10.1073/pnas.1105456108>.
- Balliet, D., Mulder, L.B., & Lange, P.A.M. van (2011). Reward, punishment, and cooperation: A meta-analysis. *Psychological Bulletin*, 137(4), 594-615. <https://doi.org/10.1037/a0023489>.
- Becker, G.S. (1968). Crime and Punishment: An Economic Approach. *Journal of Political Economy*, 76(2), 169-217.
- Berg, G.J. van den, Klaauw, B. van der, & Ours, J.C. van (2004). Punitive Sanctions and the Transition Rate from Welfare to Work. *Journal of Labor Economics*, 22(1), 211-241. <https://doi.org/10.1086/380408>.
- Berg, G.J. van den, & Vikström, J. (2014). Monitoring job offer decisions, punishments, exit to work, and job quality. *Scandinavian Journal of Economics*, 116(2), 284-334. <https://doi.org/10.1111/sjoe.12051>.
- Bhuller, M., Dahl, G.B., Løken, K.V., & Mogstad, M. (2020). Incarceration, Recidivism, and Employment. *Journal of Political Economy*, 128(4), 1269-1324. <https://doi.org/10.1086/705330>.
- Boef, A.G.C., Cessie, S. le, & Dekkers, O.M. (2013). Instrumentele-variabele-analyse. *Nederlands Tijdschrift Voor Geneeskunde*, 157:A5481. [www.ntvng.nl/artikelen/instrumentele-variabele-analyse/volledig](http://www.ntvng.nl/artikelen/instrumentele-variabele-analyse/volledig).

- Boren, J.J., & Colman, A.D. (1970). Some experiments on reinforcement principles within a psychiatric ward for delinquent soldiers. *Journal of Applied Behavior Analysis*, 3(1), 29-37. <https://doi.org/10.1901/jaba.1970.3-29>.
- Browning, M., & Lusardi, A. (1996). Household Saving: Micro Theories and Micro Facts. *Journal of Economic Literature*, 34(4), 1797-1855. [www.jstor.org/stable/2729595](http://www.jstor.org/stable/2729595).
- Byron, K., & Khazanchi, S. (2012). Rewards and creative performance: A meta-analytic test of theoretically derived hypotheses. *Psychological Bulletin*, 138(4), 809-830. <https://doi.org/10.1037/a0027652>.
- Choi, J.M., & Cho, Y.S. (2020). Impaired cognitive control during reward pursuit and punishment avoidance. *Motivation and Emotion*, 44, 832-845. <https://doi.org/10.1007/s11031-020-09837-3>.
- Deci, E.L. (1971). Effects of Externally Mediated Rewards on Intrinsic Motivation. *Journal of Personality and Social Psychology*, 18(1), 105-115.
- Deci, E.L., Koestner, R., & Ryan, R.M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, 125(6), 627-668. <https://doi.org/10.1037/0033-2909.125.6.627>.
- Delprato, D.J. (2002). Countercontrol in behavior analysis. *The Behavior Analyst*, 25(2), 191-200. <https://doi.org/10.1007/BF03392057>.
- Dijk, E. van, & Dreu, C.K.W. de (2021). Experimental Games and Social Decision Making. *Annual Review of Psychology*, 72(1), 415-438. <https://doi.org/10.1146/annurev-psych-081420-110718>.
- Edwards, M.A., & Roy, S. (2017). Academic research in the 21st century: Maintaining scientific integrity in a climate of perverse incentives and hypercompetition. *Environmental Engineering Science*, 34(1). <https://doi.org/10.1089/ees.2016.0223>.
- Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature*, 415(6868), 137-140. <https://doi.org/10.1038/415137a>.
- Fehr, E., & Rockenbach, B. (2003). Detrimental effects of sanctions on human altruism. *Nature*, 422(6928), 137-140. <https://doi.org/10.1038/nature01474>.
- Fredriksson, P., & Holmlund, B. (2006). Improving incentives in unemployment insurance: A review of recent research. *Journal of Economic Surveys*, 20(3), 357-386. <https://doi.org/10.1111/j.0950-0804.2006.00283.x>.
- Garbers, Y., & Konradt, U. (2014). The effect of financial incentives on performance: A quantitative review of individual and team-based financial incentives. *Journal of Occupational and Organizational Psychology*, 87(1), 102-137. <https://doi.org/10.1111/joop.12039>.
- Gerhart, B., & Fang, M. (2015). Pay, intrinsic motivation, extrinsic motivation, performance, and creativity in the workplace: Revisiting long-held beliefs. *Annual Review of Organizational Psychology and Organizational Behavior*, 2, 489-521. <https://doi.org/10.1146/annurev-orgpsych-032414-111418>.
- Gneezy, U., Meier, S., & Rey-Biel, P. (2011). When and why incentives (don't) work to modify behavior. *Journal of Economic Perspectives*, 25(4), 191-210. <https://doi.org/10.1257/jep.25.4.191>.
- Gneezy, U., & Rustichini, A. (2000). A fine is a price. *Journal of Legal Studies*, 29(1), 1-17.
- Hoekendijk, G., & Kommer, M. (2011). Strafdoele en tenuitvoerlegging: Perspectief op een nieuwe verenigingstheorie? *Sancties: Tijdschrift over Straffen En Maatregelen*, 4, 212-222.
- Hooft, E.A.J. van, Kammeyer-Mueller, J.D., Wanberg, C.R., Kanfer, R., & Basbug, G. (2021). Job search and employment success: A quantitative review and future research agenda. *Journal of Applied Psychology*, 106(5), 674-713. <https://doi.org/10.1037/apl0000675>.
- Hooft, E.A.J. van, Wanberg, C.R., & Hoyer, G. van (2013). Moving beyond job search quantity: Towards a conceptualization and self-regulatory framework of job search quality. *Organizational Psychology Review*, 3, 3-40. <https://doi.org/10.1177/2041386612456033>.

- Imada, H., Chan, W., Ng, Y.K., Man, L.H., Wong, M., Cheng, B., & Feldman, G. (2021). Rewarding more is better for soliciting help, yet more so for cash than for goods: Revisiting and reframing the Tale of Two Markets with replications and extensions of Heyman and Ariely (2004). *Preprint*. <https://doi.org/10.13140/RG.2.2.29661.33763/1>.
- Inspectie Verkeer en Waterstaat (2009). *Jaarverslag 2009*.
- Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. *American Psychologist*, 39(4), 341-350. <https://doi.org/10.1037/0003-066X.39.4.341>.
- Kerr, S. (1975). On the Folly of Rewarding A, While Hoping for B. *Academy of Management Journal*, 18(4), 769-783. <https://doi.org/10.5465/255378>.
- Killias, M., Villettaz, P., & Nunweiler-Hardegger, S. (2016). Higher fines – fewer traffic offences? A multi-site observational study. *European Journal on Criminal Policy and Research*, 22(4), 619-634. <https://doi.org/10.1007/s10610-016-9301-6>.
- Klaauw, B. van der, & Ours, J.C. van (2013). Carrot and Stick: How Re-Employment Bonuses and Benefit Sanctions Affect Exit Rates from Welfare. *Journal of Applied Econometrics*, 28(2), 275-296. <https://doi.org/10.1002/jae.1265>.
- Kleven, H.J., Knudsen, M.B., Kreiner, C.T., Pedersen, S., & Saez, E. (2011). Unwilling or unable to cheat? Evidence from a tax audit experiment in Denmark. *Econometrica*, 79(3), 651-692. <https://doi.org/10.3982/ECTA9113>.
- Kohn, A. (1993). *Punished by rewards: The trouble with gold stars, incentive plans, a's, praise, and other bribes*. Houghton Mifflin Harcourt.
- Kucher, M., & Götte, L. (1998). Trust me. An empirical analysis of taxpayer honesty. *FinanzArchiv / Public Finance Analysis*, 55(3), 429-444.
- Lalive, R., Ours, J.C. van, & Zweimüller, J. (2005). The effect of benefit sanctions on the duration of unemployment. *Journal of the European Economic Association*, 3(6), 1386-1417. <https://doi.org/10.1162/154247605775012879>.
- Locke, E.A., & Latham, G.P. (1990). *A theory of goal setting & task performance*. Prentice-Hall, Inc.
- Locke, E.A., & Latham, G.P. (2019). The development of goal setting theory: A half century retrospective. *Motivation Science*, 5(2), 93-105. <https://doi.org/10.1037/mot0000127>.
- Machin, S., & Marie, O. (2006). Crime and benefit sanctions. *Portuguese Economic Journal*, 5(2), 149-165. <https://doi.org/10.1007/s10258-006-0010-9>.
- Maslow, A.H. (1943). A theory of human motivation. *Psychological Review*, 50(4), 370-396. <https://doi.org/10.1037/h0054346>.
- McVicar, D. (2008). Job search monitoring intensity, unemployment exit and job entry: Quasi-experimental evidence from the UK. *Labour Economics*, 15(6), 1451-1468. <https://doi.org/10.1016/j.labeco.2008.02.002>.
- Moller, A., & Deci, E.L. (2014). The psychology of getting paid: An integrated perspective. In E. Bijleveld & H. Aarts (Eds.), *The psychological science of money*. Springer. [https://link.springer.com/chapter/10.1007/978-1-4939-0959-9\\_9](https://link.springer.com/chapter/10.1007/978-1-4939-0959-9_9).
- Mulder, L.B., Dijk, E. van, Cremer, D. de, & Wilke, H.A.M. (2006). Undermining trust and cooperation: The paradox of sanctioning systems in social dilemmas. *Journal of Experimental Social Psychology*, 42(2), 147-162. <https://doi.org/10.1016/j.jesp.2005.03.002>.
- Mulder, L.B., Vegt, G. van der, & Ponsioen, S. (2014). *Omgaan met regelovertreding in vertrouwensrelaties: Effecten van dialoog- en sanctiefactoren en de invloed van de relatie tussen toezichthouder en ondernemer* (Uitgave in de reeks Handhaving en Gedrag). Boom Lemma uitgevers.
- Müller, K.-U., & Steiner, V. (2008). *Imposed Benefit Sanctions and the Unemployment-to-Employment Transition – The German Experience* (DIW Discussion Papers nr. 792).
- Muller, R.Z. (2018). *The tyranny of metrics*. Princeton University Press.
- Nelissen, R.M.A., & Mulder, L.B. (2013). What makes a sanction 'stick'? The effects of financial and social sanctions on norm compliance. *Social Influence*, 8(1), 70-80. <https://doi.org/10.1080/15534510.2012.729493>.

- O'Donohue, W., & Ferguson, K.E. (2001). *The psychology of B F Skinner*. Sage.
- Ouwerkerk, J., Harreveld, F. van, & Pligt, J. van der (2009). *Streng doch rechtvaardig? Interactie tussen handhaaver en gehandhaafde als determinant van regelnaleving, strafacceptatie en perceptie van rechtvaardigheid* (Uitgave in de reeks Handhaving en Gedrag). Boom Juridische uitgevers.
- Pink, D.H. (2011). *Drive: The surprising truth about what motivates us*. Riverhead.
- Pligt, J. van der, Koomen, W., & Harreveld, F. van (2007). *Bestrafen, belonen en beïnvloeden: Een gedragswetenschappelijk perspectief op handhaving* (Uitgave in de reeks Handhaving en Gedrag). Boom Juridische uitgevers.
- Prooijen, J.-W. van, Gallucci, M., & Toeset, G. (2008). Procedural justice in punishment systems: Inconsistent punishment procedures have detrimental effects on cooperation. *British Journal of Social Psychology*, 47(2), 311-324. <https://doi.org/10.1348/014466607X218212>.
- Ramos, M., Thijssen, L., & Coenders, M. (2021). Labour market discrimination against Moroccan minorities in the Netherlands and Spain: A cross-national and cross-regional comparison. *Journal of Ethnic and Migration Studies*, 47(6), 1261-1284. <https://doi.org/10.1080/1369183X.2019.1622824>.
- Ryan, R.M., & Deci, E.L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68-78. <https://doi.org/10.1037/0003-066X.55.1.68>.
- Sampat, B., & Williams, H. (2019). How Do Patents Affect Follow-On Innovation? Evidence from the Human Genome. *American Economic Review*, 109(1), 203-236. <https://doi.org/10.3386/w21666>.
- Saultz, A., Murphy, K.M., & Aronson, B. (2016). What can we learn from the Atlanta cheating scandal? *Phi Beta Kappan*, 97, 48-52.
- Schmelz, K. (2021). Enforcement may crowd out voluntary support for COVID-19 policies, especially where trust in government is weak and in a liberal society. *Proceedings of the National Academy of Sciences*, 118(1). <https://doi.org/10.1073/pnas.2016385118>.
- Schmitz, L., Bijleveld, E., & Veling, H. (2020). Cognitive labor shapes the desire for social and monetary compensation. *Motivation and Emotion*, 44(6), 797-809. <https://doi.org/10.1007/s11031-020-09856-0>.
- Skinner, B.F. (1971). *Beyond freedom and dignity*. Hackett.
- Svarer, M. (2011). The Effect of Sanctions on Exit from Unemployment: Evidence from Denmark. *Economica*, 78, 751-778.
- Sykes-Muskett, B.J., Prestwich, A., Lawton, R.J., & Armitage, C.J. (2015). The utility of monetary contingency contracts for weight loss: A systematic review and meta-analysis. *Health Psychology Review*, 9(4), 434-451. <https://doi.org/10.1080/17437199.2015.1030685>.
- Thorndike, E.L. (1949). *Personnel selection*. Wiley.
- Verboon, P., & Dijke, M. van (2011). When do severe sanctions enhance compliance? The role of procedural fairness. *Journal of Economic Psychology*, 32(1), 120-130. <https://doi.org/10.1016/j.joep.2010.09.007>.
- Vliek, M. (2017). *Schandpaal of schouderklop? Handhaving door 'naming en shaming' en 'naming en faming'* (Uitgave in de reeks Handhaving en Gedrag). Boom criminologie.

## BIJLAGE A: MODELSPECIFICATIE

We passen een *two-stage least squares* (2SLS) procedure toe om de causale effecten van sancties en waarschuwingen te schatten. In deze procedure gebruiken we de strengheid van adviseurs als instrumentele variabele voor een daadwerkelijke behandeling.

In de eerste stap (*first stage*) schatten we de verbanden tussen de instrumenten ( $Z$ ) en de behandelingen ( $X$ ). Hierbij wordt een nieuwe variabele gecreëerd ( $\hat{X}$ ), die we in de tweede stap kunnen gebruiken als vervanger voor de ‘problematische’ variabele  $X$ . De vergelijkingen in de eerste stap zien er als volgt uit:

$$X_{i,0}^{\text{Sanctie}} = \gamma Z_{j,-i}^{\text{Sanctie}} + \alpha Z_{j,-i}^{\text{Waarschuwing}} + \mu Z_{j,-i}^{\text{Geen}} + C_i' \theta + \pi_i + \nu_{i,0} \quad (1)$$

$$X_{i,0}^{\text{Waarschuwing}} = \pi Z_{j,-i}^{\text{Sanctie}} + \sigma Z_{j,-i}^{\text{Waarschuwing}} + \tau Z_{j,-i}^{\text{Geen}} + C_i' \theta + \pi_i + \varepsilon_{i,0} \quad (2)$$

$$X_{i,0}^{\text{Geen}} = \phi Z_{j,-i}^{\text{Sanctie}} + \psi Z_{j,-i}^{\text{Waarschuwing}} + \phi Z_{j,-i}^{\text{Geen}} + C_i' \theta + \pi_i + \vartheta_{i,0} \quad (3)$$

De variabele  $X_{i,0}^{\text{Behandeling}}$  is een dummyvariabele die aangeeft of een WW-gerechtigde  $i$  de respectievelijke behandeling heeft ontvangen.  $Z_{i,0}^{\text{Behandeling}}$  is de instrumentele variabele. De instrumentele variabele kwantificeert de strengheid van adviseur  $j$ , aan wie de WW-gerechtigde  $i$  werd toegewezen, voor alle casussen, behalve de casus van  $i$  zelf. Deze index wordt ook wel *leave-out mean* genoemd. De vector  $C_i'$  is een set van onafhankelijke variabelen met daarin geslacht, leeftijd, opleidingsniveau, sector, maandloon laatste baan, resterende WW-maanden en twee indicatoren voor werk respectievelijk ontvangen van een WW-uitkering in  $t-1$ .  $\pi_i$  zijn kantoor- en maandspecifieke effecten.  $\nu_{i,0}$ ,  $\varepsilon_{i,0}$  en  $\vartheta_{i,0}$  zijn storingstermen.

In de tweede stap (*second stage*) schatten we het verband tussen behandeling ( $X$ ) en uitkomst ( $Y$ ). In deze stap gebruiken we  $\hat{X}$  in plaats van  $X$  als onafhankelijke variabele. De vergelijking voor de tweede stap ziet er als volgt uit:

$$Y_{i,t} = \delta_{1t} \hat{X}_{i,0}^{\text{Sanctie}} + \delta_{2t} \hat{X}_{i,0}^{\text{Waarschuwing}} + \delta_{3t} \hat{X}_{i,0}^{\text{Geen}} + C_i' \theta_t + \pi_i + \omega_{i,t} \quad (4)$$

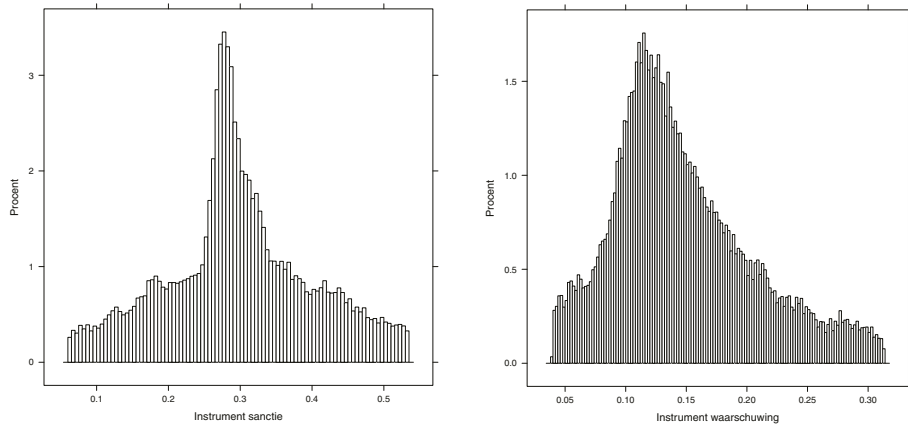
In deze vergelijking is  $Y_{i,t}$  de respectievelijke uitkomst in maand  $t$  na overtreding. De overige factoren zijn dezelfde als in vergelijking 1 tot en met 3. We zijn geïnteresseerd in de parameters  $\delta_{1t}$  en  $\delta_{2t}$ . Deze parameters schatten het causale effect van een sanctie respectievelijk waarschuwing op een uitkomst in

maand  $t$ . De causale effecten zijn *local average treatment effecten* (LATE), dat wil zeggen: het effect van een sanctie respectievelijk waarschuwing voor degenen die een andere behandeling hadden kunnen krijgen als zij terecht zouden zijn gekomen bij een andere adviseur.



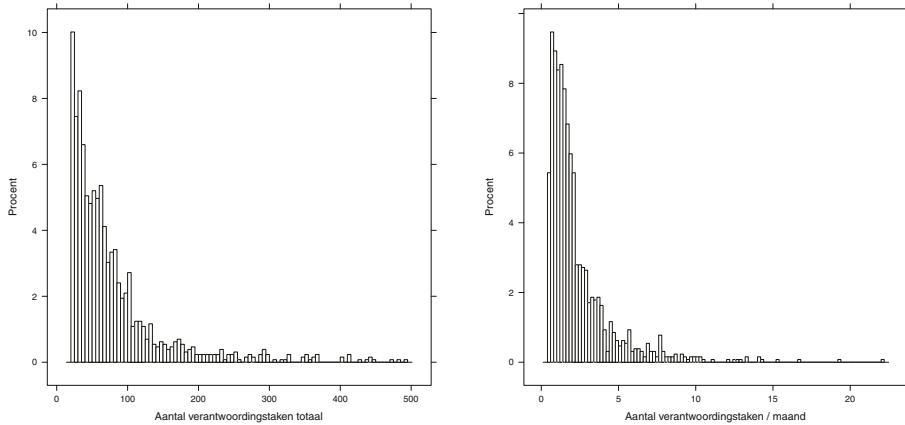
## BIJLAGE B: AANVULLENDE FIGUREN

**Figuur B.1** Variatie in instrumenten per kantoor en maand voor sancties (paneel A) en waarschuwingen (paneel B)



*Noot:* De geplotte waarden zijn residuen van een regressie waarbij de respectievelijke instrumentele variabele de afhankelijke variabele is en kantoor/maand *fixed effects* en achtergrondkenmerken van WW-gerechtigden de onafhankelijke variabelen. De residuen zijn *mean-standardized* en de hoogste en laagste 5% zijn verwijderd. De figuur volgt Bhuller et al. (2020).

**Figuur B.2** Verdeling van het totale aantal verantwoordingsstaken (paneel A) en het aantal verantwoordingsstaken per maand (paneel B) per adviseur



*Noot:* Veertien adviseurs (1%) hebben in totaal meer dan 500 verantwoordingsstaken afgehandeld. Deze adviseurs zijn niet weergegeven.

## BIJLAGE C: AANVULLENDE TABELLEN

Tabel C.1 Relevantie van de first stage

	Kans op			
	Sanctie	Sanctie	Waarschuwing	Waarschuwing
	(1)	(2)	(3)	(4)
Instrument sanctie	0,525***	0,517***	0,026*	0,026*
	(0,021)	(0,021)	(0,012)	(0,012)
Instrument waarschuwing	-0,119***	-0,115***	0,905***	0,890***
	(0,026)	(0,026)	(0,031)	(0,030)
Controle kantoor/maand	Ja	Ja	Ja	Ja
Controle kenmerken	Nee	Ja	Nee	Ja
Gemiddelde afhankelijke variabele	0,249	0,249	0,181	0,181
F-statistiek	105	107	654	630
R <sup>2</sup> (adj.)	0,328	0,336	0,230	0,240
N	51.848	51.848	51.848	51.848

*Noot:* Elke kolom rapporteert de geschatte coëfficiënten van een OLS-regressie. De afhankelijke variabelen zijn een dummy voor de respectievelijke behandeling (sanctie of waarschuwing). De onafhankelijke variabelen zijn de instrumenten (strengheid). Strengheid wordt gemeten als de *leave-out mean* van een adviseur. Standaardfouten worden geclusterd op niveau van de adviseurs ( $N = 1.303$ ) en gerapporteerd tussen haakjes.

\*  $p < 0,05$ .

\*\*  $p < 0,01$ .

\*\*\*  $p < 0,001$ .

Tabel C.2 Onafhankelijkheid van de instrumenten

	Kans op			
	Sanctie	Sanctie	Waarschuwing	Waarschuwing
	(1)	(2)	(3)	(4)
Vrouw	0,0434*** (0,0040)	0,0034* (0,0015)	-0,0245*** (0,0036)	-0,0024* (0,0009)
Leeftijd	-0,0011*** (0,0002)	-0,0000 (0,0001)	-0,0009*** (0,0002)	-0,0002** (0,0001)
Opleiding middelbaar	-0,0148*** (0,0044)	-0,0000 (0,0019)	0,0048 (0,0039)	0,0006 (0,0010)
Opleiding hoog	-0,0556*** (0,0058)	-0,0097*** (0,0026)	0,0280*** (0,0055)	0,0036* (0,0015)
Opleiding onbekend	-0,0205*** (0,0049)	-0,0020 (0,0023)	-0,0012 (0,0044)	0,0033* (0,0016)
Sector Zorg	-0,0023 (0,0067)	0,0006 (0,0027)	0,0069 (0,0062)	0,0026 (0,0018)
Sector Handel	-0,0055 (0,0065)	0,0025 (0,0025)	0,0118* (0,0060)	0,0038* (0,0017)
Sector Dienstverlening	-0,0089 (0,0061)	-0,0019 (0,0025)	0,0109 (0,0060)	0,0038* (0,0017)
Sector Transport	-0,0017 (0,0073)	0,0026 (0,0029)	0,0083 (0,0069)	0,0040* (0,0018)
Sector Uitzendbureau	-0,0021 (0,0069)	-0,0032 (0,0027)	0,0031 (0,0056)	0,0005 (0,0016)
Andere sector	-0,0041 (0,0061)	0,0040 (0,0024)	-0,0049 (0,0056)	0,0005 (0,0016)
Maandloon laatste baan (x € 1.000)	-0,0141*** (0,0020)	-0,0019* (0,0008)	0,0118*** (0,0020)	0,0028*** (0,0006)
Resterende WW-maanden (x 10)	-0,0126*** (0,0028)	-0,0018 (0,0012)	0,0090*** (0,0027)	0,0018* (0,0008)
Werkzaam ( <i>t</i> -1)	0,0062 (0,0036)	-0,0022 (0,0014)	-0,0116** (0,0038)	0,0028* (0,0011)
WW-uitkering ( <i>t</i> -1)	0,0390*** (0,0052)	0,0011 (0,0021)	0,1004*** (0,0067)	0,0078*** (0,0014)
Controle kantoor/maand	Ja	Ja	Ja	Ja
<i>N</i>	51.848	51.848	51.848	51.848

*Noot:* Elke kolom rapporteert coëfficiënten van een OLS-regressie. In kolom 1 en 3 is de afhankelijke variabele een dummy voor de respectievelijke behandeling (sanctie of waarschuwing). In kolom 2 en 4 is de afhankelijke variabele het instrument (strengheid). De achtergrondkenmerken van WW-gerechtigden zijn de onafhankelijke variabelen. Alle variabelen voor achtergrondkenmerken, met uitzondering van leeftijd, resterende WW-maanden en maandloon, zijn dummyvariabelen. De referentiecategorie is *laag* voor opleidingsniveau en *industrie* voor sector. Strengheid wordt gemeten als de *leave-out mean* van een adviseur. Standaardfouten worden geclusterd op niveau van de adviseurs ( $N = 1.303$ ) en gerapporteerd tussen haakjes.

\*  $p < 0,05$ .

\*\*  $p < 0,01$ .

\*\*\*  $p < 0,001$ .

Tabel C.3 Exclusie: verband tussen instrumenten en aanvullende dienstverlening

	Onafhankelijke variabelen		Afhankelijke variabelen			
	Gemiddelde	SD	Kans op sanctie	Instrument sanctie	Kans op waarschuwing	Instrument waarschuwing
	(1)	(2)	(3)	(4)	(5)	(6)
Workshop	0,049	0,216	0,005 (0,008)	0,005 (0,003)	0,024** (0,008)	0,002 (0,002)
Coachingsgesprek	0,078	0,269	0,000 (0,007)	0,000 (0,003)	0,005 (0,007)	-0,002 (0,002)
Ontheffing	0,182	0,386	-0,021*** (0,004)	-0,000 (0,002)	0,017*** (0,005)	0,000 (0,001)
Gem. afhankelijke variabele			0,249	0,297	0,181	0,152
<i>N</i>			51.848	51.848	51.848	51.848

*Noot:* Kolom 1 en 2 tonen het gemiddelde en de standaardafwijking van de onafhankelijke variabelen. In kolom 3 en 5 is de afhankelijke variabele een dummy voor de respectievelijke behandeling (sanctie of waarschuwing). In kolom 4 en 6 is de afhankelijke variabele het instrument (strengheid). De drie dummyvariabelen voor aanvullende dienstverlening zijn de onafhankelijke variabelen. Deze variabelen geven aan of men in de zes maanden na overtreding in aanraking is gekomen met de respectievelijke dienstverlening. Alle regressies controleren voor kantoor, maand en achtergrondkenmerken. Strengheid wordt gemeten als de *leave-out mean* van een adviseur. Standaardfouten worden geclusterd op niveau van de adviseurs ( $N = 1.303$ ) en gerapporteerd tussen haakjes. De test op exclusie volgt Arni & Schiprowski (2019).

\*  $p < 0,05$ .

\*\*  $p < 0,01$ .

\*\*\*  $p < 0,001$ .

Tabel C.4 Monotonieit instrument sanctie

	Basisinstrument	Reverse-sample instrument
	Kans op sanctie	Kans op sanctie
	(1)	(2)
<b>Analyse-dataset</b>		
Instrument sanctie	0,517*** (0,021)	
Gemiddelde afhankelijke variabele	0,249	
<i>N</i>	51.848	
<b>Paneel A: Opleidingsniveau</b>		
<i>1. Sub-sample: Opleidingsniveau laag</i>		
Instrument sanctie	0,523*** (0,031)	0,641*** (0,030)
Gemiddelde afhankelijke variabele	0,305	0,305
<i>N</i>	12.596	12.589
<i>2. Sub-sample: Opleidingsniveau middelbaar</i>		
Instrument sanctie	0,545*** (0,026)	0,651*** (0,025)
Gemiddelde afhankelijke variabele	0,281	0,281
<i>N</i>	19.670	19.670
<i>3. Sub-sample: Opleidingsniveau hoog</i>		

	Basisinstrument	Reverse-sample instrument
	Kans op sanctie	Kans op sanctie
	(1)	(2)
Instrument sanctie	0,484***	0,578***
	(0,043)	(0,042)
Gemiddelde afhankelijke variabele	0,195	0,195
<i>N</i>	7.285	7.285
<b>Paneel B: Geslacht</b>		
<i>1. Sub-sample: Geslacht man</i>		
Instrument sanctie	0,499***	0,598***
	(0,027)	(0,025)
Gemiddelde afhankelijke variabele	0,219	0,219
<i>N</i>	24.618	24.613
<i>2. Sub-sample: Geslacht vrouw</i>		
Instrument sanctie	0,526***	0,637***
	(0,023)	(0,022)
Gemiddelde afhankelijke variabele	0,277	0,277
<i>N</i>	27.230	27.228
<b>Paneel C: Leeftijd</b>		
<i>1. Sub-sample: ≤ 40 jaar</i>		
Instrument sanctie	0,532***	0,645***
	(0,024)	(0,023)
Gemiddelde afhankelijke variabele	0,261	0,261
<i>N</i>	26.799	26.799
<i>2. Sub-sample: &gt; 40 jaar</i>		
Instrument sanctie	0,485***	0,586***
	(0,027)	(0,025)
Gemiddelde afhankelijke variabele	0,237	0,237
<i>N</i>	25.049	25.042
<b>Paneel D: Sectoren</b>		
<i>1. Sub-sample: Sector Industrie</i>		
Instrument sanctie	0,544***	0,659***
	(0,042)	(0,041)
Gemiddelde afhankelijke variabele	0,270	0,270
<i>N</i>	6.607	6.606
<i>2. Sub-sample: Sector Zorg</i>		
Instrument sanctie	0,506***	0,603***
	(0,041)	(0,040)
Gemiddelde afhankelijke variabele	0,230	0,230
<i>N</i>	7.096	7.094
<i>3. Sub-sample: Sector Handel</i>		
Instrument sanctie	0,534***	0,634***
	(0,037)	(0,037)
Gemiddelde. afhankelijke variabele	0,243	0,243

	Basisinstrument	Reverse-sample instrument
	Kans op sanctie	Kans op sanctie
	(1)	(2)
<i>N</i>	8.290	8.290
<i>4. Sub-sample: Sector Dienstverlening</i>		
Instrument sanctie	0,458***	0,563***
	(0,036)	(0,037)
Gemiddelde afhankelijke variabele	0,214	0,214
<i>N</i>	9.427	9.425
<i>5. Sub-sample: Sector Transport</i>		
Instrument sanctie	0,503***	0,619***
	(0,058)	(0,057)
Gemiddelde afhankelijke variabele	0,266	0,266
<i>N</i>	4.068	4.068
<i>6. Sub-sample: Sector Uitzendbureau</i>		
Instrument sanctie	0,541***	0,651***
	(0,043)	(0,042)
Gemiddelde afhankelijke variabele	0,291	0,291
<i>N</i>	6.234	6.233

*Noot:* Elke kolom geeft de resultaten weer van een toets op monotoniciteit. Kolom 1 toont de *first stage* coëfficiënten voor de volledige steekproef alsmede voor veertien sub-samples. Kolom 2 toont *first stage* coëfficiënten voor dezelfde veertien sub-samples, echter gebruikmakend van *reverse-sample* instrumenten. *Reverse-sample* instrumenten zijn de *leave-out mean* van een adviseur voor alle casussen buiten de respectievelijke sub-sample. Alle regressies controleren voor kantoor, maand en achtergrondkenmerken. Standaardfouten worden geclusterd op het niveau van de adviseurs ( $N = 1.303$ ) en gerapporteerd tussen haakjes. De toets op monotoniciteit volgt Bhuller et al. (2020).

\*  $p < 0,05$ .

\*\*  $p < 0,01$ .

\*\*\*  $p < 0,001$ .





## BIJLAGE D: BERICHTEN UWV

### Bericht 1: Verzoek om verantwoording

**Onderwerp:** Geef reden voor het doorgeven van onvoldoende sollicitatieactiviteiten

U heeft over de periode [StartDat] t/m [EindDat] te weinig sollicitatieactiviteiten doorgegeven. We horen graag wat hiervoor de reden is. Heeft u wel voldoende activiteiten gedaan, maar dit niet doorgegeven? Geef dan de reden hiervoor door en vul uw activiteiten aan in 5 stappen:

1. Klik op '+ Sollicitatieactiviteiten aanvullen'. U komt dan op een andere pagina.
2. Klik daar op 'Nieuwe sollicitatieactiviteit doorgeven' en vul in wat u gedaan heeft.
3. Klik op 'Terug naar taak' om terug te gaan naar deze pagina.
4. Vul onderaan deze taak uw reden in bij de toelichting.
5. Klik op 'Nu doorgeven'.

Heeft u niet voldoende activiteiten gedaan? Laat ons dan in 2 stappen weten wat de reden hiervoor is:

1. Vul onderaan deze taak uw reden in bij de toelichting.
2. Klik op 'Nu doorgeven'.

Met uw reactie kunt u voorkomen dat u tijdelijk minder uitkering krijgt. Geeft u geen reden door? Dan moeten we aannemen dat u geen geldige reden heeft. Wordt uw reden door ons gezien als niet geldig? Dan ontvangt u mogelijk tijdelijk minder uitkering.

**Bericht 2: Overtreding is niet-verwijtbaar**

**Onderwerp:** Geldige reden voor het niet uitvoeren van sollicitatietaak

Beste heer/mevrouw,

Wij vroegen u om ons te laten weten waarom u te weinig of te laat sollicitatie-activiteiten had doorgegeven. U heeft in uw reactie een geldige reden aan ons doorgegeven. Daarom zijn er geen gevolgen voor uw uitkering.

Lukt het u in de toekomst niet om een verplichting die bij uw uitkering hoort, na te komen? Dan is het belangrijk dat u zo snel mogelijk contact met ons opneemt. Een adviseur bekijkt dan samen met u wat er mogelijk is in uw situatie.

Met vriendelijke groet,  
UWV

**Bericht 3: Overtreding is verwijtbaar**

**Onderwerp:** Vervolg niet nagekomen verplichting

Beste heer/mevrouw,

U bent een verplichting die bij uw uitkering hoort, niet nagekomen. U heeft bijvoorbeeld een taak in uw Werkmap niet (op tijd) gedaan, of u bent niet naar een afspraak gekomen. Wij vroegen u om aan ons door te geven wat hiervoor de reden is.

Wij gaan nu beoordelen of het niet nakomen van de verplichting gevolgen heeft voor uw uitkering. Als u een reden heeft doorgegeven, dan gebruiken wij deze informatie bij onze beoordeling.

U ontvangt onze beslissing binnen 4 weken per brief.

Met vriendelijke groet,  
UWV