

Appendix 1: Puntstellingen en 95%-betrouwbaarheidsintervallen WAO, WW en ABW/ WWB 2004

Tabel 1.1/ Puntstellingen en 95%-betrouwbaarheidsintervallen overtredingen WAO 2004 (proporties)

<i>Regels met betrekking tot inkomsten uit arbeid (afgelopen jaar/ 12 maanden wel eens):</i>	Gemiddelde	Ondergrens	Bovengrens
- Kleinere klusjes gedaan voor geldelijke vergoeding voor of via bekenden zonder dat door te geven aan uitkeringsinstantie	0,181	0,137	0,230
- Baan of uitzendwerk gehad naast de uitkering zonder door te geven	-	-	-
- Zwart werk gedaan naast de uitkering	0,060	0,002	0,105
<i>Regels met betrekking tot gezondheidssituatie (wel eens):</i>			
- Constatering van geneesheer dat arbeidsongeschiktheid klachten zijn afgenomen en dit niet doorgegeven	0,056	0,016	0,099
- Zieker voorgedaan tijdens controle door uitkeringsinstantie	0,070	0,031	0,117
- Zelf herstel van arbeidsongeschiktheidsklachten bemerkt bij (vrijwilligers)werk of klusjes in huis zonder dit door te geven aan uitkeringsinstantie	0,137	0,094	0,185
- Langere tijd duidelijk sterker en gezonder voelen en in staat meer te werken zonder dit door te geven	0,121	0,079	0,166
<i>Sancties (afgelopen 12 maanden wel eens):</i>			
- Boete of korting gehad	0,067	0,027	0,110

Tabel 1.2/ Puntstellingen en 95%-betrouwbaarheidsintervallen overtredingen WW 2004 (proporties)

<i>Regels met betrekking tot inkomsten uit arbeid (afgelopen 12 maanden wel eens):</i>	Gemiddelde	Ondergrens	Bovengrens
• Kleinere klusjes voor geldelijke vergoeding voor of via bekenden zonder dat door te geven aan uitkeringsinstantie	0,218	0,176	0,265
• Zwart werk naast de uitkering	0,114	0,073	0,157
<i>Regels met betrekking tot medewerkingverplichtingen (afgelopen 12 maanden wel eens):</i>			
• Langer dan 4 weken op vakantie zonder dit door te geven	0,034	- 0,002	0,073
• Passend werk afgewezen of er bewust voor gezorgd om niet te worden aangenomen	0,124	0,085	0,167
• Langere tijd bewust weinig gesolliciteerd	0,251	0,208	0,297
• Dagopleiding gevolgd zonder door te geven aan uitkeringsinstantie	0,000	-0,035	0,035
• Onbetaald (vrijwilligers) werk zonder door te geven aan uitkeringsinstantie	0,255	0,210	0,301
<i>Sancties (afgelopen 12 maanden wel eens):</i>			
• Boete of korting gehad	0,206	0,16	0,251

*Tabel 1.3/ Puntschattingen en 95%-betrouwbaarheidsintervallen overtredingen ABW/WWB 2004
(proporties)*

	Gemiddelde	Ondergrens	Bovengrens
<i>Regels met betrekking tot inkomsten uit arbeid (afgelopen jaar/ 12 maanden wel eens):</i>			
- Kleinere klusjes gedaan voor geldelijke vergoeding voor of via bekenden zonder dat door te geven aan uitkeringsinstantie	0,277	0,234	0,325
- Baan of uitzendwerk gehad naast de uitkering zonder door te geven - Zwart werk gedaan naast de uitkering	0,128	0,090	0,174
<i>Regels met betrekking tot informatieve verplichtingen:</i>			
- Inkomsten gehad naast uitkering zonder op te geven (vb. alimentatie, studiebeurs, giften, rente) (afgelopen 12 maanden wel eens)	0,077	0,041	0,120
- Heeft spaarrekeningen, aandelen, waardepapieren, lijfrente, levensverzekering, koopsompolis niet doorgegeven	0,074	0,031	0,112
-Heeft waardevolle bezittingen niet doorgegeven ^a	0,063	0,022	0,100
-Heeft ooit onjuist of onvolledige inlichtingen gegeven over woonsituatie, gezinssamenstelling, burgerlijke staat	0,062	0,021	0,099
<i>Regels met betrekking tot medewerkingsverplichtingen:</i>			
- De afgelopen 12 maanden wel eens passend werk afgewezen of er bewust voor gezorgd om niet te worden aangenomen ^b	0,122	0,059	0,187
- In de afgelopen 12 maanden wel eens langere tijd bewust weinig gesolliciteerd ^b	0,434	0,364	0,504
<i>Sancties (afgelopen 12 maanden wel eens):</i>			
- Boete of korting gehad	0,136	0,095	0,179

^a in 2002 alleen Capi@home, n = 253

^b in 2002 alleen Capi@home, met sollicitatieplicht n = 79

Appendix 2: steekproef en uitvoering van het veldwerk

Voor de tweede en derde meting van het onderzoek (POROSZ 2 en 3) is gekozen voor benadering via TNS NIPObase CASI. Dit is een zogenaamde *accesspool* van respondenten (huishoudens) die vragenlijsten via een modemverbinding ontvangen en ingevuld weer terugzenden aan het TNS-NIPO. De *accesspool* bestaat thans uit circa 35.000 huishoudens met circa 100.000 individuen. De *access pool* is representatief op de binnen onderzoek gebruikelijke achtergrondkenmerken en is opgebouwd op basis van de daarover bekende CBS gegevens en gegevens uit de Minicensus. Deelnemers aan de *access pool* zijn en worden geworven via onderzoek met traditionele instrumenten waarbij sprake is van a-selecte landelijk gespreide steekproeven. Door aan moeilijke doelgroepen extra aandacht te besteden wordt er naar gestreefd deze groepen in de juiste proportie in de *access pool* op te nemen.

Steekproef, uitvoering en respons

In het onderzoek zijn mensen met een WAO- of andere arbeidsongeschiktheidsuitkering betrokken (WAZ en Wajong), mensen met een WW-uitkering en mensen met een ABW/WWB-uitkering.

De TNS-NIPO *access pool* bevatte te weinig respondenten met een ABW/WWB uitkering. Een deel van de enquêtes is daarom afgenomen via een *pool* van het marktonderzoekbureau GfK PanelServices Benelux bv dat de enquêtes voor het parallel lopende onderzoek naar naleving en overtreding van de regels van de volksverzekeringen heeft verzorgd. De *pool* van GfK heeft een vergelijkbare opzet als de *access pool* van TNS-NIPO.

Tabel 2.1 geeft een overzicht van de respons op de enquête. Er blijken betrekkelijk veel respondenten te zijn die de enquête hebben ingevuld zonder dat deze voor hen was bedoeld

Er is aan het begin van de vragenlijst een lijstje uitkeringen aan de respondenten voorgelegd waarbij ze konden aangeven welke van de uitkering(en) ze ontvangen.

Overigens gaat het zeker niet om respondenten die de gehele enquête zouden hebben ingevuld terwijl deze niet voor hen bedoeld was, maar enkel om de twee eerste (filter)vragen. Dit is een gevolg van de manier waarop de screening is ingericht. Het *accesspanel* is een panel op huishoud-niveau. Het komt voor dat de screeningsvragen worden beantwoord door het ene lid van het huishouden voor een ander lid van het huishouden. Vervolgens wordt bij deze de enquête afgenomen. Het kan dan blijken dat het eerste lid niet goed geïnformeerd was.

Wij geven geen responspercentages omdat er bij het trekken van de steekproef van een *access panel* is gebruik gemaakt. De relatie tussen een *access panel* en de populatie waarnaar gegeneraliseerd dient te worden is niet helder, omdat de samenstelling van het panel niet op een statistisch gecontroleerde wijze plaatsvindt. In POROSZ 2 (2002) zijn voor de WAO zowel *access panel* data als data verkregen uit een traditionele steekproef verzameld. Na weging naar populatiekenmerken bleken deze twee typen data grote overeenstemming te vertonen. Om deze reden kan er vertrouwen worden gesteld in data verzameld met een *access panel*.

Globaal aangegeven heeft ruim de helft van de benaderde respondenten binnen de veldwerkperiode een bruikbare vragenlijst ingevuld.

Tabel 2.1| Respons Capi@home WAO, WW, en ABW/ WWB 2004

	Wet:	Wao	WW	ABW/ WWB
		N	N	N
a.	Totaal uitgezet	1556	1599	1543
b.	Totaal ingevuld binnen veldwerkperiode	1273	1164	1148
		1273	1164	1148
c.	Niet de betreffende uitkering	443	294	268
d.	Netto enquêtes	830	870	880

Representativiteit en weging

De uitkomsten van het onderzoek zijn gewogen met cijfers van het CBS over uitkeringsgerechtigden in de verschillende regelingen (WAO, WW en ABW/WWB) in maart 2004. Ook bij de analyse van de verbanden tussen variabelen met logistische regressiemodellen en met het programma Answer Tree is gewerkt met gewogen cijfers.

De representativiteit van de ondervraagde steekproeven voor de WAO en de WW blijkt uit de beperkte correcties die noodzakelijk waren bij de weging voor deze regelingen (zie tabellen 2.2 en 2.3). De steekproef voor de ABW/WWB is minder representatief (tabel 2.4). Deze populatie staat bekend als moeilijk toegankelijk voor enquêtes en was ook in de eerdere onderzoeken minder evenwichtig in de steekproef vertegenwoordigd. Weging heeft hier grotere correcties tot gevolg.

Tabel 2.2| WAO-steekproef naar geslacht, leeftijd en mate van arbeidsongeschiktheid, percentages van totale steekproef voor en na weging op basis van CBS-cijfers 2004 (n = 830)

Leeftijd	Mannen				Vrouwen			
	Gedeeltelijk arbeidsongeschikt		Geheel arbeidsongeschikt		Gedeeltelijk arbeidsongeschikt		Geheel arbeidsongeschikt	
	Ongewogen %	Gewogen %	Ongewogen %	Gewogen %	Ongewogen %	Gewogen %	Ongewogen %	Gewogen %
15-24	-	-	0	2	-	-	0	2
25-34	1	1	1	3	2	1	4	4
35-44	2	3	3	6	5	3	9	7
45-54	5	6	10	9	6	4	13	9
55-64	7	8	17	17	4	3	12	12

Tabel 2.3 | WW steekproef naar geslacht en leeftijd, percentages van totale steekproef voor en na weging op basis van CBS-cijfers 2004 (n = 870)

Leeftijd	Mannen		Vrouwen	
	Ongewogen %	Gewogen %	Ongewogen %	Gewogen %
15-24	2	4	2	3
25-34	7	12	13	11
35-44	11	14	13	11
45-54	13	13	10	9
55-64	22	16	7	7

Tabel 2.4 | ABW/WWB steekproef naar veldwerkbureau, geslacht en leeftijd, percentages van totale steekproef voor en na weging op basis van CBS-cijfers 2004

Leeftijd	TNS-NIPO (n = 578)				GfK (n = 302)			
	Mannen		Vrouwen		Mannen		Vrouwen	
	<i>Ongewogen</i> %	<i>Gewogen</i> %	<i>Ongewogen</i> %	<i>Gewogen</i> %	<i>Ongewogen</i> %	<i>Gewogen</i> %	<i>Ongewogen</i> %	<i>Gewogen</i> %
Tot 25	2	4	4	5	2	4	4	6
25-34	4	9	20	12	9	9	26	12
35-44	4	11	24	14	7	11	20	14
45-54	6	9	16	11	6	9	13	12
55-64	5	8	11	11	5	8	9	11
65+	1	3	2	3	-	-	0	4

Appendix 3: Lineaire trend

Inleiding

Het POROSZ onderzoek is drie maal uitgevoerd, namelijk in het jaar 2000, 2002 en 2004. Dit maakt het mogelijk om te onderzoeken of de percentages regelovertreiding trendmatig door de tijd heen verandert. Men spreekt hier dan ook van trendanalyses.

Voor elke regelovertreiding van elke wet afzonderlijk wordt een trendanalyse uitgevoerd. De vraagstelling laat zich het gemakkelijkst illustreren aan de hand van Tabel 3.1 van deze Appendix, die grotendeels identiek is aan Tabel 1 uit de algemene samenvatting van het rapport. In de WAO zijn voor de regelovertreiding “kleinere klusjes uitgevoerd” van 2000-2002-2004 de schattingen 26%, 21% en 18% geconstateerd. De vraag die wij hier willen beantwoorden is of de verandering door de tijd heen trendmatig is.

Hierbij dient te worden opgemerkt dat de schattingen in Tabel 3.1 tot stand zijn gekomen met de methodologie die besproken is in de Appendix getiteld “Het gehanteerde model voor de schatting van de kans op regelovertreiding.”. Relevant is hier verder dat tussen 2000 en 2002 een methodewisseling heeft plaatsgevonden waardoor de verandering in de schattingen tussen 2000 en 2002 met enige voorzichtigheid dienen te worden geïnterpreteerd. Hierdoor dienen ook de uitkomsten van deze trendanalyses met de nodige voorzichtigheid te worden geïnterpreteerd.

Methoden

Wij hanteren de standaard aanpak bij het onderzoeken van trend binnen percentages zoals beschreven in Hagnaars (1990).¹ Deze standaardaanpak is door ons aangepast om rekening te houden met het gegeven dat de percentages met wisselende Randomized response designs zijn verzameld, namelijk in 2000 met het zgn. Kuk design (speelkaarten) en 2002 en 2004 met forced choice (dobbelstenen).² Daarnaast zijn de frequenties aangepast door rekening te houden met de aanwezigheid van weigeraars in de data (vergelijk de Appendix getiteld “Het gehanteerde model voor de schatting van de kans op regelovertreiding.”).

Resultaten

De resultaten zijn opgenomen in Tabel 3.1.

In de WAO is er lineaire afname van de regelovertreiding voor “Kleine klusjes”, “Zwart werk”, en “in staat meer te werken zonder dit door te geven”. Voor de overige regelovertreidingen is geen sprake van een significante af- of toename.

In de WW is er alleen een lineaire afname van de regelovertreiding voor “Passend werk afgewezen of er bewust voor gezorgd om niet te worden aangenomen” en “Dagopleiding zonder toestemming”.

Voor de overige regelovertreidingen is geen sprake van een significante af- of toename. In de WWB is er nergens een lineaire af- of toename, waarbij opgemerkt dient te worden dat voor “Passend werk afwijzen” de steekproeven klein zijn omdat niet alle respondenten sollicitatieplicht hebben. Indien getoetst wordt op 10 % niveau is deze verandering wel significant.

¹ J.A.P. Hagnaars (1990). *Categorical longitudinal data: log-linear panel, trend and cohort analysis*. Newbury Park, CA: Sage. Standaardvormen van deze modellen worden besproken in Agresti (2002, sectie 5.3). In de toets procedure wordt allereerst een model aangepast op de data met een lineaire trend (model 1). Indien dit model past, dan wordt vervolgens, *gegeven dit model*, het model getoetst waarin er geen lineaire trend is (model 2). Deze toets heeft een hogere power dan de toets van model 2 tegen de data. Model 1 paste in alle gevallen, behalve bij “Zelf bij werk bemerkt herstel van arbeidsgeschiktheid niet doorgegeven” in de WAO, “klusjes doen” in de WW, en “vrijwilligerswerk zonder toestemming” in de WW. In deze drie gevallen is de verandering niet. In alle andere gevallen kon de conditionele toets van model 2 gegeven dat model 1 waar is zinvol worden uitgevoerd.

² Zie Van den Hout, A., en P.G.M. van der Heijden. (2004) *The analysis of multivariate misspecified data, with special attention to randomized response data*. *Sociological Methods and Research*, 32, 310-336; Van den Hout, A. (ter publicatie aangeboden). *On the use of a 2x2 design where randomized response is one of the factors*.

Tabel 3.1 Geschatte omvang van regelovertreding voor WAO, WW en ABW/ WWB, 2000, 2002 en 2004

	WAO				WW				Abw		WWB	
	2000	2002	2004	Lin. trend	2000	2002	2004	Lin. trend	2000	2002	2004	Lin. trend
<i>Regels mbt inlichtingen over:</i>	1308	1760	830		1126	955	870		1211	761	880	
<i>Inkomsten uit arbeid (afgelopen 12 maanden wel eens):</i>	%	%			%	%			%	%		
• Kleinere klusjes voor geldelijke vergoeding	26	21	18	Ja	23	16	22	Nee	30	25	28	Nee
• Zwart werk naast uitkering	15	11	6	Ja	13	12	11	Nee	17	15	13	Nee
<i>Medewerkingverplichting en (afgelopen 12 maanden wel eens):</i>												
• Langer dan 4 weken op vakantie zonder toestemming	-	-	-		7	2	3	Nee	-	-	-	
• Passend werk afwijzen	-	-	-		19	12	12	Ja	20	18	12	Nee
• Weinig solliciteren	-	-	-		31	24	25	Nee	49	40	43	Nee
• Dagopleiding zonder toestemming	-	-	-		7	0	0	Ja	-	-	-	
• Vrijwilligerswerk zonder toestemming	-	-	-		20	29	26	Nee	-	-	-	
<i>informatieve verplichtingen met betrekking tot:</i>												
• Inkomsten naast uitkering (bv. alimentatie, studiebeurs) (afgelopen 12 maanden wel eens)	-	-	-		-	-	-		13	12	8	Nee
• Vermogen (spaarrekeningen, waardepapieren, etc.)	-	-	-		-	-	-		12	10	7	Nee
• Waardevolle bezittingen	-	-	-		-	-	-		8	3	6	Nee
• woonsituatie, samenstelling huishouden	-	-	-		-	-	-		6	5	6	Nee
<i>Inlichtingen met betrekking tot gezondheidssituatie:</i>												
• Medische diagnose afname arbeidsongeschiktheid niet doorgegeven	10	5	6	Nee	-	-	-		-	-	-	
• Zieker voorgedaan tijdens controle	12	5	7	Nee	-	-	-		-	-	-	
• Zelf bij werk bemerkt herstel van arbeidsongeschiktheid niet doorgegeven	16	10	14	Nee	-	-	-		-	-	-	
• In staat meer te werken zonder dit door te geven	26	17	12	Ja	-	-	-		-	-	-	

Appendix 4: nalevingsmonitor 2002-2004

In het rapport worden in tabel 1 voor elke regelovertreiding afzonderlijk de schattingen van de percentages regelovertreiding gepresenteerd. In deze appendix wordt de tabel gegeven als Tabel 4.1.

Vanuit beleidsmatig oogpunt is het wenselijk om voor *groepen* regelovertreidingen de naleving vast te stellen. Hierbij is naleving gedefinieerd als het percentage respondenten dat geen enkele regel uit de onderzochte groep regels overtreedt. De schatting van dit percentage is mogelijk met het model dat besproken wordt in Appendix 5: “Het gehanteerde model voor de schatting van de kans op regelovertreiding”.

Er zijn nalevingsmaten geschat voor de volgende groepen regels (vergelijk Tabel 4.1):

1. WAO: regels met betrekking tot het rapporteren van *inkomen uit arbeid*, d.w.z. de regels:
 - Kleinere klusjes voor geldelijke vergoeding
 - Baan of uitzendwerk naast uitkering
 - Zwart werk naast uitkering
2. WAO: regels met betrekking tot het rapporteren van de *gezondheidstoestand*, d.w.z. de regels:
 - Medische diagnose afname arbeidsongeschiktheid niet doorgegeven
 - Zieker voorgedaan tijdens controle
 - Zelf bij werk bemerkt herstel van arbeidsgeschiktheid niet doorgegeven
 - In staat meer te werken zonder dit door te geven
3. WW: regels met betrekking tot het rapporteren van *inkomen uit arbeid*, d.w.z. de regels:
 - Kleinere klusjes voor geldelijke vergoeding
 - Baan of uitzendwerk naast uitkering
 - Zwart werk naast uitkering
4. WW: *medewerkingsverplichtingen*, d.w.z. de regels:
 - Passend werk afwijzen
 - Weinig solliciteren
5. Abw: regels met betrekking tot rapporteren van *inkomen uit arbeid*, d.w.z. de regels:
 - Kleinere klusjes voor geldelijke vergoeding
 - Baan of uitzendwerk naast uitkering
 - Zwart werk naast uitkering
6. Abw: regels met betrekking tot rapporteren van *inkomen* (uit arbeid en andere bronnen), d.w.z. de regels:
 - Kleinere klusjes voor geldelijke vergoeding
 - Baan of uitzendwerk naast uitkering
 - Zwart werk naast uitkering
 - Inkomsten naast uitkering

Tabel 4.1 Geschatte omvang van regelovertreiding voor WAO, WW en ABW/WWB, 2000, 2002 en 2004

	Jaar	WAO							Abw		ABW/ WWB		
		2000	2002	2004	2000	2002	2004	2000	2002	2004			
<i>Regels mbt inlichtingen over:</i>													
<i>Inkomsten uit arbeid (afgelopen 12 maanden wel eens):</i>		%	%	%	%	%	%	%	%	%	%		
• Kleinere klusjes voor geldelijke vergoeding		26	21	18	*1,3	23	16	22	*1,2	30	25	28	
• Baan of uitzendwerk naast uitkering		14	3	-	*1	10	2	-	*1	10	2	-	*1
• Zwart werk naast uitkering		15	11	6	*2,3	13	12	11		17	15	13	
<i>Medewerkingverplichtingen (afgelopen 12 maanden wel eens):</i>													
• Langer dan 4 weken op vakantie zonder toestemming		-	-	-		7	2	3		-	-	-	
• Passend werk afwijzen		-	-	-		19	12	12	*1,3	20	18	12	
• Weinig solliciteren		-	-	-		31	24	25	*1,3	49	40	43	
• Dagopleiding zonder toestemming		-	-	-		7	0	0	*1,3	-	-	-	
• Vrijwilligerswerk zonder toestemming		-	-	-		20	29	26	*1	-	-	-	
<i>Informatieve verplichtingen met betrekking tot:</i>													
• Inkomsten naast uitkering (bv. alimentatie, studiebeurs) (afgelopen 12 maanden wel eens)		-	-	-		-	-	-		13	12	8	
• Vermogen (spaarrekeningen, waardepapieren, etc.)		-	-	-		-	-	-		12	10	7	
• Waardevolle bezittingen		-	-	-		-	-	-		8	3	6	
• Woonsituatie, samenstelling huishouden		-	-	-		-	-	-		6	5	6	
<i>Inlichtingen met betrekking tot gezondheidssituatie:</i>													
• Medische diagnose afname arbeidsongeschiktheid niet doorgegeven		10	5	6	*1	-	-	-		-	-	-	
• Zieker voorgedaan tijdens controle		12	5	7	*1	-	-	-		-	-	-	
• Zelf bij werk bemerkt herstel van arbeidsongeschiktheid niet doorgegeven		16	10	14	*1	-	-	-		-	-	-	
• In staat meer te werken zonder dit door te geven		26	17	12	*1,3	-	-	-		-	-	-	
<i>Sancties (afgelopen 12 maanden wel eens):</i>													
• Boete of korting gehad				7				21				14	

N.B.: De percentages in de tabel zijn dus niet optelbaar tot een algemene maat voor overtreiding, hetgeen hierna wordt toegelicht.

* = significant verschil op 5%-niveau.

1 = verschil 2000 – 2002; 2 = verschil 2002 – 2004; 3 = verschil 2000 – 2004. De verschillen tussen de jaren 2002 en 2000 zijn mogelijk (deels) te verklaren door een verandering in Randomized response methode.

Voordat wij de schattingen presenteren, maken wij de volgende kantekeningen:

- In Appendix 5: “Het gehanteerde model voor de schatting van de kans op regelovertreding” wordt aangegeven dat het hiervoor gehanteerde model relatief minder stabiel was voor de gegevens uit het jaar 2000. Daarom zijn alleen schattingen gemaakt voor het jaar 2002 en 2004.
- Voor de WAO zijn voor het jaar 2002 twee schattingen gemaakt, namelijk een schatting voor de FtF steekproef en een schatting voor de Capi steekproef. Deze leiden, in drie cijfers nauwkeurig, tot exact dezelfde schatting.
- Om tot vergelijkbare schattingen te komen, zijn de percentages van degenen die niet volgens het randomized-response-design antwoorden als volgt gefixeerd (vgl. de Appendix “Het gehanteerde model voor de schatting van de kans op regelovertreding”): WAO werk Capi: 20%; WAO werk FtF: 13 %; WAO gezondheid: 15%; de overige vier groepen regels: 10 %.

De resultaten staan gegeven in Tabel 4.2. Per groep regelovertradingen worden de volgende schattingen gepresenteerd:

- het nalevingspercentage voor de groep regels; bijvoorbeeld, voor WAO, inkomen uit arbeid, is dit voor de Capi steekproef in 2002 76.5 %; dit wil zeggen dat 76.5 % van de respondenten geen enkele van de regels op het gebied van Inkomen uit arbeid overtreedt; in andere woorden: $100 - 76.5 = 23.5$ % van de respondenten overtreedt minimaal een van de regels op het gebied van Inkomen uit arbeid.
- het 95-% betrouwbaarheidsinterval berekend met een niet-parametrische methode, de zgn. bootstraphmethode. Voor de schatting van 76.5 is het betrouwbaarheidsinterval dus 70.6 – 82.4.

Tabel 4.2 *Nalevingsmaten 2002 en 2004*

		2002		2004	
		%	95-% interval	%	95-% interval
<i>WAO</i>					
Inkomen uit arbeid	Capi	76.5	70.6-82.4	81.1	74.2-87.3
	FtF	76.5	69.4-83.2		
Gezondheid		76.9	71.0-82.9	77.8	69.5-87.8
<i>WW</i>					
Inkomen uit arbeid		79.6	73.7-86.0	74.6	68.0-80.6
Medewerkingsverplichtingen		68.6	62.8-74.0	65.5	60.4-71.2
<i>ABW/WWB</i>					
Inkomen uit arbeid		72.3	63.9-79.4	71.1	64.2-77.3
Inkomen		67.3	58.5-77.8	70.7	63.8-77.0

Tabel 4.2 leidt tot de conclusie dat de veranderingen tussen 2002 en 2004 niet significant verschillen.

Appendix 5: het gehanteerde model voor de schatting van de kans op regelovertreding

Inleiding

Voor het schatten van de prevalentie van regelovertreding wordt door het Ministerie van Sociale Zaken en Werkgelegenheid gebruik gemaakt van vragenlijstonderzoek. Hiermee wordt in aanvulling op de uit de administratieve bestanden reeds bekende omvang van fraudeconstateringen ook kwantitatief inzicht verkregen in het feitelijk voorkomen van fraudegedrag. Deze lijn van onderzoek wordt ook gekozen door verschillende andere Ministeries, en ondersteund door het Expertisecentrum Rechtshandhaving van het Ministerie van Justitie.

De vragen die men in vragenlijstonderzoek aan uitkeringgerechtigden wil voorleggen, zijn gevoelig van aard. Het kan daarom niet verwacht worden dat respondenten, wanneer deze vragen zonder meer worden voorgelegd, eerlijk zullen antwoorden. Daarom is in vragenlijstonderzoek waarin deze vragen zijn gesteld, gebruik gemaakt van een speciale vraagtechniek voor gevoelige vragen, namelijk Randomized response. Het Ministerie van SoZaWe maakt van deze methode gebruik sinds 1996 (validatieonderzoek), gevolgd door een pilot-onderzoek (1998), de eerste meting van de monitor 'Periodiek onderzoek regelovertreding sociale zekerheid' (POROSZ 1; 2000), de tweede meting, POROSZ 2 (2002) en de voorliggende meting, POROSZ 3 (2004).

De Randomized response methode is speciaal ontwikkeld om in enquêtes informatie te verkrijgen over dergelijke gevoelige onderwerpen. Zij biedt de respondenten een betere waarborg voor anonimiteit van hun antwoorden, waardoor zij minder moeite zullen hebben om het overtreden van een regel toe te geven.

De kern van de methode is dat door de introductie van toevalsfactoren in de vragen of antwoorden de enquêteur noch de onderzoeker weten op welke vragen de respondent naar waarheid heeft geantwoord. Toch kan achteraf door correctie voor de bekende toevalsfactor, worden geschat welke proportie van de respondenten het omstreden gedrag heeft toegegeven. De betrouwbaarheid van de antwoorden wordt met deze procedure aanmerkelijk vergroot.

De Randomized response vraagtechniek kan op verschillende manieren worden ingezet. In dit onderzoek is gebruik gemaakt van de dobbelsteenmethode. De werkwijze is als volgt. De respondent wordt een 'gevoelige' vraag voorgelegd (bv. 'Hebt u zwart gewerkt?'), die hij echter niet direct - ten overstaan van de enquêteur - hoeft te beantwoorden. In plaats daarvan wordt hem gevraagd met twee dobbelstenen te gooien, en wel zodanig dat de enquêteur de uitkomst niet te zien krijgt. De uitkomst van de worp is volstrekt willekeurig. Is de uitkomst van de worp 2, 3 of 4, dan dient de respondent altijd de vraag met 'ja' te beantwoorden. Is de uitkomst 11 of 12, dan dient het antwoord altijd 'nee' te luiden. Bij een uitkomst van 5 tot en met 10 dient de respondent naar waarheid te antwoorden. De verhouding 25% 'verplicht ja of nee' antwoorden en 75% 'naar waarheid' geeft volgens de literatuur over Randomized response aan individuele respondenten een zodanige bescherming dat de meeste van hen, wanneer daarom via de uitkomst van de worp wordt gevraagd, naar waarheid antwoorden.³ Voor alle respondenten gezamenlijk kan later worden geschat welk deel inderdaad zwart gewerkt heeft.⁴

³ Fox, J.A. & Paul E. Tracy, The randomized response approach: applicability to criminal justice, research and evaluation, Evaluation Review, Vol. 4, no. 5, October 1980, 601-622; Moriarty, M. & F. Wiseman, On the choice of a randomization technique with the randomized response model, Proceedings of the Social Statistics Section of the American Statistical Association, 1976: 624-626; Soeken, Karen L. & George B. Macready, Respondents perceived protection when using randomized response, Psychological Bulletin, 1982, Vol. 92, No. 2, 487-489.

⁴ Door een programmeerfout zijn de door het NIPO gebruikte kansen niet identiek aan $P(\text{verplicht ja}) = 1/6$, $P(\text{verplicht nee}) = 1/12$ en $P(\text{eerlijk antwoorden}) = 3/4$, maar $P(\text{verplicht ja}) = .1868$, $P(\text{verplicht nee}) = .0671$ en $P(\text{eerlijk antwoorden}) =$

Uit een meta-analyse⁵ blijkt dat Randomized response meer valide antwoorden oplevert dan andere methoden om valide antwoorden te verkrijgen, vooral indien de onderwerpen gevoeliger van aard worden.

Het gebruik van de Randomized response methode brengt met zich dat bij de analyse van de aldus verzamelde data specifieke technieken en modellen moeten worden toegepast. Het Ministerie van SoZaWe en het Ministerie van Justitie spelen een actieve rol in de verdere ontwikkeling van de bij de Randomized response methode toe te passen technieken en modellen. Zij dragen bij in het onderzoek dat naar dit onderwerp wordt verricht bij de Utrechtse afdeling van de Onderzoeksschool IOPS (Interuniversitaire Onderzoeksschool voor Psychometrie en Sociometrie). Een nieuwe ontwikkeling op het gebied van Randomized response, waar in POROSZ 3 gebruik van is gemaakt, wordt in deze appendix beschreven.

Verdere ontwikkeling van de bij de Randomized response methode toe te passen analysemethoden

De Randomized response methode zoals die in bovenaangehaald onderzoek is gebruikt, kent enkele beperkingen. Wij bespreken er hier drie.

In de eerste plaats levert de methode weliswaar meer valide schattingen op van de mate waarin uitkeringsgerechtigden de regels omtrent de uitkering overtreden dan andere vraagmethoden, maar zij is niet afdoende. Er is naar alle waarschijnlijkheid nog altijd sprake van een onderschatting. Hoe groot die onderschatting is, is tot op heden onbekend. De onderschatting is een gevolg van het feit dat niet alle respondenten de RR procedure zullen volgen, bijvoorbeeld omdat zij de RR procedure niet snappen of omdat zij zich niet veilig genoeg wanen. Om die reden antwoorden zij soms ‘nee’ hoewel de dobbelsteenmethode aangeeft dat zij ‘ja’ (omdat zij 2, 3 of 4 hebben gegooid) zouden dienen te antwoorden. Dit leidt tot de genoemde onderschatting van de regelovertrading.

Bijvoorbeeld: in POROSZ 2 was er een vraag naar het doen van “kleinere klusjes voor geldelijke vergoeding voor of via bekenden zonder dat door te geven aan uitkeringsinstantie”, en voor de WAO werd in 2002 het percentage uitkeringsgerechtigden dat zich hieraan schuldig maakte geschat op 16%. Dit percentage is een onderschatting van het werkelijke percentage regelovertraders indien een deel van de ondervraagde uitkeringsgerechtigden de Randomized response procedure niet volgt door ‘nee’ te zeggen, terwijl zij ‘ja’ zouden moeten zeggen (omdat zij 2, 3 of 4 gooiden, dan wel omdat zij tussen 5 en 10 gooiden en in werkelijkheid de regel hadden overtreden). Het werkelijke overtreedingspercentages zal in dit geval hoger liggen dan 16%. Ten tijde van POROSZ 2 was het niet mogelijk tot een betere schatting te komen voor deze schatting van de prevalentie .

In de tweede plaats was het rekentechnisch nog niet mogelijk om de percentages voor verschillende overtredingen bij elkaar op te tellen om zo te komen tot één algemene maat van regelovertrading. Met andere woorden, indien er drie verschillende vragen naar regelovertrading zijn, welke percentage van de respondenten heeft dan 0 regels overtreden, welk percentage 1 regel, 2 regels en 3 regels?

.7461. De in dit rapport gehanteerde software is zodanig aangepast dat in de uitkomsten rekening gehouden is met deze gewijzigde kansen.

⁵ Een meta-analyse vat al het gepubliceerde onderzoek op een specifiek gebied samen. Hierbij wordt een kwantitatieve aanpak gehanteerd. Meta-analyses hebben tegenwoordig een hogere status dan een kwalitatieve aanpak bij de beschrijving van literatuur omdat zij minder willekeurig is en beter controleerbaar is. De referentie van de meta-analyse waar hier naar wordt gerefereerd is Lensvelt-Mulders, Gerty J.L.M.; Hox, Joop J.; der Heijden, Peter G.M. van; Maas, Cora J.M. (2005). Meta-Analysis of Randomized Response Research : Thirty-Five Years of Validation. Sociological Methods and Research, 33, 319-348.

Voor “gewone” vragen naar regelovertreding is dit te berekenen door een telvariabele te maken die telt hoe vaak iemand “ja” heeft gezegd. Echter, hier zijn de vragen gesteld met Randomized response, waarbij een “ja” of “nee” antwoord niet eenduidig verwijst naar het al dan niet overtreden van een regel omdat een “ja” of “nee” antwoord afdwongen kan zijn door de dobbelsteen. Een voor de hand liggende optelling met grote beleidsrelevantie is het percentage respondenten is dat *geen* van de drie vragen heeft overtreden. Dit percentage wordt verder aangeduid als het ‘nalevingsniveau’ of ‘nalevingspercentage’.

In de derde plaats kan met de Randomized response methode wel een kans op regelovertreding geschat worden, maar dit maakt niet duidelijk welke bedragen er met deze overtreding gemoeid zijn. In de statistische literatuur zijn wel uitbreidingen van Randomized response bekend naar kwantitatieve variabelen, maar deze zijn slechts theoretisch uitgewerkt in de zin dat er in de praktijk weinig tot geen ervaring mee is opgedaan.

Op alle drie punten is sinds POROSZ 2 meer ervaring opgedaan.

- In deze Appendix zal uitgebreid worden ingegaan op de voortgang met betrekking tot het eerste punt, omdat vrijwel alle gegevens in het voorliggende rapport tot stand zijn gekomen met een model dat corrigeert voor het feit dat er respondenten zijn die de Randomized response methode niet volgen.
- Het tweede punt, dat van de naleving, waarbij de score op verschillende fraudehandelingen tot een overallindicator worden samengevoegd, is inmiddels vormgegeven. Daarbij is een model ontwikkeld dat is toegepast op gegevens van POROSZ 2⁶. De resultaten hiervan zijn reeds verwerkt in de Rijksbegroting 2005 SoZaWe.
- Het derde punt, de kwantitatieve uitkomsten, is in POROSZ 3 uitgetoetst. Inmiddels is geconstateerd dat voor een goede verwerking een meer specifieke analysetechniek moet worden ontwikkeld.

Een statistisch model voor een meer valide schatting van de prevalentie van regelovertreding

Recent is een statistisch model ontwikkeld dat poogt de onderschatting van de prevalentie, zoals dat binnen de Randomized response methode voorkomt, te ondervangen. Wij hebben hiertoe samengewerkt met Ulf Bockenholt.⁷ Een gedetailleerde beschrijving van het model is te vinden in twee manuscripten ter publicatie aan tijdschriften zijn aangeboden en momenteel worden beoordeeld, en in een congresbijdrage.⁸

⁶ Rapportage in opdracht van het Ministerie van SoZaWe: P.G.M. van der Heijden, U. Bockenholt en G. van Gils (2005). Een nieuw model voor het meten van naleving op basis van POROSZ 2. Utrecht: BOA.

⁷ Ulf Bockenholt is momenteel als hoogleraar werkzaam aan McGill, Montreal, Canada. Hij is in 2004 verkozen tot President of the Psychometric Society.

⁸ Congresbijdrage: Bockenholt, U. and P.G.M. van der Heijden (2004). Measuring noncompliance in insurance benefit regulations with randomized response methods for multiple items. In: A. Biggeri, E. Dreassi, C. Lagazio and M. Marchi, In (Eds.), 19th International Workshop on Statistical Modelling. Florence, Italy, pp. 106-110. Manuscripten: (1) Bockenholt, U. and van der Heijden, P. G. M. (2004). Randomized--Response Models for Measuring Noncompliance: Risk--Return Perceptions, Social Influences, and Self-Protective Responses. Unpublished manuscript. McGill University. (2) Bockenholt, U. and van der Heijden, P. G. M. (2005). Do randomized--response designs eliminate response biases? An empirical study of non--compliance behavior with social--security--benefit regulations in the Netherlands.

Dit model biedt in principe de mogelijkheid om:

- (i) te schatten welk percentage van de respondenten de RR procedure niet heeft gevolgd;
- (ii) volgend uit (i), voor iedere afzonderlijke regel te komen met meer valide schattingen van regelovertreiding, met andere woorden, schattingen waarin de onderschatting een minder grote rol speelt.

Het model kan worden toegepast indien sprake is van een aantal (minimaal drie) inhoudelijk verwante vragen naar regelovertreiding. Men kan hier binnen de WAO bijvoorbeeld denken aan een aantal vragen naar regelovertreiding op het gebied van arbeid, dan wel regelovertreiding op het gebied van gezondheid.

Het model heeft twee onderdelen⁹:

- a) er wordt een schatting gemaakt van het percentage respondenten dat de RR procedure niet heeft gevolgd. Voor deze groep nemen we aan dat ze systematisch ‘nee’ zeggen, wat ze ook voor vraag wordt voorgelegd en ongeacht de uitkomst van de worp met de dobbelstenen. We noemen deze groep “degenen die niet volgens het Randomized-response-design antwoorden”.
- b) Voor de mensen die de RR procedure wel hebben gevolgd, wordt onderzocht of de vragen waarop het model wordt toegepast hetzelfde zogenaamde ‘latent kenmerk’ meten. Een latent kenmerk is een kenmerk dat je niet kunt zien maar waarvan je aanneemt dat het er is. Hier gaat het om het latente kenmerk ‘regelovertreiding’. Bij drie vragen loopt dat latente kenmerk van 0 tot 3. Het model dat wij hiervoor gebruiken is erg populair in de psychometrie en staat bekend als het Rasch model. Het Rasch model is een van de bouwstenen van het model. Het Rasch model is een probabilistische versie van de zgn. Guttman schaal.

Onderdeel (a) is een simpele aanname over het gedrag van respondenten die de Randomized response procedure niet volgen. Hoe wordt dit percentage geschat? De schatting van het percentage maakt onderdeel uit van de schatting van het gehele model. De schatting van het gehele model vindt plaats met de maximum likelihood methode. Kort gezegd komt hierbinnen de schatting van het percentage van degenen die niet volgens het Randomized-response-design antwoorden als volgt tot stand: bij drie items zijn er $2 \times 2 \times 2$ frequenties met ja/nee antwoorden geobserveerd. Volgens een Rasch model moeten die frequenties een bepaald patroon volgen. Nu vinden we in de data dat de frequentie voor “nee-nee-nee” veel groter is dan volgens het Rasch model mag. Een stuk van die frequentie, namelijk dat deel dat de frequentie te groot is, leidt tot een schatting van het percentage van degenen die niet volgens het Randomized-response-design antwoorden.

Indien er geen respondenten zijn die niet volgens het Randomized-response-design antwoorden, zouden zijn, zou de kans op overtreiding van een bepaalde regel als volgt tot stand komen:

$$P(2,3,4) = 1/6 \text{ en } P(5,6,7,8,9,10) = 3/4,$$

dus

$$P(\text{ja}) = 1/6 + 3/4 * P(\text{regelovertreiding}),$$

⁹ Het model is veel algemener dan hier beschreven. Wij beschrijven hier alleen die delen die in deze rapportage worden gebruikt. Het model is ook gebruikt bij het schatten van de eerder genoemde nalevingspercentages. Enkele andere uitbreidingen: (i) Het is ook mogelijk om meer dan een latent kenmerk te veronderstellen. Dit is bijvoorbeeld zinnig bij de WAO door een latent kenmerk aan te nemen voor de items A1, A2 en A3, en een latent kenmerk voor items D1, D2, D3 en D4. Het eerste latente kenmerk is dan ‘naleving van regels voor het rapporteren van inkomen uit arbeid’ en het tweede is dan ‘naleving van regels voor het rapporteren van de gezondheidstoestand’. Voor elk latent kenmerk kan apart een percentage uit onderdeel a geschat worden.

zodat

$$P(\text{regelovertreiding}) = 4/3 * [P(\text{ja}) - 1/6].$$

Indien deze er er niet zouden zijn, zou een schatting van $P(\text{ja})$ gelijk zijn aan:

$$\text{Schatting } P(\text{ja}) = (\text{aantal ja-antwoorden}) / (\text{totaal aantal respondenten}). \quad (1)$$

Nu er sprake is van mensen die de Randomized response procedure niet hebben gevolgd, dient het totaal aantal respondenten hiermee verminderd te worden. Stel het geschatte percentage van die categorie is 10 %, dan wordt:

$$\text{Schatting } P(\text{ja}) = (\text{aantal ja-antwoorden}) / (90 \% \text{ van totaal aantal respondenten}). \quad (2)$$

Dit laat zien dat de schattingen op regelovertreiding omhoog gaan door rekening te houden met de aanwezigheid van respondenten die niet volgens het Randomized-response-design antwoorden.

Wij maken nog enkele aanvullende opmerkingen om de consequenties van het model duidelijker te maken.

Het model maakt geen onderscheid tussen respondenten die randomized response niet snappen en respondenten die zich niet veilig genoeg wanen. Het model licht een deel van die respondenten er uit die systematisch ‘nee’ hebben gezegd. Dus de respondenten die RR niet snappen zitten alleen in onze kans voor “nee-zeggen” als niet-snappen leidt tot systematisch ‘nee’ zeggen.

Voor wat betreft component b) van het model merken we op dat het zinvol is om, als tussenstap, het Raschmodel te gebruiken. Het Rasch model is heel geschikt om te onderzoeken of vragen hetzelfde latente kenmerk meten. Het Rasch model maakt het vervolgens mogelijk om component a) (het percentage van degenen die niet volgens het Randomized-response-design antwoorden), te schatten indien het aantal vragen minimaal drie is. Als er namelijk drie vragen zijn, dan heb je $2 \times 2 \times 2 = 8$ frequenties, en dit leidt tot 7 onafhankelijke kansen. Geschat worden 3 parameters voor de vragen, 1 zogenaamde discriminatieparameter, en de kans op “nee”-zeggen. Dit model heeft dan $7 - 3 - 1 - 1 = 2$ vrijheidsgraden. Indien er echter slechts twee vragen zijn, dan is het model met zowel component a) als b) niet te schatten: indien er twee vragen zijn, dan zijn er $2 \times 2 =$ vier frequenties, dus 3 onafhankelijke kansen. Je kunt dan maximaal drie parameters schatten. Indien het aantal vragen kleiner is dan drie, dus twee, dan heeft het Rasch model twee parameters voor de twee vragen nodig, en 1 discriminatieparameter. Je hebt dan geen vrijheidsgraden meer om het percentage van degenen die niet volgens het Randomized-response-design antwoorden te schatten. De wijze waarop we hiermee in voorkomende gevallen zijn omgegaan is door dit percentage niet voor die specifieke dataset te schatten, maar een gefixeerde waarde te gebruiken die we vinden binnen de analyse van andere groepen regelovertredingen waar dit percentage wel valt te schatten.

Er zijn twee redenen om vertrouwen te hebben in het voorliggende model.

In de eerste plaats biedt het model een redelijke verklaring voor een tot dusverre moeilijk te verklaren fenomeen dat in POROSZ 2 optrad. In POROSZ 2 werden voor de WAO twee verschillende dataverzamelingmethoden gehanteerd, namelijk een methode interviewde respondenten face-to-face, en andere methode maakte gebruik van een accespanel van het NIPO waar mensen over het internet werden benaderd.

Indien we ongecorrigeerde schattingen maken met (1) gaf voor de vraag naar zwart werk in de face-to-face steekproef 2.7 % van de respondenten toe zwart te werken terwijl dit voor de access panel steekproef 7.3 % was; kortom een verschil van 4.6 %. Indien we de gecorrigeerde schattingen maken met (2) vinden we voor de face-to-face steekproef 9,8 % terwijl we voor de access panel steekproef 12,1 % vinden; kortom een veel kleiner verschil van 2.3 %.¹⁰ Dit komt omdat voor het access panel het aantal mensen dat niet volgens het Randomized-response-design antwoordt, geschat is op 20 % terwijl dit voor de face-to-face steekproef 13 % was. De verklaring die dit model dus biedt, is dat respondenten binnen de face-to-face steekproef vaker de Randomized response procedure hebben gevolgd, en dus vaker eerlijk hebben geantwoord. In de Appendix Naleving is een vergelijkbaar maar nog extremer voorbeeld te vinden dat vertrouwen geeft in het model: het nalevingspercentage voor zowel de face-to-face steekproef als voor de access panel steekproef is exact hetzelfde: 76.5 %.

Een tweede reden om vertrouwen te hebben in het voorliggende model is dat het model goed bij de data past. Dit blijkt uit zgn. chi-kwadraattoetsen, waarmee (bijvoorbeeld voor een set van drie vragen $2 \times 2 \times 2 = 8$) geobserveerde frequenties worden vergeleken met schattingen van verwachte frequenties die het model oplevert. Voor vrijwel alle analyses die wij hebben uitgevoerd is de chikwadraattoets niet significant, wat betekent dat het opstellen van een ingewikkelder model niet veel oplevert. Er zijn ingewikkeldere aannames mogelijk voor het niet volgen van de Randomized response procedure. Dergelijke ingewikkeldere aannames zijn onderzocht in Bockenholt en Van der Heijden (2005; zie noot 6) en het bleek dat de gegevens geen aanleiding geven ingewikkelder veronderstellingen te maken dan de twee simpele aannamen a) en b). Bijvoorbeeld, men zou zich ook kunnen afvragen of er ook mensen zijn die systematisch “ja” zeggen, ongeacht de vraag die aan hun wordt voorgelegd. Uit de geanalyseerde gegevens blijkt echter dat deze groep niet bestaat, althans, niet groot genoeg is om hier in het model een parameter voor op te nemen. Het model is dus probaat genoeg om de voorliggende data goed adequaat te beschrijven.

In het licht van deze laatste opmerking is het zinnig de vraag te bespreken of de voorliggende schattingen, die gebaseerd zijn op de respondenten die het Randomized-response-design *wel* volgen, nu nog steeds onderschattingen kunnen zijn. Er zijn twee redenen die er toe kunnen leiden dat de voorliggende (dus voor degenen die niet volgens het Randomized-response-design antwoorden) gecorrigeerde schattingen nog steeds een onderschatting laten zien.

In de eerste plaats weten we niet wat het percentage regelovertreeders is binnen die categorie. Door schattingen met formule (2) te berekenen nemen we feitelijk aan dat binnen die categorie de kans op regelovertreiding identiek is aan die van de degenen die wel volgens het Randomized-response-design antwoorden. Indien er echter onder de eerstgenoemde categorie een grotere kans is op regelovertreiding, dan zijn de gerapporteerde kansen nog steeds te interpreteren als onderschattingen.

¹⁰ Dit verschil binnen het betrouwbaarheidsinterval rondom de waarde nul, zodat het onverstandig is voor dit verschil een inhoudelijke verklaring naar voren te brengen.

In dit licht zijn resultaten van belang uit Bockenholt en Van der Heijden (2004)¹¹ die laten zien dat voor de WAO in POROSZ 2 er een relatie is tussen het zijn van bedrieger enerzijds en het minder goed begrijpen van de instructies en het hebben van een lage opleiding anderzijds. Uit de voorliggende POROSZ 3 rapportage blijkt echter het tegenovergestelde voor de variabele opleiding: voorzover er relaties zijn tussen regelovertrading en opleiding, overtreden juist mensen met een hogere opleiding vaker een regel. Dit maakt het minder aannemelijk dat de groep die niet volgens het Randomized-response-design antwoordt voor een groter deel bestaat uit mensen die regel overtreden dan de mensen die het Randomized-response-design hebben gevolgd. Wij concluderen daarom voorzichtig dat een interpretatie als onderschatting ten gevolge van een overschot aan regelovertraders onder de degenen die niet volgens het Randomized-response-design antwoorden niet waarschijnlijk is.

In de tweede plaats is het mogelijk dat respondenten *wel* het Randomized-response-design volgen indien zij geforceerd 'ja' dienen te zeggen (d.w.z. bij een worp 2, 3 of 4) maar *niet* het Randomized-response-design volgen indien zij eerlijk dienen te antwoorden (d.w.z. bij een worp 5, 6, 7, 8, 9 of 10) en een regelovertrading zouden dienen toe te geven: in dit laatste geval liegen ze. Een dergelijke overtrading van het Randomized-response-design is in zijn algemeenheid door het model heel moeilijk op te sporen (het kan leiden tot een schending van het Rasch deel van het model), en niet indien dit vooral optreedt bij de regelovertradingen met een relatief lagere prevalentie (dan is het Rasch deel van het model niet geschonden). Wij hebben geen aanwijzingen of dergelijke overtradingen van de Randomized response procedure wel of niet plaatsvinden. Indien zij echter geregeld plaatsvinden, dan leidt dit voor de betreffende regelovertrading tot een onderschatting.

Hoewel er dus theoretisch gezien twee redenen zijn dat de voorliggende schattingen nog steeds een onderschatting te zien kunnen geven, is het aannemelijk, door de redelijkheid van de veronderstellingen, dat wij met dit model een grote stap hebben gemaakt om het realiteitsgehalte van de schattingen te vergroten. Het is ook duidelijk dat nader onderzoek naar wat kenmerken van de categorie die niet volgens het Randomized-response-design antwoordt zijn, gewenst is.

Enkele resultaten voor de onderzochte modellen

De modellen zijn onderzocht op zeven groepen van regelovertradingen:

- WAO: vragen naar regelovertrading m.b.t. werk, in 2000 en 2002 drie vragen, in 2004 twee vragen;
- WAO: vragen naar regelovertrading m.b.t. gezondheid, elk jaar vier vragen
- WW: vragen naar regelovertrading m.b.t. werk, in 2000 en 2002 drie vragen, in 2004 twee vragen;
- WW: vragen naar regelovertrading m.b.t. medewerkingsverplichtingen, elk jaar vijf vragen
- ABW/WWB: vragen naar regelovertrading m.b.t. werk, in 2000 en 2002 drie vragen, in 2004 twee vragen;
- ABW/WWB: vragen naar regelovertrading m.b.t. medewerkingsverplichtingen, elk jaar twee vragen
- ABW/WWB: vragen naar regelovertrading m.b.t. informatieve verplichtingen, elk jaar vier vragen

Voor de formulering van de vragen naar regelovertrading verwijzen wij naar het rapport. Wij rapporteren hier allereerst de geschatte percentage van degenen die niet volgens het Randomized-response-design antwoorden voor ieder van de modellen. Wij memoreren dat dit percentage niet te schatten is indien het aantal regelovertradingen twee is (zie hierboven). Daarnaast heeft de dataverzameling in de WAO in POROSZ 2 op twee manieren plaatsgevonden, namelijk met huisbezoek (face-to-face, dobbelsteenmethode, waarbij de antwoorden op Randomized response vragen door de respondent zelf op de computer werden ingetikt) en via internet (Casi, computer assisted self interviewing, dobbelsteenmethode, bij respondenten die deel uit maken van een herhaaldelijk benaderde groep respondenten, een zgn. access panel). In 2000 hebben alle interviews plaatsgevonden met FtF, maar met de kaartmethode voorgesteld door Kuk.

¹¹ Bockenholt, U. and van der Heijden, P. G. M. (2004). Randomized-Response Models for Measuring Noncompliance: Risk-Return Perceptions, Social Influences, and Self-Protective Responses. Unpublished manuscript. McGill University.

In 2002 hebben alle WW en ABW interviews plaatsgevonden met Casi, de dobbelsteenmethode, evenals alle WAO, ABW en ABW/WWB interviews in 2004. De resultaten voor de schattingen zijn te vinden in Tabel 5.1.

Tabel 5.1: geschatte percentages van degenen die niet volgens het Randomized-response-design antwoorden

WAO	2000	2002	2002	2004
		FtF	Casi	
Werk	.239	.127	.195	---
Gezondheid	.128	.144	.161	.110
WW	2000	2002	2004	
Werk	.153	.095	---	
Medewerk.	.055	.100	*	
ABW/WWB	2000	2002	2004	
Werk	.152	.109	---	
Medewerk.	---	---	---	
Informatie	.167	.095	.074	

---: kans kon niet geschat worden omdat groep vragen slechts twee groot was

* model niet geconvergeerd

Voor het jaar 2000 (POROSZ1) bleek het model relatief instabiel. Dit is het gevolg van het feit dat er in 2000 gebruik is gemaakt van de kaartmethode. Deze kaartmethode geeft de respondenten erg veel veiligheid (in de zin dat zij een relatief geringere kans hebben om eerlijk te moeten antwoorden), met als gevolg dat er minder informatie over regelovertreding in de data zit. Hierdoor zijn de schattingen van het model voor 2000 minder betrouwbaar dan die in 2002 (POROSZ2) en 2004 (POROSZ3). Wij hebben daarom aan de schattingen voor het jaar 2000 minder waarde gehecht dan aan de andere schattingen. In de hoofdtekst van het rapport hebben wij dit tot uitdrukking gebracht door aan te geven dat veranderingen in de percentages tussen 2000 en 2002 mogelijk het gevolg zijn van de methodenwisseling tussen deze jaren (namelijk van de face-to-face kaartmethode naar de casi-dobbelsteenmethode). Daarnaast waren er in 2004 bij de WW voor de medewerkingsverplichtingen convergentieproblemen, waardoor het model niet geschat kon worden.

De schatting van de percentages regelovertreding voor de afzonderlijke vragen is gevoelig voor de schatting van de percentages van degenen die niet volgens het Randomized-response-design antwoorden, omdat bij het verwijderen van die categorie uit de gegevens het relatieve aantal 'ja'-zeggere groter wordt en hierdoor dus de percentages regelovertreding stijgen. Om er voor te zorgen dat de schattingen van de percentages regelovertreding voor de afzonderlijke vragen in 2002 vergelijkbaar zijn aan die in 2004, dient het percentage van degenen die niet volgens het Randomized-response-design antwoorden in 2002 identiek te zijn aan dit percentage in 2004. Dit pleit voor het fixeren van deze percentages op een over 2002 en 2004 identieke waarde. Wat hier ook voor pleit is dat het niet aannemelijk is dat de schattingen van de percentages in tabel 5.1 over de jaren heen snel als feitelijke ontwikkelingen kunnen worden beoordeeld, omdat zij namelijk grote betrouwbaarheidsintervallen kennen (deze zijn niet bepaald voor de schattingen uit Tabel 5.1, maar resultaten uit Bockenholt en Van der Heijden (2005) laten zien dat aldaar een schatting van 13.3 % een betrouwbaarheidsinterval heeft dat loopt van 7.4% tot 19.5 %). Om deze twee redenen hebben wij niet de schattingen uit tabel 5.1 direct gebruikt, maar de schattingen gefixeerd op bepaalde waarden en de modellen opnieuw geschat waarbij het percentage van degenen die niet volgens het Randomized-response-design antwoorden is gefixeerd.

Dit maakt het ook mogelijk het model te schatten in die situaties dat het aantal regelovertradingen per groep twee is. Dit geldt voor de WAO 2004 werk, WW 2004 werk, WWB 2004 werk en ABW/WWB 2000-2004 medewerkingsverplichtingen.

De percentages waarop deze zijn gefixeerd, zijn te vinden in Tabel 5.2. Hierbij zijn als waarden waarop de percentages zijn gefixeerd steeds ronde getallen genomen. We hebben voor het bepalen van de gefixeerde ronde getallen voornamelijk gebruik gemaakt van de schattingen in tabel 5.1 voor 2002 en 2004, en hier de schattingen voor 2000 op afgestemd:

- voor Werk WAO zijn in 2002 twee waarden bepaald omdat de schattingen in tabel 5.1 zoveel uiteenliepen, en de resultaten een groot deel van het verschil voor Zwart werk kon verklaren tussen de FtF 2002 meting en de Casi 2002 meting (zie hierboven); de 2004 meting is een Casi meting en daarom is ook hier 20% genomen;
- voor Gezondheid is het percentage bepaald op 15%, ruwweg het gemiddelde over 2002 en 2004;
- voor de WW zijn alle percentages gefixeerd op 10% (vgl. de schattingen in tabel 5.1 voor 2002), behalve voor de medewerkingsverplichtingen, die op 5% zijn gefixeerd; hier is voor gekozen omdat anders de schattingen voor 2000 nog meer uit de pas zouden lopen met de andere jaren (dit illustreert de eerder genoemde onbetrouwbaarheid van de resultaten voor het jaar 2000);
- voor de ABW zijn de percentages gefixeerd op 10%, ruwweg het gemiddelde voor de drie aanwezige schattingen in 2002 en 2004;
- voor de vraag naar Korting of Boete is voor de WAO een percentage van 15% gehanteerd, en voor de WW en ABW/WWB een percentage van 10%.

Tabel 5.2: gefixeerde percentages van degenen die niet volgens het Randomized-response-design antwoorden

WAO	2000	2002	2002	2002	2004
		FtF	Casi	samen	
Werk	.200	.130	.200		.200
Gezondheid	.150			.150	.150
WW	2000	2002	2004		
Werk	.100	.100	.100		
Medewerk.	.050	.100	.100		
ABW/WWB	2000	2002	2004		
Werk	.100	.100	.100		
Medewerk.	.100	.100	.100		
Informatie	.100	.100	.100		

De schattingen van de modellen op basis van de gefixeerde percentages van deze categorie leveren ook schattingen op van de frequenties van antwoord patronen. Bijvoorbeeld voor WAO Werk 2000 zijn er schattingen voor de frequenties van de antwoordpatronen 000, 001, 010, 011, 100, 101, 110, 111 (0 is 'nee', 1 is 'ja'). Van de schatting van de frequentie voor 000 wordt vervolgens 20 % van de steekproefomvang afgetrokken. Vervolgens worden op basis van de aldus ontstane acht schattingen van frequenties met formule (2) (zie hierboven) de univariate schattingen op regelovertrading berekend, die gevonden zijn in Tabel 1 van het hoofdrapport.

Consequenties voor overige analysetechnieken

Naast aanpassing van de schatting van de univariate percentages op regelovertreding dienen de volgende analyse technieken te worden aangepast:

- betrouwbaarheidsintervallen voor de univariate percentages;
- logistische regressie;
- answer tree.

In deze aanpassingen dient rekening gehouden te worden met het percentage van degenen die niet volgens het Randomized-response-design antwoorden.

Voor de betrouwbaarheidsanalyses is voor elke afzonderlijke regelovertreding een percentage nee-zeggere verwijderd dat gelijk is aan het percentage van die groep. Vervolgens is met een bootstrap procedure het betrouwbaarheidsinterval geschat, rekening houdend met het feit dat de gegevens verzameld zijn met het Randomized-response-design.

Voor de logistische regressie is als volgt met de deze categorie rekening gehouden. In de ideale situatie zouden de degenen die niet volgens het Randomized-response-design antwoorden zelf verwijderd zijn. Dit zou ideaal zijn omdat de samenhang tussen de degenen die niet volgens het Randomized-response-design antwoorden en de afhankelijke variabele kan verschillen van de samenhang tussen degenen die *wel* het Randomized-response-design volgen en de afhankelijke variabele. In dit tijdbestek is het echter niet mogelijk de personen te identificeren die met de grootste mogelijke aannemelijkheid het Randomized-response-design niet volgen. Daarom is er voor gekozen alle ‘nee-zeggere’ een dusdanig lager gewicht te geven dat uit de totale groep ‘nee-zeggere’ een percentage wordt verwijderd dat gelijk is aan het percentage van degenen die het Randomized-response-design niet volgen. Hierdoor zijn de percentages die de logistische regressie oplevert in overeenstemming met de percentages van regelovertreding die het model heeft opgeleverd.

Voor de answer tree speelt dezelfde problematiek als voor de logistische regressie en is een identieke aanpak gekozen als die gebruikt is voor de logistische regressie.

Rest te vermelden dat alle analyses zijn uitgevoerd met gewichten zodat de gegevens zijn gewogen naar bekende populatiekenmerken.

Appendix 6: De T11 in een staafdiagram

Inleiding

De tweede presentatievorm van de T11 is een staafdiagram dat de percentages respondenten weergeeft die worden gekwalificeerd als 'kwetsbaar', 'middengroep' en 'veilig'. De donkere arcering wijst op kwetsbaarheid, de grijze arcering geeft het 'veilige' gebied aan en de lichte arcering het middengebied. Een respondent is 'kwetsbaar' ten aanzien van een dimensie, als hij of zij voor die dimensie van een houding getuigt die eerder met overtreding dan met naleving samengaat.

De WAO

Tussen 2002 en 2004 is vooral het 'middengebied' gegroeid, vooral ten kosten van het 'kwetsbare' gebied, maar ook door kleine verschuivingen vanuit het 'veilige' gebied. WAO-ers zijn dus over het algemeen iets neutraler komen te staan tegenover factoren die een rol kunnen spelen bij naleving en overtreding van de regels.

Binnen de sanctiedimensie lijkt zich nu een nieuwe ontwikkeling voor te doen. Tussen 2000 en 2002 nam hier het 'veilige' gebied af en het midden- en 'kwetsbare' gebied groeide. Tussen 2002 en 2004 is een iets grotere groep WAO-ers een 'veilige' positie gaan innemen ten aanzien van de twee sanctiedimensies T10 (de kans op sanctie) en T11 (de ernst van een sanctie).

De WW

Over de periode 2000-2004 is de meest duidelijke en consistente ontwikkeling voor de WW voor bijna alle dimensies een kleinere of grotere verschuiving naar het middengebied. De WW-ers zijn dus vooral neutraler komen te staan tegenover factoren die naleving en overtreding van de uitkeringsregels kunnen beïnvloeden. Deze ontwikkeling gaat vooral ten koste van het 'veilige' gebied.

Het minder 'veilig' worden van de dimensie 'kans op een sanctie' (T10) die in 2002 was te constateren, lijkt tot staan gebracht. De kans op een sanctie wordt iets hoger ingeschat en de dimensie is daarmee iets veiliger geworden tussen 2002 en 2004.

De ABW/WWB

Bij de ABW/WWB veranderen tussen 2002-2004 de dimensies die vooral een kenmerk zijn van de uitkeringsgerechtigden zelf, zoals de kennis van de regels (T1a), mate waarin de regels van de uitkering duidelijk worden geacht (T1b), de mate waarin men denkt dat naleven moeite kost (T2), de acceptatie van de regels (T3) en de vrees voor verklikken (T6). Deze dimensies veranderen in negatieve zin. Het 'veilige' gebied neemt af, in veel gevallen ten voordele van het middengebied en in een aantal gevallen ook ten voordele van het 'kwetsbare' gebied. Voor dimensies trouw aan de wet (T4), en de vrees om te worden verklikt (T6) of te worden veroordeeld door de omgeving (T5) is er (vooralsnog) vooral sprake van een meer neutrale, mogelijk een meer afwachtende houding. Bij deze dimensies groeit vooral het middengebied.

Ook bij de ABW/WWB lijkt de beduchtheid voor sancties iets toe te nemen. Het 'veilige' gebied voor de dimensies T10 (kans op een sanctie) en T11 (ernst van de sanctie) groeit met enige procentpunten tussen 2000 en 2004.

Diagram 6.1a

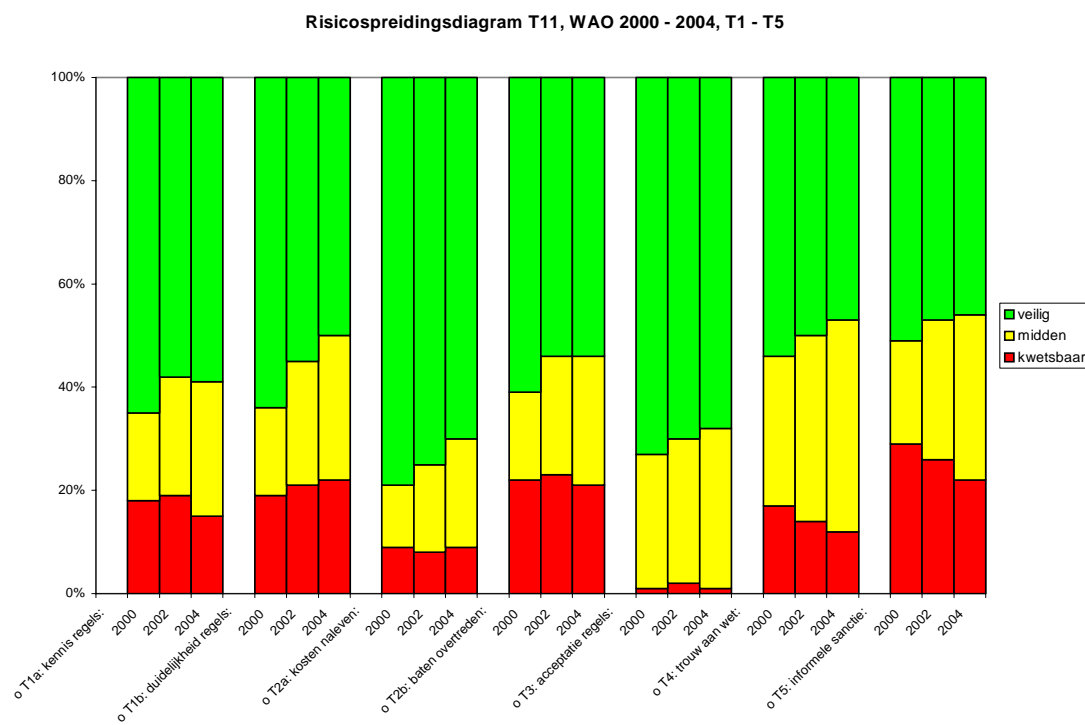


Diagram 6.1b

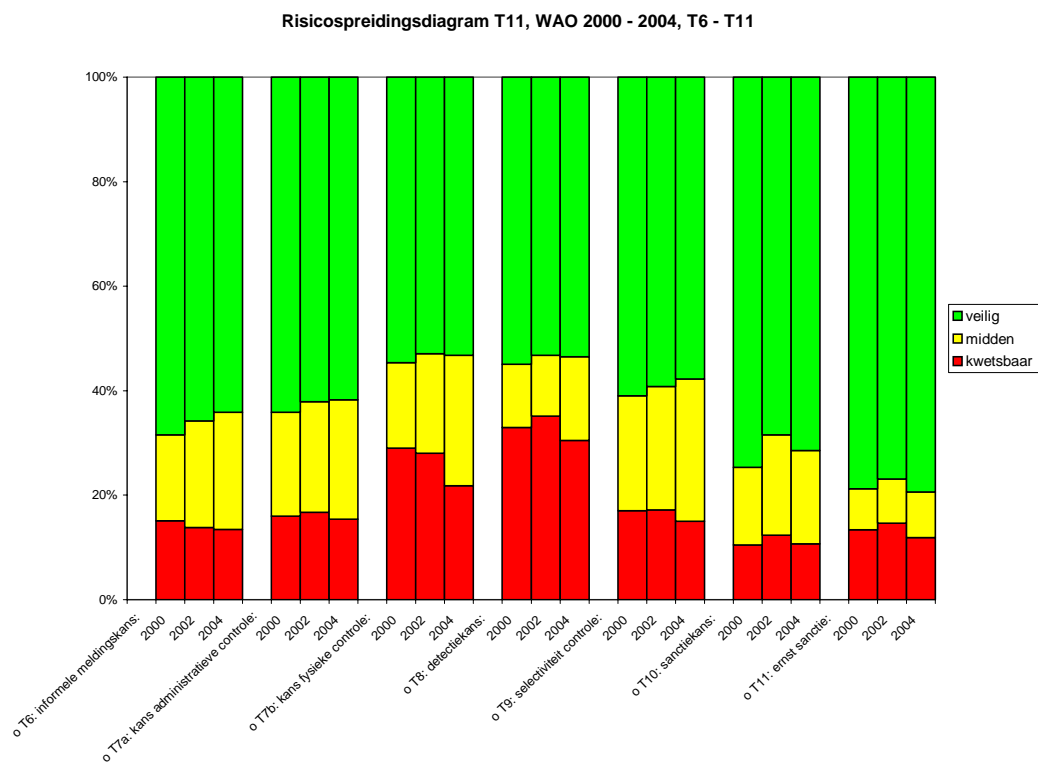


Diagram 6.2a

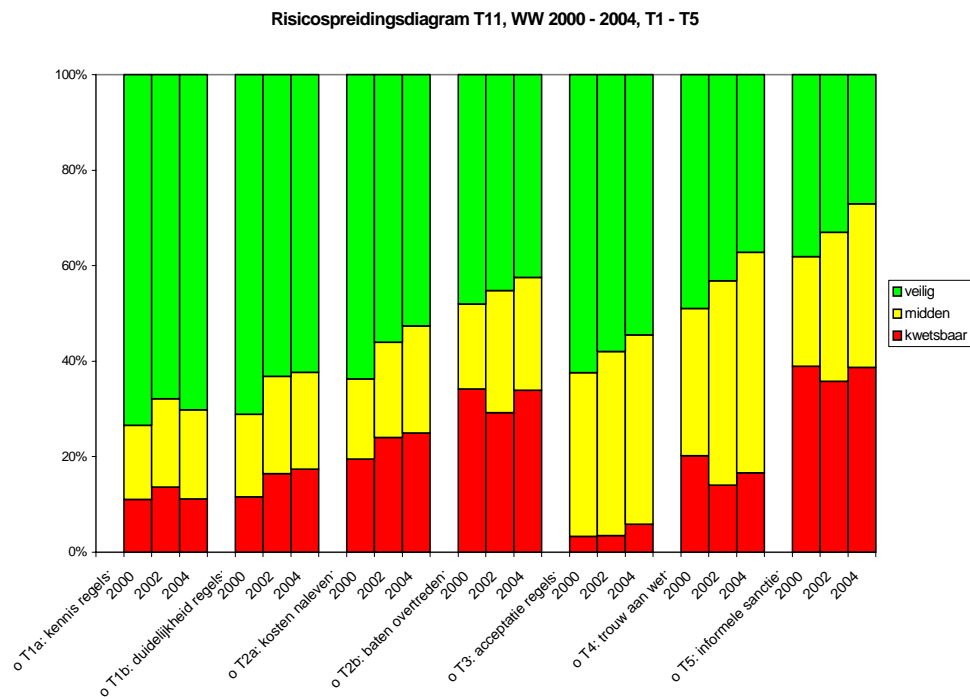


Diagram 6.2b

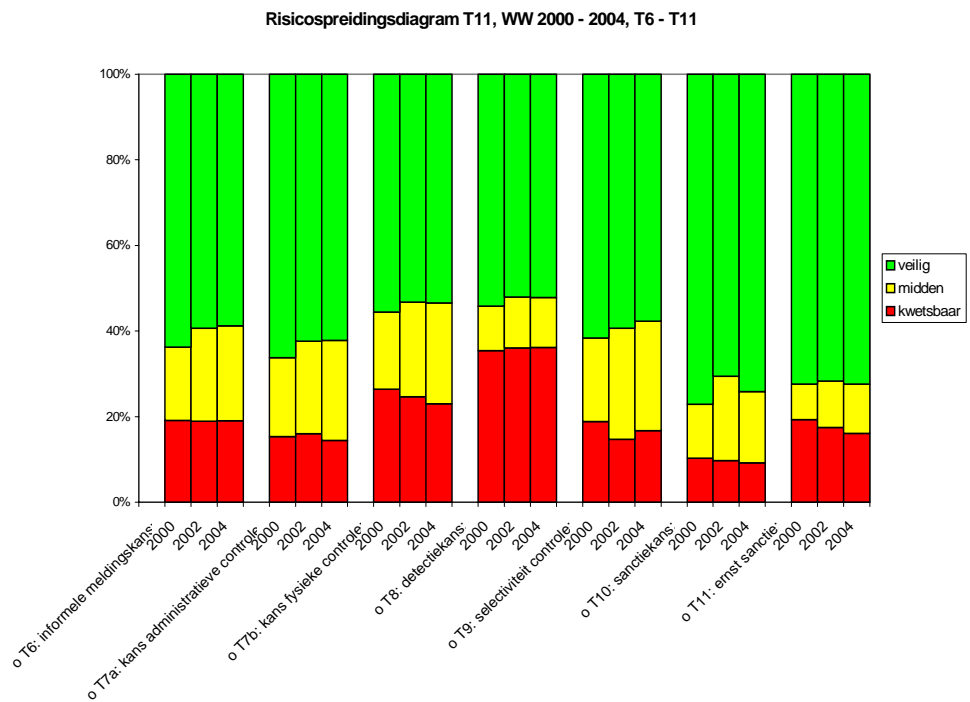


Diagram 6.3a

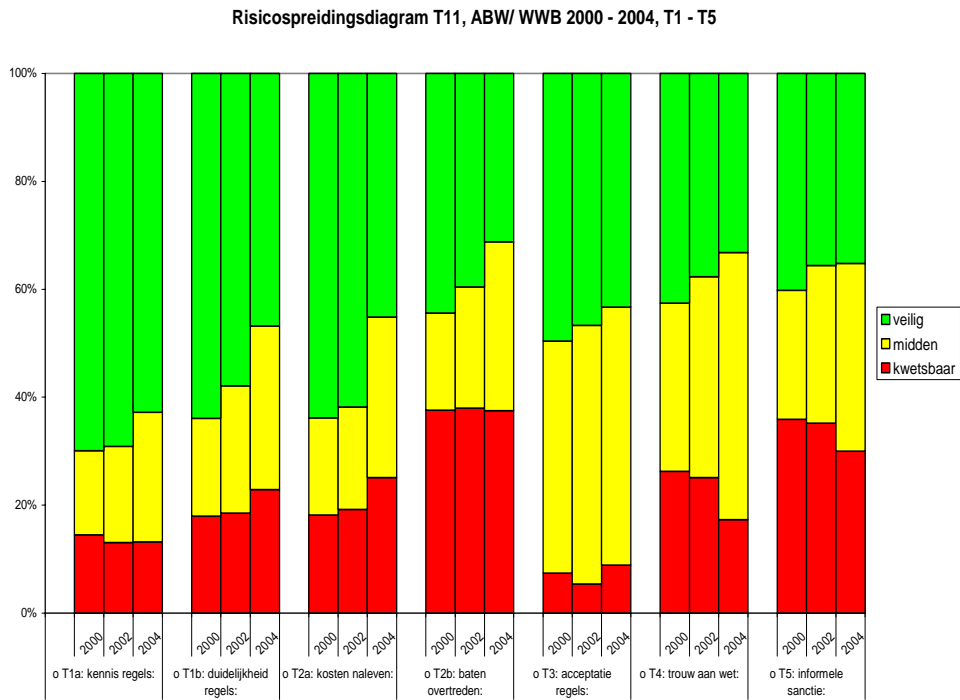


Diagram 6.3b

