

Effect- en impactmeting van de Inspectie Werk en Sociale Economie (IWSE) in Vlaanderen: Analyse en aanbevelingen

Bert van Landeghem
Frank Cörvers

ROA Technical Report

ROA-TR-2013/2

Research Centre for Education and the Labour Market
Maastricht University
P.O. Box 616, 6200 MD Maastricht, The Netherlands
T +31 43 3883647 F +31 43 3884914

secretary-roa-sbe@maastrichtuniversity.nl
www.roa.nl

Effect- en impactmeting van de Inspectie Werk en Sociale Economie (IWSE) in Vlaanderen: Analyse en aanbevelingen

Bert van Landeghem
Frank Cörvers

ROA-TR-2013/2
juni 2013

Research Centre for Education and the Labour Market

Maastricht University

P.O. Box 616, 6200 MD Maastricht, The Netherlands

T +31 43 3883647 F +31 43 3884914

secretary-roa-sbe@maastrichtuniversity.nl

www.roa.nl



Een onderzoek in opdracht van de Vlaamse minister bevoegd voor Werk, in het kader van het VIONA-onderzoeksprogramma.

Inhoud

Voorwoord	iv
Resumé	v
1 Inleiding	1
1.1 Onderzoeksopdracht	1
1.2 Begrippenkader	2
1.3 Hoofdstukindeling	5
2 De Inspectie Werk en Sociale Economie (IWSE)	6
2.1 Situering en structuur	6
2.2 Bevoegdheden en werking	7
2.3 Takenpakket van een inspecteur	8
2.4 Allocatie van mankracht en het stellen van prioriteiten	9
2.5 Aanbevelingen	11
3 Methodologie van de impactmeting: theorie en voorbeelden	12
3.1 Inleiding	12
3.2 Achtergrond	13
3.3 Veel voorkomende terminologie en concepten bij kwantitatieve data-analyse	15
3.4 Types databestanden	20
3.5 Correlatie versus causaliteit of impact	23
3.6 Methoden voor het meten van impact	25
3.7 Voorbeelden uit de literatuur	30
3.8 Aanbevelingen	38
4 Praktijkervaringen in België en Nederland	40
4.1 Inleiding	40
4.2 België: Sociale Inspectie	40
4.3 Nederland: Inspectieraad en Inspectie SZW	41
4.4 Aanbevelingen	46
5 Handreikingen	47
5.1 Inleiding	47
5.2 Administratieve bestanden als basisbouwstenen voor impactstudies	47
5.3 Kwantitatieve rapportagebestanden als basisbouwstenen voor impactstudies	51
5.4 Impact op microniveau: twee concrete voorbeelden met kwantitatieve rapportagebestanden	52
5.5 Impact op microniveau: Samenwerking met derde instanties voor evaluatie en bijsturing risicoanalyse	62

5.6 Impact op microniveau: Samenwerking met derde instanties voor succesvolle implementatie en evaluatie van preventiecampagnes	63
5.7 Impact op macroniveau: het meten van maatschappelijke impact	65
5.8 Aanbevelingen	69
Mogelijke initiatieven door IWSE	70
Literatuur	71
Lijst van afkortingen	74
Glossarium	75

Voorwoord

Het voorliggende onderzoek is uitgevoerd in opdracht van de Vlaamse minister bevoegd voor Werk, in het kader van het zogenaamde VIONA-onderzoeksprogramma. Binnen het model voor strategisch arbeidsmarktonderzoek in Vlaanderen ("Vlaams Programma Strategisch Arbeidsmarktonderzoek") worden jaarlijks middelen ingezet voor studieopdrachten of onderzoeks- en ontwikkelingsopdrachten met het oog op wetenschappelijke ondersteuning van het werkgelegenheidsbeleid.

In de oproep voor de voorliggende studieopdracht werden de voornaamste onderzoeksvragen als volgt geformuleerd:

"1) Meer inzicht te vergaren op de concrete inhoudelijke omschrijving van effect en effectiviteit van het inspectieoptreden;

(2) Relevante kadermethodieken, instrumenten, parameters en/of indicatoren aangereikt te krijgen voor het meten van gewijzigde normconform handelen bij de geïnspecteerde en impact op het beleid van het inspectieoptreden, hetzij generaliserend toepasselijk op alle inspectiematerialen hetzij specifiek per inspectiematerie;

(3) Een stramen en tijdsduur van uitbouw en implementatie van voormelde kadermethodieken, parameters en /of indicatoren."

De grondvraag die IWSE op termijn wenst te beantwoorden is hoe doelgericht haar handhavingsactiviteiten en risicoanalyses zijn en welke resultaten die inspanningen op het beleid en op het normconform handelen door de geïnspecteerde opleveren. Volgens de IWSE bestaat er al een zeer goede risicoanalyse en wenst men nu meer zicht te krijgen op het effect en de impact van het inspectieoptreden. De impactmeting vormt derhalve het belangrijkste en meest interessante deel van deze studieopdracht. De opdracht werd toevertrouwd aan het Researchcentrum voor Onderwijs en Arbeidsmarkt (ROA), Universiteit Maastricht, op basis van het uitgebrachte onderzoeksvoorstel.

De onderzoekers danken de leden van de begeleidingscommissie van het onderzoek voor hun suggesties bij de uitvoering van het onderzoek, en de aanvullingen en het constructieve commentaar op eerdere versies van dit rapport. De onderzoekers willen bovendien alle overige personen bedanken die op één of andere wijze aan het onderzoek hebben meegewerkt. Het betreft de volgende personen uit de begeleidingscommissie en daarbuiten: An Adams (FOD Sociale Zekerheid), Jan Boeykens (IWSE), Anne Bogaerts (Inspectie SZW), Chris Brijs (KSZ), Bjorn Cuyt (UNIZO), Ronald de Bode (Inspectie SZW), Dirk De Rijck (Agentschap Ondernemen), Astrid Depickere (ADSEI), Karel Deridder (RIZIV), Raymond Graindor (IWSE), Gaël Kermarrec (Sociale Inspectie), Evi Martens (IWSE), Paul Michel (RWO), Helena Muyltermans (Kabinet van Philippe Muyters), Thomas Rossie (IWSE), Sonja Teughels (VOKA), Lise-Ange Traufler (SIOD), Koen Van Osselaer (IWSE), Geert Van der Elst (Agentschap Ondernemen), Vincent Vandenameele (Departement WSE), Jan Vansevenant (Privacycommissie), Marc Willems (IWSE).

Resumé

Abstract

Dit rapport is in de eerste plaats een handleiding voor de Vlaamse Inspectie Werk en Sociale Economie (IWSE) teneinde een beeld te geven van de verschillende strategieën om een kwantitatieve impactstudie uit te voeren. Het rapport overloopt enkele statistische begrippen, en het bevat een overzicht van verschillende methodologieën die courant worden gebruikt voor het evalueren van impact. Hierbij wordt aangestipt dat er niet zoiets als een “gouden standaard” bestaat: elke methodologie vereist assumpties, en het is aan beleidsmakers en onderzoekers om een haalbare onderzoeksmethode te kiezen die geloofwaardige resultaten oplevert in een bepaalde context. Het rapport bevat praktijkvoorbeelden vanuit zowel een meer academische als meer beleidsgeoriënteerde invalshoek. Tot slot bevat het rapport enkele voorstellen voor impactstudies geschreven op maat van IWSE.

Doel van het rapport

De Inspectie Werk en Sociale Economie (IWSE) heeft gedurende de voorbije jaren gewerkt aan een kwantitatieve risicoanalyse. Dit model moet haar in staat stellen een prioriteit toe te kennen aan de verschillende materies onder haar bevoegdheid. De inspectie wil echter een stap verder gaan en wil via een kwantitatieve analyse haar impact meten op het normconform handelen van de geïnspecteerden, maar ook op de besluitvorming van regering en parlement, en op de normen en waarden in de maatschappij.

Dit rapport heeft dan ook als doel om IWSE een aantal ideeën aan te reiken omtrent de implementatie van impactstudies. Meer in het bijzonder heeft het rapport tot doel:

- Het definiëren van het jargon dat in de impactevaluatieliteratuur wordt gebruikt;
- Het geven van een overzicht van de verschillende methodologieën die worden gebruikt om impact te meten, en het aanstippen van hun eigenschappen, sterktes en zwaktes;
- Het documenteren van praktijkvoorbeelden;
- Het schetsen van enkele voorstellen tot impactstudies op maat gemaakt voor de Inspectie WSE;
- Op korte termijn dient het rapport een aantal rudimentaire doch potentieel erg interessante impactanalyses mogelijk te maken. Op langere termijn zal het rapport hopelijk kunnen dienen als fundament voor een dialoog met regionale en nationale instellingen, en met onderzoekers gespecialiseerd in impactstudies, teneinde te komen tot het ontwerp en de implementatie van meer complexe en ambitieuze studies.

Definities

Het rapport vangt aan met een bespreking van de centrale concepten in dit rapport: effect, impact, effectiviteit. In Van Dale vinden we de volgende omschrijvingen.

- Effect (eerste betekenis): uitwerking, gevolg.
- Impact: kracht die van iets uitgaat. Synoniemen: invloed, inwerking, draagwijdte.
- Effectiviteit: doeltreffendheid.

Het derde concept is duidelijk verschillend van de eerste twee. De betekenissen voor de begrippen “effect” en “impact” liggen echter veel dichterbij elkaar. De begrippen lijken in de literatuur vaak door elkaar te worden gebruikt, maar impact lijkt toch een iets zwaardere lading te hebben, en lijkt te duiden op een bundeling of reactieketen van de verschillende effecten. In de context van het rapport is vooral het begrip impact gebruikt.

Verder documenteert het rapport enkele basisbegrippen uit de statistiek die geregeld terugkomen in de uiteenzetting. Van bijzonder belang voor het begrijpen van de verschillende methodes om impact te meten zijn de verschillende types van databestanden. De belangrijkste zijn hieronder weergegeven.

- Cross-sectionele data: De data bevatten informatie van een steekproef voor slechts één tijdperiode.
- Paneeldata: Paneeldata bevatten gegevens van een populatie (of steekproef van deze populatie) voor meer dan één periode.
- Cross-sectionele tijdreeksen: Men observeert meerdere subjecten voor een heel aantal periodes. De structuur van cross-sectionele tijdreeksdata is een spiegeling van die van paneeldata. Bij paneeldata observeert men relatief veel subjecten en relatief weinig tijdperiodes, en bij cross-sectionele tijdreeksen is dit omgekeerd.

Twee uitdagingen in impactstudies

Voor inspectiediensten lijken de volgende twee uitdagingen steeds terug te komen wanneer men een impactstudie wil implementeren.

1. Het meten van normconform handelen

Het is niet altijd eenvoudig om het niveau van, of de trends in, normconform handelen te meten. Soms kan men gebruik maken van proxies, dat zijn variabelen die het normconform handelen niet rechtstreeks meten maar die er wel zeer sterk mee gecorreleerd zijn. Bedrijven met een stelselmatig negatief BTW-saldo zijn niet per definitie schuldig aan BTW-fraude, maar de kans is wel groot. Men kan daarom opteren om een indicator voor negatief BTW-saldo te gebruiken om meer inzichten te krijgen in het niveau van alsook de trend in BTW-fraude. Het rapport bespreekt in het kort enkele technieken om via surveys betrouwbare antwoorden te krijgen op gevoelige vragen over nonconform handelen, zoals de Randomized Response Techniek en de Unmatched Count Techniek.

2. Het isoleren van de impact van de inspectiedienst van andere invloeden
Zoals in vele domeinen waarbinnen een impactstudie wordt geïmplementeerd, is het een grote uitdaging om de impact van de inspectiedienst te isoleren van andere factoren die eveneens een invloed hebben op de afhankelijke variabele. In het rapport verwijzen we naar de positieve correlatie tussen de grootte van een politiemacht en criminaliteit, zoals deze vaak wordt vastgesteld in de literatuur. Deze correlatie impliceert echter niet dat een grotere politiemacht leidt tot meer criminaliteit, maar eerder dat beleidsmakers ervoor kiezen om tijdens periodes van onrust of in probleemregio's een grotere politiemacht te ontplooiën. Na het ontwarren van deze verschillende mechanismen komt de literatuur vaak tot de conclusie dat er weliswaar een positieve correlatie bestaat tussen politiemacht en criminaliteit, maar dat de impact van een grotere politiemacht op de criminaliteitsgraad negatief is (criminaliteit neemt dus af).

Methodologieën voor impactmeting

Verschillende methodologieën kunnen worden aangewend voor het meten van impact (eerder dan correlatie), en de context alsook de aannames die men wenselijk vindt zullen bepalen welke methodologie het meest geschikt is. Een aantal courante technieken worden hieronder samengevat.

Controle-variabelen

Bij het meten van een impact kunnen we controle-variabelen toevoegen aan het model, dit zijn variabelen die gecorreleerd zijn met de hoofdverklarende variabele evenals met de verklaarde variabele, om zo tot een meer zuivere impactmeting te komen. Indien men beschikt over paneeldata kan men zelfs technieken gebruiken die controleren voor alle subject-specifieke karakteristieken die niet veranderen over de tijd (zoals geslacht). Men moet er echter rekening mee houden dat in realiteit niet alle factoren waarvoor men zou willen controleren (voldoende accuraat) observeerbaar zijn.

Instrumentele variabelen

Indien men een variabele meent te hebben waarvan men vermoedt dat deze een impact heeft op de hoofdverklarende variabele maar niet op de verklaarde variabele, kan men speciale regressietechnieken gebruiken om via dit instrument de causale impact te schatten van de hoofdverklarende variabele op de verklaarde variabele. Het is echter niet eenvoudig om instrumenten te vinden, en de assumptie dat het instrument een impact heeft op de verklarende variabele maar niet op de verklaarde variabele kan niet direct getest worden.

Gerandomiseerde experimenten

Gerandomiseerde experimenten winnen snel terrein in de sociale wetenschappen teneinde de impact van beleidsmaatregelen op doelvariabelen te meten. Men deelt een populatie "at random" op in een controlegroep en één of meerdere treatmentgroepen, en verschillen in de gemiddeldes van de doelvariabele tussen de verschillende groepen kan ons dan iets leren over de causale impact van een interventie op de doelvariabele. Hoewel een gerandomiseerd experiment een erg

nuttige tool kan zijn, is het vaak om praktische of juridische redenen niet mogelijk om “at random” te interveniëren in de werking van een instelling. Bovendien moet men waken over de validiteit van een experiment: zo mogen subjecten uit de treatment groep niet ex post overschakelen naar de controlegroep en vice versa.

Natuurlijke experimenten

Soms kunnen beleidsveranderingen of (onverwachte) gebeurtenissen de populatie verdelen in twee gelijkaardige groepen, waarbij de ene groep wordt blootgesteld aan een bepaald regime en de andere groep niet. Dergelijke fenomenen worden natuurlijke experimenten genoemd en helpen bij het opzetten van een impactstudie. Opportuniteiten kunnen zich bijvoorbeeld voordoen wanneer een nieuw beleid of programma toegangscriteria kent met scherpe drempelwaardes zoals minimumleeftijd, minimumloon, een exacte startdatum van het programma etc. Subjecten net boven en net onder zo'n drempelwaarde zijn gemiddeld gezien erg gelijkaardig, en gemiddeldes van de doelvariabele voor deze groepen kunnen dan vergeleken worden bij een impactmeting.

Difference-in-difference schatting

Veronderstel een situatie waarbij een groep subjecten een treatment krijgt (bijvoorbeeld een cursus) en een andere groep niet. De cursus start op tijdstip één, en indien we beschikken over paneeldata voor beide groepen met observaties in periode één en twee, kan de impact van de treatment gemeten worden aan de hand van een difference-in-difference schatting. Hoeveel verschilt de groei in de doelvariabele tussen controle- en treatmentgroep? Deze techniek levert echter enkel correcte resultaten op als we kunnen veronderstellen dat de groei in de treatmentgroep even groot zou geweest zijn als in de controlegroep indien de treatmentgroep niet onderworpen was geweest aan de treatment. Als de opdeling in groepen gebeurde via een gerandomiseerd of natuurlijk experiment, kan men met meer overtuiging de validiteit van deze assumptie claimen.

Resultaten

Het rapport stelt de volgende praktische suggesties voor aan IWSE inzake mogelijke impactstudies.

1. Het verder kwantificeren van de inspectieverslagen in kwantitatieve paneeldatasets
In haar jaarrapport documenteert IWSE het totaal aantal inspecties per materie, het aantal inbreuken en zelfs de trend in vastgestelde inbreuken. Deze cijfers laten echter geen impactevaluatie toe omdat een stijging van het aantal gedetecteerde inbreuken niet noodzakelijk duidt op een lage impact, maar ook kan betekenen dat de risicoanalyse verbeterd is. Daarom stellen we voor dat IWSE kwantitatieve rapportagebestanden aanmaakt, met één rij per inspectieoptreden. Dit databestand kan dan gebruikt worden om na te gaan of bedrijven die eerder reeds geïnspecteerd werden meer kans hebben om normconform te handelen bij een tweede en volgende inspectierondes. Zo'n impactstudie is inderdaad vrij rudimentair, maar kan wel op korte termijn geïmplementeerd worden, zonder dat daarvoor de goedkeuring of medewerking

vereist is van andere instanties. Andere vormen van studies kunnen op langere termijn haalbaar zijn als er goede samenwerkingsverbanden tot stand kunnen worden gebracht.

2. Het opzetten van preventiecampagnes, en evaluatie via gerandomiseerde experimenten

Preventiecampagnes zijn vooral nuttig in domeinen waarin de regelgeving niet goed bekend is bij geïnspecteerden, of waarin de dekkingsgraad (ratio van inspecties en mogelijk te inspecteren subjecten) laag is. Men kan “at random” de groep van potentieel geïnspecteerden opdelen in een controle- en treatmentgroep, en de treatmentgroep vervolgens aanschrijven in het kader van een preventiecampagne. Om de impact van de preventiecampagne te testen kan IWSE vervolgens uit de controle- en treatmentgroep een steekproef trekken en nagaan of er een verschil is tussen de twee groepen in normconform handelen. Ook kan men via administratieve data nagaan of beide groepen zich gemiddeld gezien anders gaan gedragen.

3. Enkele suggesties voor IWSE om haar impact te meten op de politieke besluitvorming, of op normen en waarden in de maatschappij, zijn als volgt:

- Het analyseren van (trends in) woordpatronen in de dagbladen, waarbij we denken aan woordgroepen zoals “Vlaamse Arbeidsinspectie” of steekwoorden die geassocieerd zijn met de domeinen waarvoor IWSE bevoegd is. Men kan gebruik maken van de data op de webstek van Mediargus, het digitale platform van de Vlaamse Dagbladpers;
- Een gelijkaardige oefening kan worden gemaakt door het scannen van kamerstukken op bepaalde woordgroepen;
- Om trends in normen en waarden in de maatschappij te meten, of globale trends in normconform handelen, kan men ervoor ijveren om bepaalde vragen op te nemen in enquêtes die op zeer regelmatige basis worden uitgevoerd, zoals de Enquête naar de Arbeidskrachten.

Tot slot willen we nog de navolgende aanbevelingen meegeven teneinde de kans op succesvolle én nuttige impactstudies te vergroten.

- IWSE kan nagenoeg meteen starten met eenvoudige impactstudies door de kwantificering van de inspectieverslagen verder door te zetten.
- Voor meer substantiële impactstudies lijkt een goede samenwerking met andere regionale of nationale instanties onontbeerlijk. Analisten, ICT personeel, software en apparatuur kunnen dan gedeeld worden, waardoor de vaste kosten enorm afnemen. Op de lange termijn zouden beleidsmakers ervoor kunnen kiezen om deze samenwerkingsverbanden institutioneel te verankeren.
- Het kan nuttig zijn om te investeren in goede connecties met de academische wereld. Exploratieve analyses kunnen een gewild onderwerp uitmaken van een bachelor- of masterscriptie. Als de uitkomst van een bepaalde studie van belang is voor een breder (academisch) publiek, kan er zich een gelegenheid voordoen om samen te werken met academische onderzoekers en doctoraatstudenten.
- Men kan bovendien trachten om andere financieringsbronnen dan VIONA aan te boren, indien het onderzoek van belang is voor een breder publiek.

Vele fondsenverstrekkers voor academisch onderzoek beschouwen samenwerking met (of cofinanciering door) een partij buiten de academische wereld vaak als een pluspunt (of zelfs noodzakelijke voorwaarde) bij het toekennen van subsidies.

1 Inleiding

1.1 Onderzoeksopdracht

Een inspectieorgaan dat op nationaal niveau efficiënt wil opereren heeft zonder meer nood aan een goed uitgewerkte systematische aanpak, die voorschrijft hoe de beschikbare middelen binnen het bestaande wettelijke kader worden ingezet. In 2008 is de Inspectie Werk en Sociale Economie (IWSE) dan ook begonnen met de introductie van een kader voor risicoanalyse. Deze risicoanalyse is onder andere gestoeld op de VIONA-studie Pacolet en de Wispelaere (2008) van het HIVA, die een overzicht geeft van het juridisch kader, de beschikbare data alsook verschillende methodes van risicoanalyse. Behalve op deze studie deed het IWSE ook beroep op de expertise van andere organisaties, zoals OVAM, VDAB en Kind en Gezin.

Dit onderzoek is een studieopdracht uitgevoerd binnen het VIONA-onderzoeksprogramma, en heeft als doel zich vooral te richten op de huidige behoefte bij het IWSE naar meer kennis over de impact van het optreden van de inspectie. Het IWSE wil beschikken over tools en methoden om een (continue) evaluatie en bijsturing van de werking van IWSE mogelijk te maken. Men is geïnteresseerd in de impact op microniveau, dat is de impact op het conform handelen van de geïnspecteerde. Ook is men geïnteresseerd in de macro-effecten: hebben inspecties een invloed op sociale normen, en zetten rapportages van de IWSE de overheid ertoe aan om nieuwe maatregelen te treffen, of eerder genomen maatregelen aan te passen?

Op basis van de eerder genoemde wetenschappelijke HIVA-studie (2008) werkt IWSE jaarlijks een risicoanalyse uit ten behoeve van het inspectieoptreden. De nieuwe, voorliggende VIONA-studieopdracht heeft als doel om ter ontsluiting van de inspectieketen, in de toekomst de (beleids)impact en effectiviteit van het optreden van de Inspectie WSE te kunnen meten. Het IWSE wil immers beschikken over tools en methoden om een (continue) evaluatie en bijsturing van de werking van IWSE mogelijk te maken, of zoals het in de projectoproep is geformuleerd:

“De afdeling inspectie wenst nu binnen haar kwaliteitsbenadering haar inspectieketen volledig te ontsluiten door na inspecties, vaststellingen en analyses te komen tot het gebruiken van methoden, technieken, parameters en/of indicatoren voor het meten van:

- Het effect van het inspectieoptreden bij de geïnspecteerde. Met andere woorden het onderzoeken naar het verband tussen naleving van regels en het optreden van inspectie;
- De effectiviteit van het inspectieoptreden. Met andere woorden het onderzoeken in welke mate handhaving naast andere beleidsmatige aansturingen bijdraagt aan een oplossing van achterliggende maatschappelijke problemen of het voorkomen van maatschappelijke risico's. Hier wordt voornamelijk beoogd om impact te creëren op een te wijzigen of een nieuw te creëren beleid.”

In dit hoofdstuk wordt in paragraaf 1.2 ingegaan op het hier te gebruiken begrippenkader. In paragraaf 1.3 volgt een korte beschrijving van de inhoud van de hoofdstukken 2 tot en met 5 van het rapport.

1.2 Begrippenkader

Vaak ontstaat er verwarring tussen verschillende disciplines, instanties, of culturen, omdat de verschillende partijen een bepaald begrip associëren met een andere definitie. De Inspectie Werk en Sociale Economie (IWSE) is zich bewust van dit gevaar, en heeft opgemerkt dat begrippen als effect, impact, effectiviteit door elkaar gebruikt worden. Eén van de vragen van de inspectie op dit vlak luidt of een begrip als “effect” een micro-gevoelswaarde heeft en dus duidt op de draagwijdte voor het gedrag en mentaliteit van de geïnspecteerde, terwijl een begrip als “impact” duidt op de draagwijdte op macroniveau (de overheden).

Impactmeting heeft vele raakvlakken met risicoanalyses. Zo is de ontsluiting van databanken niet enkel noodzakelijk voor risicoanalyses, maar ook voor impactstudies. Ook kunnen de risicoanalyses zelf geëvalueerd worden, onder meer door hun impact na te gaan. Hierdoor kunnen de risicoanalyses worden verbeterd. Er zijn heel wat verschillende definities in omloop voor de begrip impact en risico, wat voor de nodige verwarring kan leiden. Daarom leek het ons nuttig om hier ook aandacht aan te besteden en enkele suggesties te formuleren om de kans op misverstanden te minimaliseren.

In deze paragraaf zullen we aan de hand van definities uit woordenboeken en quote analyses uit de literatuur proberen een beeld te schetsen van de lading of gevoelswaarde van een aantal kernbegrippen in dit rapport. Deze oefening zal dan leiden tot een voorstel van definitie voor elk van deze begrippen, die dan ook zal gehanteerd worden gedurende de rest van het rapport.

Impact en effect

Om te beginnen citeren we hier de betekenis die wordt gegeven in de Dikke Van Dale voor beide begrippen:

Effect (eerste betekenis): uitwerking, gevolg.

Impact: kracht die van iets uitgaat. Synoniemen: invloed, inwerking, draagwijdte.

Uit de Oxford English Dictionary lijkt het verschil tussen “effect” en “impact” iets duidelijker naar voren te komen dan in de Van Dale.

- Effect (first sense): a change which is a result or consequence of an action or other cause.
- Impact (second sense): A marked effect or influence.

De betekenissen uit de woordenboeken illustreren de sterke gelijkheid tussen de twee begrippen, hoewel het woord “impact” toch een zwaardere lading lijkt te hebben, en meer lijkt te wijzen op de finale consequenties voor de lange termijn.

De definities geven ons echter geen antwoord op de vraag van de inspectie of het ene begrip eerder duidt op het individueel niveau, terwijl het andere eerder betrekking heeft op overheden/overkoepelende structuren. Om de intuïtie verder aan te scherpen, onderzoeken we vervolgens enkele titels uit drie vooraanstaande vakbladen, zijnde the American Economic Review, the Journal of Political Economy and de Quarterly Journal of Economics, waarin de termen “impact” en “effect” voorkomen.

Enkele titels met de term “impact” zijn:

- The Impact of Rosenwald Schools on Black Achievement;
- The Impact of Regulations on the Supply and Quality of Care in Child Care Markets;
- The Impact of the Transatlantic Slave Trade on Ethnic Stratification in Africa;
- The Medium-Term Impact of Medicare Part D on Pharmaceutical Prices.

Enkele titels met de term “effect” zijn:

- Clearing the Air? The Effects of Gasoline Content Regulation on Air Quality;
- The Effects of Lottery Prizes on Winners and Their Neighbors: Evidence from the Dutch Postcode Lottery;
- The Effect of Female Education on Fertility and Infant Health: Evidence from School Entry Policies Using Exact Date of Birth;
- The Effect of Medicare Part D on Pharmaceutical Prices and Utilization.

Het lijkt er wederom op dat “effect” en “impact” vrij gelijklopende begrippen zijn, maar dat “impact” toch een zwaarder gewicht, een bredere invulling lijkt te hebben, dat het gaat om een verzameling van effecten. Noteer immers dat het woord “impact” nooit in het meervoud werd gebruikt terwijl het woord “effect” in twee van de voorbeeldtitels in het meervoud staat.

We zullen daarom gedurende de uiteenzetting de term “impact” gebruiken in de betekenis van “invloed”, “draagwijdte”. Het is echter wel aangewezen om telkens een bijvoeglijke bepaling en een voorzetselvoorwerp te definiëren, zodat het duidelijk wordt waarvan men de impact wil meten en waarop.

Effectiviteit

Het begrip effectiviteit scheidt wat minder verwarring, en wordt in de Van Dale omschreven als “doeltreffendheid”. Wanneer men een “impact” identificeert houdt dit niet noodzakelijk effectiviteit in. Stel dat men een beleid invoert met als doel de criminaliteit te doen dalen. Echter, het beleid veroorzaakt het tegenovergestelde. Niettegenstaande er een impact is van het beleid op criminaliteit, is er hier allerminst sprake van effectiviteit. In de Oxford English Dictionary is deze gedachtegang terug te vinden in de definitie van “effectivity”.

- Effectivity = effectiveness: the degree to which something is successful in producing a desired result; success.

Risico

Van de vier hier besproken begrippen in dit hoofdstuk, is het begrip “risico” wellicht hetgeen waarmee de meeste uiteenlopende definities worden geassocieerd, met heel wat onduidelijkheid tot gevolg.

In het woordenboek Van Dale wordt risico omschreven als “gevaar voor onheil, schade of verlies”. In de Oxford English Dictionary vinden we voor het begrip “risk” een gelijkaardige betekenis: “(Exposure to) the possibility of loss, injury, or other adverse or unwelcome circumstance; a chance or situation involving such a possibility”.

Deze betekenissen illustreren meteen de complexiteit van het begrip. Een risicofactor is immers een aggregatie van meerdere dimensies, namelijk kans op gevaar en de impact wanneer het onheil of de ongewenste gebeurtenis zich voltrekt. De woordenboekdefinities zijn hier vrij vaag over: in meer formele definities, bijvoorbeeld in handboeken wiskundige micro-economie, zijn de verschillende dimensies veel beter omlijnd.

Niet alleen “risico” is een vrij complex begrip, maar ook de verschillende dimensies (de kans op het voltrekken van onheil enerzijds, en de negatieve impact wanneer de gebeurtenis zich voltrekt anderzijds) zijn moeilijk te meten en te bevatten materies. Het hoeft dan ook niet te verwonderen dat binnen en tussen culturen mensen er een wijde variatie aan invullingen van het begrip op nahouden, wat het moeilijk maakt om de (risico)bereidheid van mensen te meten en te vergelijken (Weber en Hsee, 1998).

De Vlaamse inspectie WSE ziet een risicogehalte duidelijk als een aggregatie van meerdere dimensies. Zo krijgt het domein van projecten binnen het Europees Sociaal Fonds (ESF) een risico-score “extreem ernstig”, niet zozeer omdat er voortdurend verkeerd wordt omgesprongen met ESF-subsidies, maar omdat de gevolgen bij fouten enorm groot zijn.

Bij de Nederlandse Inspectie van het ministerie van Sociale Zaken en Werkgelegenheid (SZW) interpreteert men risico op een gelijkaardige manier, en definieert men risicoanalyse als: “een methode waarbij risico's worden gekwantificeerd door het bepalen van de kans dat een dreiging zich voordoet in combinatie met de gevolgen ervan: risico = kans x effect. “

Het lijkt er dus op dat de Inspectie WSE de term risicoanalyse op een vrij algemeen aanvaarde manier gebruikt. Desalniettemin, gegeven de verschillende connotaties die het woord bij mensen oproept en gegeven het feit dat men vaak slechts vluchtig een samenvatting van een rapport leest, is het misschien beter om materies met een zeer hoog risicogehalte niet te categoriseren als “zeer ernstig”. Dit kan bij anderen het idee opwekken dat er heel veel fout gaat in een bepaald domein. De uitkomsten van een risicoanalyse weergegeven in een prioriteitentabel, met labels als “zeer hoge prioriteit”, “hoge prioriteit”, etc. zou de kans op misvattingen kunnen verkleinen.

1.3 Hoofdstukindeling

In hoofdstuk 2 van dit rapport wordt ingegaan op de werkwijze van de IWSE. Er zal een beknopte beschrijving worden gegeven van haar huidige werkwijze. Hiervoor wordt gebruik gemaakt van werknota's van het IWSE (zoals Rossie, 2012), maar ook van informatie verkregen via gesprekken met haar medewerkers.

In hoofdstuk 3 zal vervolgens uitvoerig worden ingegaan op verschillende strategieën (kadermethodieken) die kunnen aangenomen worden om de werking van de inspectiedienst te evalueren. Voor- en nadelen van methodologieën worden in hun algemeenheid besproken, alsook de contextuele elementen die bepalen welke methodologie het meest geschikt is. In dit deel van de studie zal gerefereerd worden aan bestaande rapporten en handboeken die handelen over optimalisatie en evaluatie van inspectieorganen, maar ook zal informatie verwerkt worden uit andere bronnen (handboeken, overzichtsartikelen) die de methodologische aspecten van impactevaluatie meer in het algemeen behandelen.

Het theoretisch kader van hoofdstuk 3 is belangrijk omdat het duidelijk maakt welke aannames nodig zijn bij het gebruik van een bepaalde methodologie, en omdat het duiding geeft welke methodologie het meest geschikt is in welke omstandigheden. Zo'n kader is echter vrij abstract zonder illustraties uit de praktijk. Bovendien vergt een goede evaluatie niet enkel een grondige kennis van de beschikbare methodologieën en technieken, maar ook de nodige creativiteit, onderzoekservaring en probleemoplossend vermogen.

Vervolgens komen in hoofdstuk 4 van dit rapport de praktijkervaringen in België, en Nederland aan de orde. Op deze wijze wordt getracht een brug te slaan tussen de meer academische benadering in hoofdstuk 3 en de manier waarop er in de wereld van politiek en beleid wordt omgegaan met impactstudies.

In hoofdstuk 5 worden op basis van de voorgaande hoofdstukken handreikingen gedaan naar de IWSE voor de implementatie van impactstudies. Hierbij worden tevens meerdere voorbeelden uit de literatuur gegeven.

Elk hoofdstuk sluit af met een paragraaf van aanbevelingen voor de IWSE. Aan het einde van het rapport worden veel van deze aanbevelingen samengevat in een tabel met mogelijke initiatieven voor de IWSE op korte, half lange en lange termijn. Voorts zijn een literatuurlijst, een lijst met afkortingen, en een glossarium met begrippen opgenomen.

2 De Inspectie Werk en Sociale Economie (IWSE)

2.1 Situering en structuur

“Werk en Sociale Economie” is één van de dertien Vlaamse Beleidsdomeinen. Binnen het beleidsdomein Werk en Sociale Economie zijn er het Departement Werk en Sociale Economie en de beleidsuitvoerende agentschappen: Vlaamse Dienst voor Arbeidsbemiddeling en Beroepsopleiding (VDAB), het Vlaams Subsidieagentschap voor Werk en Sociale Economie, het Agentschap voor Ondernemersvorming Syntra Vlaanderen en het ESF-Agentschap Vlaanderen. Net zoals in de andere beleidsdomeinen staat het Departement Werk en Sociale Economie in voor de kerntaken beleidsontwikkeling, -opvolging en -evaluatie enerzijds en anderzijds de beheers- en beleidscoördinatie. Bijkomend zorgt het Departement Werk en Sociale Economie tevens voor het handhavings- en sanctioneringsbeleid en het inspectieoptreden. Binnen het Departement Werk en Sociale Economie zijn er naast de staffuncties een afdeling Werkgelegenheidsbeleid en een afdeling Inspectie.

Het Subsidieagentschap voor Werk en Sociale Economie verleent subsidies ter ondersteuning en versterking van de werkgelegenheid in de verschillende sectoren van de economie in Vlaanderen. De VDAB biedt een uitgebreid dienstenpakket aan op het vlak van arbeidsbemiddeling, competentieontwikkeling en loopbaandienstverlening. Als regisseur faciliteert zij tevens de arbeidsmarktwerking. Het Vlaams Agentschap voor Ondernemersvorming Syntra Vlaanderen staat in voor het aansturen en richting geven van een arbeidsmarktgericht en innovatief opleidingsaanbod voor ondernemende mensen. Zij stimuleert, ontwikkelt, erkent, financiert en houdt toezicht op allerhande initiatieven daartoe.

Het ESF-Agentschap Vlaanderen waarborgt vanuit haar beheersopdracht van het ESF-programma het samenwerkings- en partnerschapsdenken door een doorgedreven gesystematiseerde samenwerking met alle actoren in het werkgelegenheidsbeleid. De ESF-werking heeft de opdracht het Vlaamse werkgelegenheidsbeleid te ondersteunen en aanvullend subsidiërend te versterken.

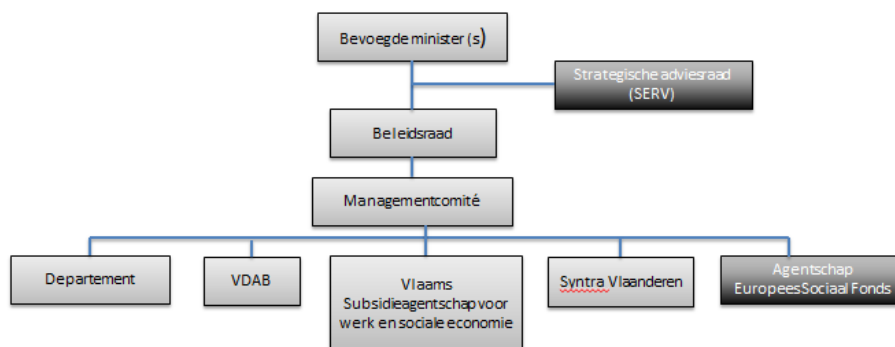
De IWSE is onderdeel van het kerndepartement, en is verantwoordelijk voor de handhaving van het werkgelegenheidsbeleid en het beleid sociale economie. Om deze kerntaak te vervullen wordt de Vlaamse Sociale Inspectie WSE (IWSE) geacht ondernemingen en in slechts enkele uitzonderingen particulieren -- die werken binnen deze ondernemingen -- te inspecteren.

De IWSE is provinciaal georganiseerd. Elke provincie heeft een team van inspecteurs. Er bestaat ook overleg tussen de verschillende teams om ervoor te zorgen dat de werkwijze uniform is. Elk provinciaal team bestaat uit een 4 tot 7 FTE. In totaal heeft de IWSE ongeveer 35 inspecteurs in dienst die ca. 26 tot 29 FTE vertegenwoordigen. Via overleg, intervisie, afsprakenkaders, het geüniformiseerd elektronisch inspectie en registratiesysteem, geüniformiseerde modeldocumenten en checklists, kwaliteitsmanagement en organisatiebeheer wordt er gestreefd naar een

gemeenschappelijke missie en visietoepassing en een maximale uniforme werkwijze tussen provincies.

Figuur 2.1

Positie van IWSE als onderdeel van het Departement Werk en Sociale Economie, ten opzichte van andere beleidsuitvoerende agentschappen



Bron: Departement Werk en Sociale Economie (2011)

2.2 Bevoegdheden en werking

De IWSE oefent toezicht uit op de uitvoering van het beleid via onafhankelijke controles op het terrein van WSE, ondersteund door een onderbouwde risicoanalyse. Aan de vaststellingen die de Inspectie Werk en Sociale Economie doet, kunnen – afhankelijk van de aard van de vaststelling of de inbreuk – verschillende gevolgen gegeven worden. Het inspectieoptreden behelst zowel een informatief, preventief als repressief optreden.

IWSE is bevoegd voor de handhaving van de volgende materies:

- Tewerkstelling van werknemers met een vreemde nationaliteit;
- Private arbeidsbemiddeling;
- Sociale Interventie;
- Evenredige Arbeidsdeelname;
- Diversiteit;
- ESF-dossiers;
- Sociale Economie;
- Tewerkstellingsprogramma's, Activeringsmaatregelen, Competentie- en Loopbaanbeleid;
- Taaldecreet.

IWSE is de kleine gewestelijke zusterorganisatie van de volgende vier federale sociale inspectiediensten:

- Rijksdienst voor Arbeidsvoorziening (RVA);
- Rijksdienst voor Sociale Zekerheid (RSZ);
- Toezicht Sociale Wetten (TSW);
- Sociale Inspectie.

Sommige bevoegdheden sluiten nauw bij elkaar aan en de inspectie op de tewerkstelling van buitenlandse werknemers is een gedeelde bevoegdheid. Het

IWSE werkt dan ook samen met deze diensten, alsook met andere gewestelijke inspecties (het Waalse en het Brusselse) en indien nodig met andere inspectiediensten buiten het domein Werk en Sociale Economie. Wanneer IWSE inbreuken vaststelt waarvoor het niet bevoegd is, kan IWSE het bevoegde inspectieorgaan hiervan inlichten. Eveneens is het mogelijk dat verschillende inspectieorganen gezamenlijk een onderneming gaan inspecteren.

De jaarverslagen geven een beeld van de werking en prestaties van het IWSE, en zijn op de website raadpleegbaar. Het gaat daarbij om de volgende informatie:

- de bezettingsgraad van de afdeling Inspectie Werk en Sociale Economie;
- het aantal uitgevoerde inspecties, met vermelding van het aantal bezochte werkgevers en opleidingscentra en de hierbij gecontroleerde werknemers of cursisten;
- het aantal opgestelde processen-verbaal en het aantal voorgestelde en getroffen maatregelen;
- de ingevorderde administratieve geldboeten;
- de nog te innen administratieve geldboeten;
- acties en initiatieven ter verbetering van de kwaliteit en het beheer van het uniform optreden en rapporteren en de organisatie;
- een rapportering en evaluatie, inclusief een voorstel met betrekking tot de wijze van optreden, de bevoegdheden en de samenwerking en uitwisseling van informatie met andere instanties.

Bij afgeronde inspectierondes over specifieke materies worden er bij markante vaststellingen globale nota's opgemaakt ter ondersteuning van de beleidsevaluatie en -aansturing.

In het laatste jaarverslag van de IWSE (Departement WSE, 2013) wordt aangekondigd dat er de komende jaren wordt gewerkt aan een meer structureel preventief optreden. Bij inspecties is men aanwezig op het terrein en geeft men informatie. Er worden eveneens waarschuwingen uitgeschreven bij inspecties, die eveneens onderdeel uitmaken van het preventieve optreden. In 2012 werden bovendien 'quick scans' uitgevoerd met betrekking tot arbeidsgerelateerde discriminatie. Daarnaast worden er veel informatievragen verwerkt via telefoon of mail. Al deze acties hebben als doel (nieuwe) onregelmatigheden te vermijden.

2.3 Takenpakket van een inspecteur

Een inspecteur heeft doorgaans expertise in enkele materies. Immers, elke materie waarvoor de Inspectie WSE bevoegd is, kent zijn eigen complexiteit. Toen de inspectiedienst nog veel kleiner was werd een brede inzetbaarheid verwacht, en deed elke inspecteur inspecties op alle terreinen. Heden ten dage echter, nu het team groter is, is het gemakkelijker om mensen zich te laten specialiseren. Toch wordt nog steeds verwacht dat een inspecteur ook enig inzicht heeft in de andere materies die het IWSE onder zijn bevoegdheid heeft. Immers, indien een inspecteur langdurig afwezig is, moet deze kunnen vervangen worden, en ook verschuivingen van prioriteiten over de tijd kunnen ertoe leiden dat men buiten zijn kernmateries moet opereren.

Een inspectie behelst niet enkel actie op het terrein. Eerst is er een voorbereidingsfase, waarbij de inspecteur de nodige bronnen gaat raadplegen. Dan is er de inspectie ter plaatse. Bij de inspecties hebben de inspecteurs bepaalde politionele bevoegdheden, volgens het zogenaamde Decreet van 30 april 2004¹ en het sociaal strafwetboek van 6 juni 2010. Een belangrijke bevoegdheid is de bijzondere bewijskracht van een inspecteur. Zo ook de bevoegdheid om zich dag en nacht op een werkplaats te mogen begeven (de privévertrekken enkel na het bekomen van een visitatiebevel), en mogen inspecteurs de identiteit van alle personen opvragen (en alle andere wettelijke documenten), en iedereen ter plekke bevragen indien dit kadert in de uitvoering van hun onderzoek. Ook mag de IWSE alle documenten bij andere overheidsdiensten opvragen. Dit laatst geldt niet voor gerechtelijke documenten, daarvoor moet eerst toestemming gevraagd worden aan de onderzoeksrechter.

Tot slot is er het schrijven van een rapport. De verslagsjablonen worden sinds kort automatisch gegenereerd na ingave van diverse data. Deze hebben als voordeel dat rapporten (tot op zekere hoogte) een geüniformiseerde structuur hebben, en dat de ingave van data in de applicatie het mogelijk maakt om statistieken te genereren voor het jaarrapport.

Elk provinciaal team van inspecteurs heeft een coördinator. Deze coördineert het werk van de inspecteurs, en gaat ook zelf de baan op om voeling te houden met wat er op het terrein gebeurt. Ongeveer 15% van de bestede tijd van de inspecteur wordt ingezet aan taken die niet direct verband houden met het behandelen van concrete dossiers, zoals het schrijven van een bijdrage voor het jaarrapport. Een coördinator wordt geacht 50% van zijn tijd te besteden aan inspecties.

2.4 Allocatie van mankracht en het stellen van prioriteiten

Zoals beschreven in risicoanalyse 2012 (Rossie, 2012) worden bij de start van ieder jaar de prioriteiten bepaald op materieniveau, en dit aan de hand van – wat men noemt- een risicoanalyse. Uitdrukkelijk kan hier worden vermeld dat de risicoanalyse nog niet plaats vindt op werkgeversniveau, iets dat voor de toekomst wellicht zinvol is.

De verdeling van de beschikbare mensen en middelen over de verschillende materies wordt formeel vastgelegd in het jaaractieplan. Hierin wordt aangegeven hoeveel inspecties er moeten plaatsvinden per materie, en hoeveel tijd er voor een inspectie binnen een bepaalde materie mag uitgetrokken worden. Hoe de inspecties binnen een materie worden verdeeld, bijv. welke bedrijfstakken in welke regio's het meest dienen gecontroleerd te worden, wordt bepaald door middel van overleg in de thematische werkgroepen en over de verschillende provinciale afdelingen heen. Deel doelstellingen worden voorgelegd en gevalideerd door de bevoegde ministers.

Anno 2012 is de risicoanalyse eerder descriptief en op materieclusterniveau. Vanaf 2013 zal deze gebeuren op materieniveau onderbouwd door een bevraging van

¹ Decreet van 30/04/2004 tot uniformisering van de toezichts-, sanctie- en strafbepalingen die zijn opgenomen in de regelgeving van de sociaalrechtelijke aangelegenheden, waarvoor de Vlaamse Gemeenschap en het Vlaamse Gewest bevoegd zijn.

institutionele stakeholders. De risicoanalyse helpt om de materies te ordenen op basis van hun risicogehalte in de volgende categorieën:

- Extreem ernstig
- Zeer ernstig
- Ernstig
- Belangrijk
- Neutraal

Zoals in hoofdstuk 1 al is vermeld, is het wellicht beter om materies met een zeer hoog risicogehalte niet te categoriseren als “zeer ernstig”. Zo krijgt domein ESF-projecten een risico-score “extreem ernstig”, niet zozeer omdat er voortdurend verkeerd wordt omgesprongen met ESF-subsidies, maar omdat de gevolgen bij fouten enorm groot zijn. Het gaat hier dus om het gebruik van de term risico in de formule: risico = kans x effect. De uitkomsten van een risicoanalyse weergegeven in een prioriteitentabel, met labels als “zeer hoge prioriteit”, “hoge prioriteit”, etc. zou de kans op misvattingen kunnen verkleinen.

Vier clusters van criteria worden meegenomen in de (toekomstige) risicoanalyse:

Cluster 1: Impact

De dekkingsgraad (het quotiënt van aantal uitgevoerde inspecties en mogelijk te inspecteren subjecten) alsook het percentage inspecties met vaststellingen in het afgelopen kalenderjaar wegen mee in het bepalen van de prioriteit van een materie in het toekomstig jaar. De weging van de dekkingsgraad is daarenboven een functie van het percentage inspecties met vaststellingen. Daarnaast zijn ook de financiële massa (hoeveelheid subsidies per materie) alsook de financiële impact (totale som binnen een materie over het voorbije jaar van bedragen die IWSE voorstelde te korten of terug te vorderen) factoren binnen dit cluster.

Het uiteindelijke doel van de inspectie is een gedragswijziging van de geïnspecteerden te bereiken in de richting van het conform handelen. Een interessante vraag in dit kader is het nut van preventief optreden. Zou het effectief kunnen zijn om inspecties aan te kondigen? Dit zou inbreuken kunnen voorkomen.

Cluster 2: Wetgeving

De prioriteit die een materie krijgt hangt tevens af van de bruikbaarheid van het wettelijk kader voor inspectie.

Wetgeving speelt ook een rol bij het dossierniveau van de geïnspecteerden. De impact van het inspectieoptreden wordt grotendeels bepaald door de tools die de wetgeving biedt. De bevoegdheden van de inspectie hebben betrekking op alle documenten. Bij elektronische opgeslagen gegevens (in microdatabestanden, dat wil zeggen in batchformat) speelt de privacy van te inspecteren personen een belangrijke rol. De IWSE kan niet zomaar gegevens in batchvorm opvragen, of op regelmatige, systematische wijze alle gegevens laten bezorgen. Door analyse van databestanden zouden echter uitgebreidere voorbereidingen, meer efficiëntie en eenvormigheid, betere kwaliteit van inspecties, en minder overhead mogelijk zijn. Ook het koppelen van databestanden voor risicoanalyses zou heel zinvol zijn. Betere ontsluiting van databestanden is gewenst om de voorbereidingen te verbeteren en de

efficiëntie van de inspectie verder te vergroten. Iets soortgelijks geldt voor een uitbreiding van de juridische ondersteuning van de inspecteurs die 'op de baan' zijn.

Cluster 3: Relatieve aspect

Is de inspectie bekend met het te inspecteren object, en is het te inspecteren object bekend met de inspectie? Meer bekendheid van het object met de inspectie kan het risico op inbreuken verkleinen, terwijl de inspectie de risico's beter kan inschatten als er meer kennis is van het object.

Cluster 4: Gevoeligheid

Materies die gevoelig liggen bij de publieke opinie krijgen een hogere weging. Er is sprake van een hoger risico door de maatschappelijke impact die een inbreuk kan hebben (volgens de formule $\text{risico} = \text{kans} \times \text{effect}$).

2.5 Aanbevelingen

- De uitkomsten van een risicoanalyse weergegeven in een prioriteitentabel, met labels als "zeer hoge prioriteit", "hoge prioriteit", etc. in plaats van "zeer ernstig" en "ernstig", zou de kans op misvattingen kunnen verkleinen.
- Om een gedragswijziging van de geïnspecteerden in de richting van het conform handelen te bereiken, kunnen er ook preventieve inspecties worden uitgevoerd, of kunnen inspecties worden aangekondigd. Hiermee kunnen inbreuken worden voorkomen, en kunnen de bekendheid en het imago van de IWSE worden vergroot.
- De effectiviteit en efficiëntie van de IWSE kunnen verder worden verbeterd door onder andere de ontsluiting van databestanden met gegevens over de geïnspecteerden, het koppelen met databestanden van andere inspecties en instellingen, en de uitbreiding van de juridische ondersteuning van de inspecteurs die 'op de baan' zijn.

3 Methodologie van de impactmeting: theorie en voorbeelden

3.1 Inleiding

Een kwantitatieve, zuivere impactanalyse is bijzonder nuttig en zelfs onontbeerlijk, niet enkel voor inspectiediensten maar in vele andere domeinen van de maatschappij. Overheden (en belastingbetalende burgers) willen weten of enorme bedragen besteed aan onderwijs, tewerkstellingsprogramma's, of ontwikkelingshulp, de verwachte impact hebben en dus een gewenste bijdrage leveren aan het bereiken van de doelstellingen. Impactstudies die vaak de media halen hebben betrekking op de farmaceutische sector. Vooraleer medicijnen op de markt komen moet immers aangetoond worden dat de medicijnen effectief zijn, dus bijdragen tot de beoogde doelstelling, en dat zij geen ongewenste impact (bijwerkingen) met zich meebrengen.

Gegeven de behoefte voor kwantitatieve impactstudies op talrijke domeinen, dient het niet te verwonderen dat gedurende de laatste decennia er heel wat aandacht besteed is aan het ontwikkelen van methodologieën om een impactstudie te kunnen bewerkstelligen. Evenmin zal het verwonderen dat er voortdurend wordt gedebatteerd over welke methodologie in welke context het meest geschikt is, en dat impactstudies herhaaldelijk tegen een kritisch licht worden gehouden.

In dit hoofdstuk geven we een overzicht van enkele veelgebruikte methodes om impact te meten. Welke methode men kan toepassen of het meest verkieslijk is, hangt af van de specifieke context. We geven dan ook per methodologie aan welke elementen aanwezig moeten zijn om ze te kunnen gebruiken. Doorheen het hoofdstuk zal ook herhaaldelijk worden benadrukt dat een methode enkel correcte resultaten oplevert indien de basisveronderstellingen die de methode vereist, inderdaad opgaan. Een veronderstelling is per definitie iets wat men niet hard kan bewijzen, maar vaak kan men wel nagaan (door middel van een combinatie van *data mining* en gebruik van het gezond verstand) of een bepaalde veronderstelling in de specifieke context van de impactstudie aannemelijk is.

De methodes zullen worden geïllustreerd aan de hand van voorbeelden uit de literatuur, waarin maatregelen en factoren die normconform handelen kunnen beïnvloeden worden geëvalueerd. Succesvolle academische studies in dit gebied hebben vaak de volgende eigenschappen:

- De studies hebben een hoge mate van interne validiteit: Vaak gaan onderzoekers op zoek naar zeer specifieke case studies waarbij men op een heel overtuigende manier kan argumenteren dat de nodige aannames voor een zuivere impactstudie correct zijn.
- De studies hebben een hoge mate van externe validiteit: Onderzoekers zijn niet meteen geïnteresseerd in de impact van een specifieke instantie in een specifiek land. De case study is immers vaak enkel gekozen omwille van de opportuniteiten voor het bereiken van een hoge mate van interne validiteit. Men is echter meer geïnteresseerd om, door middel van de impactstudie,

lessen te trekken voor een veel breder publiek dan voor de instantie waarop de case study betrekking heeft. Vaak wordt de studie dan ook gekaderd binnen een bredere literatuur, en wordt de impactstudie gebruikt om een theorie te testen in verband met normconform handelen. De resultaten worden dan interessant voor een breed internationaal lezerspubliek van beleidmakers en academici.

Omdat het in de literatuur vaak gaat om daartoe uitgezochte voorbeelden, zal een inspectie in de praktijk de validiteit van de nodige assumpties niet op eenzelfde manier kunnen beargumenteren als in de literatuur. Het goede nieuws is echter dat beleidmakers en inspecteurs vaak beter zullen kunnen inschatten in hoeverre de assumpties gerechtvaardigd zijn dan een internationaal lezerspubliek, omdat zij immers bovenop de materie zitten en beschikken over heel wat praktijkervaring in de te evalueren materie.

Samengevat zijn de doelstellingen van dit hoofdstuk:

- Een geïllustreerd overzicht geven van de verschillende methodes om impact te meten;
- Duidelijk te maken dat er onder de voorgestelde methodes geen “gouden standaard” bestaat, maar dat men in elke context moet afwegen welke methode toepasbaar is en de meest zuivere resultaten oplevert;
- Duidelijk te maken bij iedere methode welke aannames vereist zijn, zodat beleidmakers een goed kader hebben om te beslissen of het verantwoord is om een bepaalde impactstudie (volgens een bepaalde methode) uit te voeren.

3.2 Achtergrond

Zoals bekend uit handboeken over programmaevaluatie, en zoals ook wordt aangegeven in documentatie inzake impactevaluatie van inspectie (bv. in Werkgroep Effectmeting IG-Beraad, 2005) is het allereerst belangrijk doelstellingen (beoogde effecten) te formuleren. Vervolgens moet men dan op zoek naar indicatoren die de effecten kunnen meten. Meestal zijn deze indicatoren proxy-variabelen: de eigenlijke doelstelling meten is vaak erg moeilijk, en daarom doet men vaak beroep op proxies of proxy-variabelen. Dat zijn variabelen die eenvoudiger te observeren zijn, en waarvan men aanneemt dat ze in hoge mate gecorreleerd zijn met het al dan niet bereiken van de doelstelling.

Enige omzichtigheid bij het kiezen van proxy-variabelen is echter nodig. Een daling van het aantal meldingen over zwartwerk kan erop wijzen dat het zwartwerk inderdaad daalt, maar kan ook op een heel aantal andere manieren geïnterpreteerd worden, zoals het optreden van meldingsmoeheid van de burger. Bovendien delen we de bezorgdheid van het IWSE dat een positief effect op normconform handelen van subjecten (bedrijven etc.) niet noodzakelijk leidt tot de gewenste maatschappelijke veranderingen. Zo is recentelijk in de media ophef ontstaan over het feit dat veel interimkantoren nog steeds, op vraag van klanten, discrimineren in het nadeel van allochtone werknemers. Indien na verloop van tijd blijkt dat door een

strengere inspectie interimkantoren niet ingaan op de vraag van bedrijven om allochtonen weg te filteren uit selecties, impliceert dit nog niet automatisch dat de eigenlijke doelstelling, het verhogen van de arbeidsmarktkansen van de allochtoon, gerealiseerd is.

Zoals aangegeven in WODC (2010) en op de webstek van het “anti-Corruption Research Network”, worden nieuwe inspectiestrategieën idealiter geëvalueerd via pilootstudies. Hierbij geldt dat een nieuw project of een nieuwe werkwijze geleidelijk wordt ingevoerd. Een willekeurig gekozen subgroep van de subjecten die onder de inspectiebevoegdheid van het inspectieorgaan valt, dient dan te worden onderworpen aan het nieuwe, experimentele beleid, terwijl voor de andere groep alles bij het oude blijft. Ook al wordt zo'n “gerandomiseerd experiment” door velen gezien als de gouden standaard, toch heeft zij ook haar beperkingen. Vaak is het immers omwille van ethische, juridische of technische redenen niet mogelijk om een willekeurig deel van de subjecten aan een ander regime te onderwerpen. Zo is het niet mogelijk om twee bedrijven die een identieke inbreuk hebben gepleegd een verschillende sanctie op te leggen. Als men een mediacampagne op touw zet om te wijzen op de noodzaak van het naleven van arbeidsvoorwaarden, is het niet mogelijk om de campagne-boodschappen enkel te laten verschijnen op de televisieschermen van een willekeurig gekozen subgroep van bedrijfsleiders en aandeelhouders.

Een ander belangrijk nadeel is dat gerandomiseerde experimenten enkel geschikt zijn voor het bepalen van micro-effecten. Macro-effecten, die enorm relevant zijn wanneer een bepaald pilootproject uit de testfase wordt gehaald en landelijk wordt ontplooid, kunnen niet geanalyseerd worden met behulp van een gerandomiseerd experiment (Imbens en Wooldridge, 2009). Veronderstel bijvoorbeeld dat een inspectieorgaan wil inzetten op het terugdringen van tewerkstelling van illegalen, en dat men een willekeurige groep van bedrijven gaat onderwerpen aan een meer frequente inspectie op dit soort van arbeid. Het is best mogelijk dat men na verloop van tijd vaststelt dat bedrijven die vaker worden geïnspecteerd ook meer conform gaan handelen naar de regels. Er wordt dan echter nog geen rekening gehouden met macro-effecten die gaan meespelen wanneer de hele populatie van bedrijven aan strenger toezicht onderworpen wordt. Omdat illegale arbeid door de meer frequente inspectie minder aantrekkelijk wordt en dus de vraag naar deze arbeid daalt, kunnen de illegale werknemers ervoor kiezen om hun diensten aan te bieden onder gunstigere voorwaarden, waardoor het opnieuw aantrekkelijker wordt voor bedrijven een beroep te doen op deze arbeid.

Om deze redenen is men vaak aangewezen op andere evaluatiemethoden. Zo kan men bijv. het gedrag van subjecten (bijv. ondernemingen), die onder de bevoegdheid van het inspectieorgaan vallen, gaan vergelijken net voor en net na de invoering van een nieuw regime (natuurlijk experiment). Bij deze werkwijze dient men er wel op te letten dat het resultaat niet wordt beïnvloed door andere factoren, die tussentijds zijn veranderd. Ook kwalitatieve benaderingen zijn mogelijk om een zicht te krijgen op de impact die het inspectie-optreden heeft. In dit hoofdstuk zal worden geïllustreerd dat een zorgvuldige afweging van voor- en nadelen bij de keuze van een bepaalde onderzoeksmethode op zijn plaats is.

De IWSE wil naast het actief opsporen van mistoestanden ook meer gaan inzetten op preventiecampagnes. Derhalve wordt er ook aan deze strategie (en aan de evaluatie daarvan) aandacht besteed in dit hoofdstuk. Naast het actief op zoek gaan naar overtredingen is de bestrijding van overtredingen, door bijvoorbeeld burgers de mogelijkheid te bieden fraudegevallen aan te geven en door mediacampagnes op te starten die waarschuwen voor de consequenties van het niet naleven van de wetgeving, een belangrijke sleutel om het handelen conform de wetgeving te stimuleren (United Nations, 2004).

3.3 Veel voorkomende terminologie en concepten bij kwantitatieve data-analyse

Vooraleer we overgaan tot het bespreken van de methodes die gebruikt kunnen worden om impact te meten, dienen eerst enkele begrippen en bouwstenen te worden besproken, die later in het hoofdstuk en het rapport aan bod zullen komen.

3.3.1 Enkele statistische basisbegrippen

Hieronder definiëren we kort enkele begrippen die in de rest van deze sectie (en het rapport) herhaaldelijk zullen voorkomen.

- **Populatie:** een groep van subjecten (individuen, bedrijven, steden...) waarop de studie betrekking heeft. Bijv., wanneer de Inspectie WSE de impact wil meten van haar optreden op het correct gebruik van de tewerkstellingspremie 50+, behelst de populatie alle bedrijven die een werknemer in dienst hebben waarvoor een dergelijke premie werd aangevraagd of werd goedgekeurd.
- **Steekproef:** een (al dan niet) willekeurige trekking van subjecten uit de populatie
- **Variabele:** Een kenmerk dat kan veranderen tussen subjecten van een populatie, of dat kan veranderen voor eenzelfde subject over de tijd (lengte van individuen, gemiddelde jaartemperatuur over landen, de gemiddelde dagtemperatuur binnen een land over het jaar etc.).

3.3.2 Soorten variabelen: opdeling volgens type

Men kan variabelen opdelen volgens type en volgens de functie in een model. In deze paragraaf wordt de type-indeling besproken, in de volgende paragraaf de functionele indeling.

Variabelen kunnen opgedeeld worden in functie van de getallenverzameling waartoe ze behoren:

Een *dummy-variabele* kan enkel de waarden 0 en 1 aannemen. Werd er een inbreuk vastgesteld nee/ja? Is men van het mannelijk geslacht, nee/ja?

Een *categorische variabele* kan een beperkt aantal waarden aannemen en beschrijft categorische data. Deze categorieën kunnen van nature geordend zijn, zoals:

- 1 = kleiner dan 10cm;
- 2 = tussen 10 en 20cm;
- 3 = groter dan 20cm.

Categorieën kunnen ook van nature ongeordend zijn, zoals bijvoorbeeld data die de mogelijke transportmodi beschrijven:

- 1 = per trein;
- 2 = per bus;
- 3 = per auto.

Discrete variabelen zijn variabelen die waarden aannemen uit de verzameling van de natuurlijke getallen (1, 2, 3 ...). Voorbeelden van discrete variabelen zijn het aantal inwoners per stad, de grootte van een veestapel van een landbouwbedrijf etc.

Een *continue variabele* neemt waarden aan uit de verzameling van de reële getallen, bijv. het aantal inwoners per vierkante kilometer in een land, het aantal kindersterftes per duizend geboortes in een land etc.

3.3.3 Soorten variabelen: opdeling volgens functie

Afhankelijke variabelen

De afhankelijke variabele is een kenmerk dat "afhangt" van andere factoren, of er door wordt bepaald. Een afhankelijk variabele is bijvoorbeeld het non-conform handelen, dat af kan hangen van het inspectieoptreden. De afhankelijke variabele meet idealiter een eindresultaat, waaruit men dus moet kunnen afleiden in welke mate een doelstelling werd gehaald. In het geval van een inspectiedienst kan men dus op zoek aan naar maatstaven die normconform handelen meten. Deze variabelen kunnen zich situeren op microniveau, dat wil zeggen, ze worden geregistreerd op het niveau van de geïnspecteerde, alsook op macroniveau, dat wil zeggen, men meet normconform handelen op een hoger niveau dan het geïnspecteerd object. Een paar voorbeelden kunnen dit verduidelijken.

Meting op microniveau:

Veronderstel dat de inspectiedienst inspecties uitvoert op illegale tewerkstelling, en tot doelstelling heeft om het aantal bedrijven dat beroep doet op illegale arbeidskrachten te doen dalen. Men kan dan een afhankelijke dummy-variabele definiëren die de waarde 1 aanneemt indien bij een inspectie in een bedrijf illegalen werden aangetroffen, en de waarde 0 in het andere geval. Men observeert nu voor elk geïnspecteerd bedrijf of deze zich aan de regels heeft gehouden of niet. Een daling van de ratio non-conform en conform handelende bedrijven over de tijd duidt er dan op dat men op de goede weg is om de doelstelling te bereiken. In hoofdstuk 5 worden voorbeelden gegeven hoe zogenaamde kwantitatieve rapportagebestanden bij de IWSE per materie kunnen worden opgebouwd.

Meting op macroniveau:

Veronderstel dat een inspectie fraude met BTW bij ondernemingen wil onderzoeken. Indien ondernemingen (die niet de consument zijn) goederen aankopen bij een leverancier, kunnen zij de BTW terugvorderen bij de BTW-administratie. Aan hun eigen klanten dienen zij echter opnieuw BTW aan te rekenen, en deze door te storten aan de BTW-administratie. In de regel is per boekhoudkundig jaar het saldo dat de onderneming moet storten aan de BTW-administratie positief: een onderneming die geen toegevoegde waarde creëert of die kampt met een stock die voortdurend aangroeit, kan op termijn niet overleven. Wanneer een bedrijf een negatief BTW-saldo opgeeft is men dus niet per definitie schuldig aan fraude, maar het is wel een groot knipperlicht. In dit geval geeft het BTW-saldo op individueel niveau dus niet aan of een bedrijf fraudeert of niet (meer onderzoek is nodig), maar gemiddeldes per sector, en de evoluties van deze gemiddeldes over de tijd, geven ons wel een goed beeld van het niveau en de evolutie van normconform handelen. Een dergelijke meting op macroniveau met behulp van administratieve data laat bovendien toe om na te gaan of er spillover-effecten bestaan, en dus of inspectieoptredens een effect hebben op de gehele populatie en niet enkel op de steekproef van geïnspecteerden. Dit is zeker interessant voor materies met een lage dekkinggraad, waarbij het onmogelijk is om iedereen op zeer regelmatige basis aan een inspectie te onderwerpen. Voor het verkrijgen van informatie over BTW zal de IWSE wel een samenwerkingsakkoord moeten sluiten met de federale overheidsdienst Financiën of dient men dit te comprimeren via de jaarrekeningen van de ondernemingen.

Onafhankelijke variabelen

Onafhankelijke, of verklarende, variabelen, zijn variabelen die in het model de afhankelijke variabele proberen te verklaren, zoals bijvoorbeeld het inspectieoptreden. Een onafhankelijke variabele kan bijv. een indicator zijn die varieert over de tijd, en 0 aanneemt wanneer een bepaalde interventie of beleidsmaatregel nog niet in voege is, en de waarde 1 vanaf het ogenblik waarop dat wel zo is. In hoofdstuk 5 wordt het gebruik van deze informatie in kwantitatieve databestanden verder uitgelegd.

Een model kan onafhankelijke variabelen bevatten die zijn gemeten op het niveau van de afhankelijke of te verklaren variabele. Maar sommige onafhankelijke variabelen kunnen gemeten zijn op een hoger niveau. Stel bijvoorbeeld dat de afhankelijke (i.e. te verklaren) variabele wordt gemeten op het niveau van de geïnspecteerde, namelijk een dummy-variabele die de waarde 1 aanneemt wanneer er een inbreuk is vastgesteld en de waarde 0 in het andere geval. Men heeft metingen voor verschillende jaren, en de volgende onafhankelijke variabelen kunnen worden gebruikt in het model om de kans op een inbreuk te verklaren:

- Het aantal tewerkgestelden in een bedrijf: Deze informatie wordt gemeten voor elke observatie apart, en kan dus variëren over de geïnspecteerden.
- Een dummy-variabele die aangeeft tot welke sector het geïnspecteerde bedrijf behoort: Deze informatie wordt niet gemeten op het niveau van de

afhankelijke variabele, en varieert dus niet per se van observatie tot observatie maar varieert wel over groepen van observaties.

- Een dummy die aangeeft in welk jaar de meeting plaats vond: Ook in dit geval varieert de waarde van de verklarende variabele niet van observatie tot observatie, maar wel over groepen van observaties. In dit geval zijn de groepen waarover de variabele varieert niet de bedrijfssectoren, maar wel de tijdsperiodes waarin de metingen plaatsvonden.

Men zal in de literatuur soms ook de term “instrumentele variabele” of “instrument” tegenkomen. *Instrumentele variabelen* worden soms gebruikt om een causaliteitsprobleem op te lossen. We komen hierop nog terug bij de verschillende methodologieën in paragraaf 6 van dit hoofdstuk.

Een proxy of *proxy-variabele* is een onafhankelijke variabele die het fenomeen wat men wil meten niet rechtstreeks meet, maar er wel in hoge mate mee gecorreleerd is. Denken we terug aan het voorbeeld van het BTW-saldo: een negatief BTW-saldo meet niet rechtstreeks non-conform handelen, maar is er wel sterk mee gecorreleerd. Later in dit hoofdstuk komen nog een aantal van zulke voorbeelden aan bod.

3.3.4 Data-analyse: begrippen en concepten

Een populatiemodel: Een populatiemodel is een model dat een bepaalde wetmatigheid beschrijft, de causale relatie tussen een aantal onafhankelijke variabelen enerzijds, en de afhankelijke variabele anderzijds. Een populatiemodel kan er als volgt uitzien:

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \varepsilon$$

In bovenstaand model is Y de afhankelijke variabele, en zijn X_1 en X_2 verklarende of onafhankelijke variabelen. α is een constante term, en β_1 en β_2 zijn parameters of coëfficiënten die de relatie tussen de onafhankelijke variabelen en de afhankelijke variabele kwantificeren. ε is een storingsterm, en bevat alle factoren die een invloed hebben op de afhankelijke variabele en die niet expliciet zijn opgenomen in het populatiemodel.

Ceteris paribus assumptie: Deze term wordt vaak gebruikt in causale impactstudies. Hoe verandert een afhankelijke variabele in functie van een onafhankelijke variabele als al het overige gelijk blijft? Wanneer we spreken over een causaal verband tussen X en Y denken we steeds in ceteris paribus termen: hoe verandert Y als X verandert en alle andere factoren ongewijzigd blijven?

Regressie(analyse): In regressie(analyses) probeert men aan de hand van beschikbare data het populatiemodel te schatten. Een regressievergelijking die bovenstaand populatiemodel schat ziet er als volgt uit:

$$Y = \hat{\alpha} + \beta_1 X_1 + \beta_2 X_2 + \hat{\varepsilon}$$

De hoedjes op de parameters en de constante geven aan dat het hier gaat om geschatte waardes. Indien het model correct gespecificeerd is zullen deze geschatte waardes convergeren naar α , β_1 en β_2 . Naarmate de steekproef groter wordt, worden de schatters steeds nauwkeuriger. Indien er zich causaliteitsproblemen voordoen, bijv. wanneer er in de storingsterm variabelen vervat zitten die zowel met de onafhankelijke als afhankelijke variabelen gecorreleerd zijn, zullen de geschatte parameters niet convergeren naar de eigenlijke parameters: de schattingen zijn dan inconsistent. $\hat{\varepsilon}$ is de schatting van de storingsterm, en wordt een restterm genoemd.

Variabelen kunnen voorkomen (in populatiemodellen en regressieanalyses) in een heel aantal transformaties. De meest eenvoudige relatie tussen de afhankelijke en een onafhankelijke variabele is de lineaire relatie, maar in realiteit zijn relaties vaak veel complexer. Zo kan de relatie tussen een onafhankelijke en afhankelijke variabele kwadratisch zijn, of logaritmisch. Een belangrijk concept dat nog zal terugkomen is de interactieterm. Een interactieterm is een product van twee onafhankelijke variabelen die ook op zichzelf in de regressie voorkomen. Een interactieterm wordt toegevoegd aan een model wanneer men verwacht dat het effect van de ene variabele afhangt van de hoogte van de andere variabele.

Laten we deze sectie afsluiten met een eenvoudig voorbeeld om de hierboven geschetste concepten en terminologieën wat te concretiseren. Een landbouworganisatie wil het verband kennen tussen opbrengst van de oogst enerzijds, en irrigatie en bemesting anderzijds. De oogst (ton graan per ha) is de afhankelijke variabele gelabeld Y . De eerste onafhankelijke variabele X_1 is het aantal liter water per ha, en de tweede onafhankelijke variabele X_2 is het aantal kilogram mest per ha. Men kan echter een interactie-effect veronderstellen. Op een kurkdroge akker zal wat extra mest niet veel uithouden, en omgekeerd zal op een onbemeste akker een goede irrigatie ook maar een matig effect hebben op de oogst. Daarom vermoedt men een interactie-effect tussen het aantal liter water per ha en het aantal kilogram mest per ha. Kort samengevat heeft men volgend populatiemodel in gedachten:

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 * X_2 + \varepsilon$$

α is een constante, β_1 , β_2 , en β_3 zijn parameters en ε is een storingsterm die allerlei factoren bevat (zoals bodemkwaliteit) die niet in het model zijn geëxpliciteerd.

Men kan nu het model aan de hand van data gaan schatten met een regressieanalyse die er als volgt uitziet:

$$Y = \hat{\alpha} + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 * X_2 + \hat{\varepsilon}$$

Bij afwezigheid van causaliteitsproblemen zijn $\hat{\alpha}$, β_1 , β_2 , β_3 en $\hat{\varepsilon}$ consistente schatters voor α , β_1 , β_2 , β_3 en ε . Jammer genoeg zijn er goede redenen om

te twijfelen aan de afwezigheid van causaliteitsproblemen. Immers, in de storingsterm zitten factoren zoals de bodemkwaliteit, en de bodemkwaliteit is een variabele die mogelijk gecorreleerd is met irrigatie en bemesting. Landbouwers zouden er bijvoorbeeld voor kunnen kiezen om een slechte bodemkwaliteit te compenseren door meer mest en water te gebruiken. De variabelen die bemesting en irrigatie meten pikken dan ook het effect op van de bodemkwaliteit: hoe meer bemesting, hoe meer water, hoe lager de bodemkwaliteit. Hierdoor verkrijgen we inconsistente schatters, en de causale impact van bemesting en irrigatie op de opbrengst van de oogst wordt onderschat.

3.4 Types databestanden

In deze paragraaf beschrijven we enkele categorieën van data. Het gaat hier niet zozeer over het type bron (primaire of secundaire data), maar wel over de structuur en type informatie die de data bevatten.

Cross-sectionele data

Men spreekt over een cross-sectionele dataset wanneer de data gegevens bevatten van een populatie (of steekproef daarvan) voor slechts één tijdsperiode. De populatie kan bestaan uit individuen, bedrijven, landen enzovoort. Elke survey die dus eenmalig wordt georganiseerd levert in principe cross-sectionele data op. Als men uit een databank (zoals de World Development Indicators van de Wereldbank) alle gegevens extraheert voor de verschillende landen voor een bepaald jaar, bijvoorbeeld 2006, verkrijgt men ook een cross-sectionele dataset. De structuur van een cross-sectionele dataset wordt geïllustreerd in Tabel 3.1.

Tabel 3.1
Structuur cross-sectionele dataset

Periode	Subject	Variabele 1	Variabele 2	Variabele 3
1	1			
1	2			
1	3			
1	4			

Herhaalde cross-sectionele data

Herhaalde cross-sectionele data bevatten gegevens van verschillende steekproeven uit een populatie. De dataverzameling gebeurt voor iedere steekproef tijdens een andere tijdsperiode. Als men de steekproeven telkens via dezelfde procedures opbouwen, zodat ze vergelijkbaar zijn, is het mogelijk om trends over de tijd te analyseren.

De periodiciteit waarmee de steekproeven worden getrokken en de data worden verzameld, verschilt van dataset tot dataset. De verkiezingspoll van de Vlaamse VRT bijvoorbeeld, wordt in verkiezingstijd maandelijks afgenomen. Een ander voorbeeld van een herhaalde cross-sectionele dataset is de Eurobarometer Survey, die wordt

georganiseerd in de lidstaten van de Europese Unie sinds 1970. Vaak worden aan de survey speciale modules toegevoegd, om in te kunnen zoomen op bepaalde aspecten die de Europese Commissie of de lidstaten op dat moment belangrijk achten. Omdat deze speciale modules vaak slechts eenmalig worden uitgezet, bekomt men voor deze modules slechts een momentopname, of een enkelvoudige cross-sectie. De standaard Eurobarometer Survey daarentegen, wordt jaarlijks in de lente en de herfst uitgevoerd en vele vragen blijven ongewijzigd over de tijd. Op die manier heeft de Eurobarometer Survey een erg belangrijke herhaalde cross-sectionele dataset opgeleverd, die niet alleen analyses per land toelaat, maar ook tussen landen en over de tijd. In België organiseert de Federale Overheidsdienst Economie om de drie maanden de Enquête naar de Arbeidskrachten (EAK), wat ook herhaalde cross-sectionele data oplevert.

Niet elke herhaalde cross-sectionele dataset kent eenzelfde strakke regelmaat. Een goed voorbeeld hiervan is de World Values Survey, een survey die wereldwijd wordt uitgezet ten einde socioculturele en politieke verschillen over landen en over de tijd mogelijk te maken. De data bevatten nu gegevens uit vijf verschillende rondes, zijnde de periodes 1981-1984, 1989-1993, 1994-1999, 1999-2004, en 2005-2008. Niet alleen is de periodiciteit van deze survey vrij onregelmatig, ook worden niet alle landen opgenomen in elke ronde. Desalniettemin heeft deze dataset een enorme bijdrage geleverd aan internationaal socio-economisch onderzoek. De structuur van een herhaalde cross-sectionele dataset wordt geïllustreerd in Tabel 3.2.

Tabel 3.2
Structuur herhaalde cross-sectionele dataset

Periode	Subject	Variabele 1	Variabele 2	Variabele 3
1	1			
1	2			
1	3			
1	4			
2	5			
2	6			
2	7			
2	8			

Paneeldata

Paneeldata bevatten gegevens van een populatie (of steekproef van deze populatie) voor meer dan één periode. Het verschil tussen paneeldata en herhaalde cross-sectionies bestaat erin dat in paneeldata men tracht dezelfde steekproef op te volgen in elke periode waarin data worden verzameld.

Een bekend voorbeeld van een paneeldataset in België is de Paneelstudie van Belgische Huishoudens (PSBH). Het project liep tussen 1992 en 2002 en bevat 11 rondes: Jaarlijks werden ruim 4.400 gezinnen of ruim 11.000 personen geënquêteerd. Tussen 1994 en 2002 werd data van de PSBH ook geïntegreerd in de ondertussen ook stopgezette European Community Household Panel (ECHP), die

paneeldata uit de verschillende lidstaten uit de Europese Unie bijeen brachten teneinde vergelijkingen over de lidstaten mogelijk te maken. De opvolger van de ECHP is de Statistics on Income and Living Conditions (SILC). Dit project werd opgestart in 2003 bij een beperkt aantal EU-lidstaten waaronder België, en sinds 2004 dragen alle EU-lidstaten bij aan deze dataset. In België is de Algemene Directie Statistiek verantwoordelijk voor de operationalisatie van het project.

Paneeldata kunnen in sommige onderzoekscontexten enorme voordelen hebben tegenover herhaalde cross-secties, zoals verder in dit rapport zal worden besproken. Hoewel minder benadrukt in de literatuur, kunnen paneeldata ook nadelen hebben tegenover herhaalde cross-secties. In het geval van enquêtes afgenomen van individuen, kan het immers heel wat inspanning en middelen kosten om alle observaties uit de steekproef te blijven volgen. Mensen kunnen immers verhuizen of niet meer geïnteresseerd zijn om aan de enquête deel te nemen. In het geval van dataverzameling van bedrijven, kan men te maken krijgen met faillissementen, fusies etc. Indien de elementen die “uitvallen” niet representatief zijn voor de steekproef, moet men opletten geen vertekend beeld te schetsen van trends of relaties over de tijd. Tabel 3.3 illustreert de structuur van een paneeldataset.

Tabel 3.3
Structuur herhaalde cross-sectionele dataset

Periode	Subject	Variabele 1	Variabele 2	Variabele 3
1	1			
1	2			
1	3			
1	4			
2	1			
2	2			
2	3			
2	4			

Pseudo-paneeldata

Soms gebeurt het dat men herhaalde cross-secties gebruikt om paneeldata te construeren, waarbij de eenheid van observatie in de pseudo-paneeldata dan op een hoger niveau ligt dan de eenheid van observatie in de herhaalde cross-secties. Stel dat men over data beschikt op ondernemingsniveau. Men kan dan voor de variabelen aanwezig in de herhaalde cross-secties een gemiddelde berekenen per periode en per sector. Als men ervan kan uitgaan dat de steekproeftrekking in elke ronde op dezelfde manier gebeurde, kan men dus een paneeldataset creëren waarbij men elke sector meermaals observeert in de dataset, namelijk een keer per periode.

(Meervoudige) tijdreeksen

Bij tijdreeksen observeert men één subject voor een heel aantal periodes. Een tijdreeks kan bijvoorbeeld zijn het Bruto Binnenlands Product (BBP) van België op kwartaalbasis van 1940 tot 2012. Bij meervoudige tijdreeksen observeert men voor

dit subject verschillende variabelen, bijvoorbeeld niet enkel het BBP maar ook werkloosheidsgraad, aantal criminele feiten, etc. Lange (meervoudige) tijdreeksen met een hoge periodiciteit kunnen vaak erg interessante verbanden blootleggen. Tabel 3.4 illustreert de structuur van een (meervoudige) tijdreeks.

Tabel 3.4
Structuur (meervoudige) tijdreeksdataset

Periode	Subject	Variabele 1	Variabele 2	Variabele 3
1	1			
2	1			
3	1			
4	1			
5	1			
6	1			
7	1			
8	1			

Cross-sectionele tijdreeksen

Een cross-sectionele tijdreeks heeft dezelfde structuur als een paneeldataset. Een paneeldataset wordt echter gekenmerkt door een hoge ratio van subjecten en periodes, terwijl dit net het omgekeerde is bij cross-sectionele tijdreeksen. Men observeert bijvoorbeeld het Bruto Binnenlands Product (en andere macro-economische indicatoren) voor de Eurozone gedurende 48 kwartalen. Er zijn dan meer periodes dan landen. De structuur van een cross-sectionele tijdreeks wordt weergegeven in Tabel 3.5.

Tabel 3.4
Structuur cross-sectionele tijdreeks

Periode	Subject	Variabele 1	Variabele 2	Variabele 3
1	1			
1	2			
2	1			
2	2			
3	1			
3	2			
4	1			
4	2			

3.5 Correlatie versus causaliteit of impact

“Only an idiot, or an anti-American bigot prepared to maintain that Americans are four times more murderous than Britons, could possibly pretend that no connection exists between those figures and the fact that 300m guns are “out there” in the United

States, more than one for every adult.” (Uit: the Economist, 22/12/2012, “Gun violence in America. Newtown’s horror”)

Via regressie-analyses, waarvan de routines in statistische softwarepakketten (en ook in Excel) geprogrammeerd zijn, is het eenvoudig om na te gaan of en welk verband er bestaat tussen twee variabelen. Een verband of correlatie vaststellen betekent echter niet noodzakelijk dat een causaal verband werd geïdentificeerd. Het is voor beleidmakers erg belangrijk om deze twee uit elkaar te houden. We illustreren dit hieronder met een fictief voorbeeld, echter geïnspireerd door bestaande literatuur die later nog uitvoeriger wordt besproken.

We definiëren een variabele Y die het jaarlijks aantal criminele feiten per duizend inwoners in een stad weergeeft, en X een variabele die het aantal politieagenten per duizend inwoners in een stad meet. Een regressie van Y op X levert ons voor X een geschatte coëfficiënt op van 0,2. Dit wil zeggen dat in steden met meer politieagenten per duizend inwoners er meer criminele feiten worden gepleegd. Indien in stad A er per duizend inwoners 5 agenten meer zijn dan in stad B, voorspelt het model in stad A 1 misdrijf per duizend inwoners meer dan in stad B. Het geschatte verband impliceert echter geenszins een oorzakelijke relatie. Indien beleidmakers aan de hand van deze schattingen in stad A zouden beslissen om de politiemacht te doen inkrimpen, is het verre van zeker dat ook de criminaliteit in de stad zal dalen, integendeel zelfs. Immers, de reden waarom men in stad A een grote politiemacht heeft uitgebouwd, heeft misschien te maken met de hoge criminaliteit aldaar. Een hoge criminaliteit veroorzaakt dan een grote politiemacht eerder dan omgekeerd.

Later in het hoofdstuk bespreken we enkele studies die een positieve correlatie tonen tussen criminaliteit en politiemacht, maar die echter argumenteren dat er een negatief oorzakelijk verband bestaat tussen politiemacht en het aantal criminele daden. In de rest van het hoofdstuk bespreken we dan ook methodes die kunnen helpen om de stap te zetten van een identificatie van een correlatie of verband tot de identificatie van een causaal of oorzakelijk verband, of impact.

Zoals in de inleiding van dit hoofdstuk is aangehaald, zijn er heel wat methodes ontwikkeld, en het is aan de onderzoeker en/of opdrachtgever om te bepalen welke methode in een bepaalde context het meest geschikt is, en of de assumpties die men maakt aannemelijk zijn.

In de meer gerespecteerde academische vakbladen is men heden ten dage geobsedeerd door causaliteit, en editors en referenten zullen vaak erg kritisch kijken naar elke bewering van een gevonden causale relatie: kosten noch moeite worden gespaard om tegenargumenten te zoeken om in twijfel te trekken dat de relatie die men heeft geïdentificeerd potentieel toch niet zuiver causaal is. Deze hoge standaarden die men oplegt om verbanden als causaal te aanvaarden is uiteraard niet geheel onterecht, zoals het voorbeeld in het begin van deze paragraaf aangeeft. Een correlatie verkeerd interpreteren als een causaal verband kan immers leiden tot beleidsinterventies die het omgekeerde van het gehoopte effect hebben. De

obsessie voor causaliteit brengt echter wel gevaren met zich mee, namelijk dat bepaalde fenomenen in de maatschappij compleet genegeerd dreigen te worden. De quote aan het begin van deze paragraaf, die we tegenkwamen bij het lezen van de kersteditie van "the Economist", leek ons een mooi voorbeeld om aan te halen als uiting van ongenoegen tegenover een soms te ver doorgedreven drang naar ontkenning van causaliteit. Als men niet staalhard kan maken dat er een causaal verband bestaat tussen het aantal wapens in omloop en het aantal moorden in een land, wil dit immers niet zeggen dat deze causale relatie niet bestaat.

Twee weken later, echter, verscheen er in hetzelfde blad een lezersbrief als reactie op deze quote, die we dan ook hieronder citeren:

"SIR – There is simply no credible documented evidence of causality between murder rates and the strictness of gun controls. Yet you stated that "Only drastic gun control could make a big difference" after the massacre at Sandy Hook ("Newtown's horror", December 22nd). The very simple-minded argument that more guns equates with more murder is false. Germany has five times the rate of private-gun ownership than Britain yet Germany's murder rate is half that of Britain. The rate of gun ownership in Russia is much lower than in the United States, yet its homicide rate is twice that of America.

The National Academy of Sciences published a critical review of firearms and violence in 2004 and could not find any credible evidence to answer even basic questions about the efficacy of gun control. You mentioned that gun bans in Chicago and Washington, DC, were struck down by the Supreme Court. What you didn't point out is that these cities had some of the highest murder rates in the country even when their ban on guns were in use. The crime and murder rates in both places after the gun bans were removed actually fell at a higher rate than the national average.

Do you still believe that "only an idiot" could argue that there is no connection between gun ownership and the murder rate?

Ben Russo

Herndon, Virginia"

Het moge dus duidelijk zijn dat de zoektocht naar impact of causale verbanden steevast leidt tot woord en wederwoord, intense en zelfs bitsige discussies naargelang het onderwerp en het maatschappelijke belang hiervan. Het voorbeeld illustreert nog maar eens dat onderzoekers/opdrachtgevers zich bewust moeten zijn van de sterktes maar ook zwaktes van de impactstudie, en dat duidelijke doch genuanceerde conclusies belangrijk zijn opdat de impactstudie kan bijdragen tot een breed en sereen maatschappelijk debat.

3.6 Methoden voor het meten van impact

Indien men poogt een causaal verband te schatten maar geen rekening houdt met het feit dat een verschil bestaat tussen correlatie en causaliteit, kan de schatter *vertekend* zijn. In het eerder gegeven fictieve voorbeeld was de schatting van 0,2 van het effect van politiemacht op criminaliteit ernstig vertekend. Immers, de correlatie tussen beide variabelen is positief maar het oorzakelijke verband negatief. Hieronder

bespreken we enkele veel gebruikte methodes om zulke vertekeningen te voorkomen.

Opnemen van controlevariabelen

Bij een impactstudie is men meestal geïnteresseerd in de impact van één enkele verklarende variabele op de verklaarde variabele. Regressie- en matching-modellen laten echter toe veel meer dan één verklarende variabele op te nemen in het model. Onderzoekers voegen dan ook vaak zogenaamde controlevariabelen toe aan het model. Men doet dit niet omdat men geïnteresseerd is in de impact van deze variabelen op de afhankelijke variabele, maar omdat deze “ceteris paribus” analyses mogelijk maken en men zo een stap dichterbij een causale analyse komt. Bijvoorbeeld: Wat is de relatie tussen politiemacht en criminaliteit, indien we andere factoren constant houden?

Het is belangrijk om een variabele in een model op te nemen als controlevariabelen indien deze laatste gecorreleerd is met de hoofdverklarende variabele en als men bovendien vermoedt dat de variabele een causale impact heeft op de verklaarde variabele. Doet men dit niet, dan zal de hoofdverklarende variabele de “impact” van de weggelaten variabele oppikken. Toch dient men omzichtig om te springen met controlevariabelen. Er zijn immers verschillende scenario's denkbaar waarin het toevoegen van controlevariabelen de vertekening eerder zal doen toenemen dan doen afnemen. Dit kan met name voorkomen wanneer de hoofdverklarende variabele zelf een causale impact heeft op de controlevariabele.

Jammer genoeg is reeds meermaals aangetoond in de literatuur dat het toevoegen van controlevariabelen in een model ons vaak niet erg veel verder brengt. Beschikbare variabelen in datasets zijn dikwijls slechts zeer povere maatstaven voor de factoren die men constant zou willen houden, en er zijn ook vaak onobserveerbare verklarende variabelen die dus helemaal niet aanwezig zijn in de dataset.

Paneeldata en pseudo-paneeldata hebben het enorme voordeel dat ze toelaten om technieken te gebruiken die controleren voor alle kenmerken die niet veranderen over de tijd. Vele karakteristieken waarvoor men zou willen controleren zijn echter wél tijdsafhankelijk. In het voorbeeld waarbij men geïnteresseerd is in het effect van politiemacht op criminaliteit, zouden we bijvoorbeeld kunnen denken aan een verwachte gebeurtenis (een voetbalkampioenschap) of een onverwachte gebeurtenis (een moord op een politiek leider) die de gemoederen doen oplaaien. Vanuit conceptueel standpunt zal het toevoegen van controlevariabelen überhaupt niet voldoende zijn indien men vermoedt dat er een simultaneïteitsprobleem bestaat, dat wil zeggen, dat de afhankelijke of verklaarde variabele rechtstreeks ook een impact heeft op de verklarende variabele: een politiemacht heeft een impact op de criminaliteit, maar criminaliteit kan ook een impact hebben op het bepalen van de grootte van de politiemacht.

Een ander probleem bij retrospectieve analyses (vaak gebruikmakend van secundaire of administratieve databronnen) met controlevariabelen is dat men een brede waaier aan modelspecificaties gaat exploreren en eerder “per toeval” een relatie gaat vinden die er in wezen niet is. Immers, geschatte parameters in een regressiemodel hebben een zekere foutenmarge. De “geschatte” waarde van een parameter kan dus afwijken van de “eigenlijke” waarde. In de statistiek wordt daarom vaak niet enkel gekeken naar de geschatte waarde van de parameter, maar onderzoekt men ook standaard of deze waarde significant afwijkt van 0. Mits enkele technische aannames is het mogelijk om een betrouwbaarheidsinterval te berekenen. Een betrouwbaarheidsinterval van 95% is het interval waarbinnen met 95% zekerheid de “eigenlijke” waarde van de parameter zich bevindt. Veelal wordt een resultaat als significant verschillend van 0 beschouwd als de waarde “0” zich buiten dat interval bevindt. Een resultaat wordt beschouwd als insignificant als de waarde “0” binnen dit interval ligt.

Uiteraard bestaat er steeds een kans dat men onterecht oordeelt dat men geen verband kan aantonen, maar het is evenzeer mogelijk dat men een verband claimt dat er niet is. Wanneer een onderzoeker heel veel regressie-specificaties draait, is het dan ook heel goed mogelijk, zeker voor kleine datasets, dat één van de specificaties “per toeval” een significant resultaat oplevert. Het gevaar bestaat dan dat men deze ene specificatie rapporteert, en men beweert een significante relatie gevonden te hebben terwijl er in werkelijkheid helemaal geen verband bestaat. Wanneer men zonder een op voorhand gespecificeerd model in data gaat zoeken naar verbanden en patronen, zal men bijna altijd verbanden vinden. Analisten die via data mining risicoanalyses ontwerpen zijn zich hiervan bewust, en later in dit rapport zullen we ingaan op hoe men dat probleem in de praktijk aanpakt.

Zoeken naar instrumentele variabelen

Een andere manier om tot een zuivere impactmeting te komen als men vermoedt dat de resultaten vertekend zijn door niet-geobserveerde variabelen of simultaneïteit, is op zoek te gaan naar instrumentele variabelen. Een instrumentele variabele is een variabele waarvan men veronderstelt dat deze, vanuit conceptueel standpunt, geen impact kan hebben op de verklaarde variabele, maar wel op de verklarende variabele. Met behulp van speciale regressie-technieken kunnen goede instrumentele variabelen helpen om een causaal verband te identificeren. Enkele voorbeelden zullen worden gegeven in de volgende paragraaf, en meteen zal ook blijken dat het niet vanzelfsprekend is om een instrument te definiëren dat waterdicht is voor kritiek.

Uitvoeren van gerandomiseerde experimenten

Gerandomiseerde experimenten kennen recentelijk een enorme opgang in de sociale wetenschappen, en zijn reeds lange tijd ingeburgerd in de medische wetenschappen. Het basisprincipe van een gerandomiseerd experiment is als volgt. Een steekproef wordt getrokken en wordt daarna aselekt of “at random” verdeeld over een controlegroep en één of meerdere treatmentgroepen. De controlegroep wordt niet onderworpen aan een interventie of treatment, en de treatmentgroepen wel.

Gerandomiseerde experimenten hebben de laatste jaren heel wat bijgedragen aan de vooruitgang van kennis in de sociale wetenschappen. Van groot belang is echter het bewaken van interne validiteit, namelijk dat de groepen inderdaad “at random” zijn ingedeeld en dat er geen overstap mogelijk is van de ene naar de andere groep. Dit is niet altijd eenvoudig of mogelijk, omdat individuen uit de ene groep, indien zij kennis nemen van het experiment, kunnen proberen over te schakelen naar een andere groep. Bovendien kunnen er ook ethische kwesties parten spelen. Hoewel een gerandomiseerd experiment een enorm krachtig instrument is om impact vast te stellen, is het dus niet in elke context inzetbaar.

Een programma of regime in de opstartfase biedt vaak opportuniteiten voor een gerandomiseerd experiment. In een beginfase start men met een beperkt pilootproject, en gedurende de tijd wordt het project langzaam uitgebreid naar de volledige doelgroep.

Gerandomiseerde experimenten hebben enkele grote troeven tegenover andere methodes zoals het gebruik van controlevariabelen of instrumentele variabelen:

- Indien gerandomiseerde experimenten goed en transparant worden uitgevoerd en indien de procedures nauwgezet worden gedocumenteerd, kan met veel meer overtuigingskracht een causaal verband worden geïdentificeerd. Het best gebeurt de implementatie van experimenten wel met een theoretisch model in gedachten. Anders blijft de oefening beperkt tot de evaluatie van een bepaald programma, en leert ons dan weinig over de onderliggende mechanismen waardoor het voor beleidsmakers moeilijk wordt om praktische besluiten te trekken;
- Bij gerandomiseerde experimenten is de kans kleiner op het verkeerd gebruik van data mining. Zoals eerder aangehaald, zijn bij niet-experimentele, retrospectieve analyses vaak meerdere modelspecificaties mogelijk, en men dreigt dan te zoeken naar een specificatie die de intuïtie bevestigt. Bij een gerandomiseerd experiment, waar de opzet en het model van tevoren in detail worden beschreven, is dit gevaar minder aanwezig.

Gebruik maken van natuurlijke experimenten

In sommige gevallen verleent de onderzoekscontext ons een “natuurlijk experiment”. De onderzoekers en beleidsmakers hoeven dan niet meer zelf subjecten willekeurig (“at random”) toe te wijzen aan controle- en treatmentgroepen: deze verdeling is reeds gebeurd door de heersende omstandigheden. Een belangrijke bron van

natuurlijke experimenten zijn regressiediscontinuïteiten. Dat wil zeggen, subjecten die een bepaalde drempelwaarde overschrijden vallen in een ander regime. Indien deze drempelwaarde niet gemanipuleerd kan worden, evenmin als de positie ten opzichte van de drempelwaarde, zijn de subjecten net onder en net boven de drempelwaarde erg gelijkend, wat helpt om de zuivere impact van het desbetreffende regime te evalueren. Drempelwaardes die vaak terugkomen in de literatuur zijn een bepaalde leeftijd, een geboortjaar of een tijdstip. Bij een groot aantal observaties kan men regressiediscontinuïteiten veel beter benutten dan bij kleinere datasets. Immers, er bevinden zich bij grote datasets over het algemeen voldoende observaties rond de drempelwaarde om een analyse te kunnen uitvoeren. Bij kleine datasets zal men het venster rond de drempelwaarde moeten vergroten, waardoor de twee groepen minder vergelijkbaar worden en er meer assumpties moeten worden gemaakt.

Difference-in-difference analyse

Deze methode bestaat erin om een “verschil in de verschillen” te meten, en is in sommige gevallen aangewezen wanneer men beschikt over paneeldata en wanneer de hoofdverklarende variabele een dummy-variabele is (zoals deelname aan een programma, de inwerkingtreding van een wetgeving of beleidsmaatregel etc.). Het principe is als volgt. Stel dat men twee groepen subjecten observeert (Groep A en Groep B), beide op twee tijdstippen, T_1 en T_2 . Groep A wordt vanaf tijdstip T_1 onderworpen aan een bepaald regime of programma, en de andere groep (Groep B) niet. Graag wil men nu weten of het programma een causale impact heeft op de variabele Y . Indien men enkel cross-sectionele data heeft voor de periode T_2 dient men een aantal sterke assumpties te maken, namelijk dat het gemiddelde niveau van Y in periode T_1 niet verschilt tussen beide groepen. Het verschil in de variabele Y tussen de twee groepen op tijdstip T_2 zou dan volledig toe te schrijven zijn aan het programma. Men kan eventueel enkele controlevariabelen opnemen in het model, maar de assumpties blijven toch zeer sterk. Als men echter beschikt over paneeldata waarbij men een meting heeft aan de start van het programma, kan men een causaal verband schatten gestoeld op veel mildere assumpties. Het populatiemodel ziet eruit als volgt:

$$Y = \alpha + \beta_1 D + \alpha_2 * T_2 + \alpha_3 * D * T_2 + \varepsilon$$

D is een dummy-variabele die de waarde 1 aanneemt wanneer men tot groep A behoort, en T_2 is een dummy die de waarde 1 aanneemt wanneer de observatie in de tweede tijdsperiode werd geregistreerd. α is een constante term. β_1 is een coëfficiënt bij D en geeft aan hoeveel de waarde van Y op tijdstip 1 in groep A verschilt van die van groep B. β_2 is een coëfficiënt bij T_2 en geeft aan hoeveel de waarde Y gemiddeld verandert tussen de twee tijdsperiodes voor Groep 2, de groep

die niet deelneemt aan het programma. β_3 is de coëfficiënt bij de interactieterm $D*T_2$, en deze coëfficiënt geeft aan hoeveel de waarde Y meer groeit tussen periode 1 en periode 2 voor groep A dan voor groep B. Met andere woorden, β_3 meet het verschil tussen de twee groepen in het verschil van Y tussen periode 2 en periode 1.

Echter, wanneer we het populatiemodel schatten met een regressiemodel, moeten we ook bij dit soort modellen assumpties maken en erop vertrouwen dat de assumpties correct zijn opdat $\hat{\beta}_3$ convergeert naar β_3 . Een belangrijke assumptie houdt in dat de groei in groep A even groot zou geweest zijn als de groei van groep B mocht groep A niet deelgenomen hebben aan het programma. Het is niet ondenkbeeldig dat deze assumptie in vele toepassingen geschonden wordt, omdat de groepen kunnen verschillen op grond van (niet-geobserveerde) karakteristieken.

Veronderstel bijv. dat de afhankelijke variabele Y het loon voorstelt, en dat het programma een scholingsprogramma behelst. Aan het scholingsprogramma zullen vooral mensen deelnemen met een hoge motivatie en veel energie, kenmerken die niet (voldoende) geobserveerd zijn in de data. Deze mensen hebben vaak een hoger loonniveau dan de groep die minder gemotiveerd is, of de groep die meer belang hecht aan andere aspecten in het leven dan werk. Het difference-in-difference model kan rekening houden met deze initiële loonverschillen. Echter, het is weinig waarschijnlijk dat de loongroei van de deelnemers even groot was geweest als die van de niet-deelnemers, mocht het programma niet bestaan hebben. Immers, mensen die van hun werk een hoge prioriteit maken, kennen vaak een steiler loonprofiel over de tijd.

Er zijn verschillende manieren om de assumpties nodig bij het schatten van een difference-in-difference model geloofwaardiger te maken. Indiende data en context het toelaten, kan men kijken naar de trends tussen de twee groepen voor de start van het programma. Men kan de techniek ook combineren met data verzameld aan de hand van een natuurlijk of gerandomiseerd experiment. Later in dit hoofdstuk wordt nog een toepassing besproken uit de literatuur. Voor uitgebreidere discussies over het difference-in-difference model, de valkuilen en strategieën om resultaten geloofwaardiger te maken, verwijzen we naar Cameron en Trivedi (2005, hoofdstuk 22.6), Angrist en Pischke (2008, hoofdstuk 5.2) en Duflo et al. (2003).

3.7 Voorbeelden uit de literatuur

3.7.1 Controlevariabelen (Locke et al., 2007)

Doel

De onderzoekers willen in deze case study nagaan:

a) Welke leveranciers van Nike het meest handelen overeenkomstig de code van goede praktijken opgesteld door Nike zelf.

b) Of een verhoogde monitoring inderdaad helpt om het normconform handelen te verhogen.

De auteurs gebruiken paneeldata om hiervan een eerste idee te krijgen.

Context

In de jaren zeventig had de schoenfabrikant Nike vooral goede contacten met Japanse producenten. Door stijgende kosten echter, besloot het bedrijf geografisch te diversifiëren, van Japan en de Verenigde Staten naar overwegend Korea en Taiwan. Verdere geografische diversificatie vond plaats toen ook in deze landen de kosten begonnen te stijgen, en tegen 2004 werden Nike-producten vervaardigd in meer dan 800 bedrijven verspreid over 58 landen. Nike blijft vooral een producent van schoenen, maar heeft ook textielproducten in zijn gamma.

De relatie met zijn leveranciers is vrij heterogeen. In sommige gevallen (voor de textielproducten) gaat het om kortetermijncontracten met bedrijven die ook produceren voor (vaak concurrerende) firma's. Met andere bedrijven (in de schoeiselbranche) bestaan langetermijnrelaties.

Nike kwam echter tijdens deze geografische expansie vaak in opspraak in de media, omdat bij vele van deze toeleveranciers arbeidsvoorwaarden en mensenrechten werden geschonden. In 1992 stelde Nike daarom een eerste gedragscode op (die ondertussen reeds is bijgewerkt) voor zijn toeleveranciers, en elke toeleverancier was verplicht deze te ondertekenen en te verspreiden in het bedrijf. Om deze gedragscode af te dwingen:

- a) werden trainingssessies georganiseerd voor leveranciers,
- b) werden een negentigtal inspecteurs opgeleid, om de leveranciers te monitoren,
- c) al het personeel van toeleveranciers bevoegd voor productie of correct naleven van de regelgeving kregen incentive training op vlak van arbeidsvoorwaarden, veiligheid, beleid rond cross-cultureel bewustzijn etc.

Er zijn drie types van audits:

- a) Een basisaudit inzake milieu, veiligheid en gezondheid,
- b) periodieke inspecties uitgevoerd door de Fair Labor Association, en
- c) een diepgaande audit inzake management en arbeidsvoorwaarden, die M-audit werd genoemd. Deze neemt ca. 48 uur in beslag en wordt uitgevoerd door de Nike-auditeurs. deze audit wordt op voorhand aangekondigd en duurt 48 uur, en is dus verspreid over verschillende dagen. De beoordeling wordt samengevat in een conformiteitsscore van 0 tot 100, waarbij 100 perfecte normconformiteit aanduidt.

De afhankelijke variabele in de verschillende analyses is de conformiteitsscore van 0 tot 100. Men focust zich dus op de meest uitgebreide van de drie types audits, de M-audit. Verschillende leveranciers hebben reeds meerdere malen zo'n M-audit gehad, dus men beschikt over paneeldata.

Resultaten

De M-score bedraagt gemiddeld om en bij de 65 punten, maar er is een enorme variatie tussen leveranciers. Via een regressieanalyse stelt men vast dat deze variatie geassocieerd kan worden met volgende karakteristieken:

- De regio;
- Het aantal werknemers in de vestiging is negatief gecorreleerd met de conformiteitscore;
- De score is lager in bedrijven die eigendom zijn van binnenlandse eigenaars dan bedrijven die in bezit zijn van buitenlanders;
- De frequentie van bezoeken van Nike-personeel (niet inspecteurs) aan de bedrijven, alsook het al dan niet gecategoriseerd zijn van een strategische partner, zijn positief gerelateerd aan de conformiteitscore. De tijdsperiode sinds samenwerking met Nike, alsook het percentage van de productiecapaciteit gewijd aan Nike, waren echter negatief gecorreleerd met de score. Dit lijkt niet erg intuïtief, maar uiteraard zijn er allerlei verschillende mechanismen die deze correlaties kunnen verklaren. Bijvoorbeeld, een bedrijf waarmee er een langere relatie bestaat is vaak ook ouder dan de nieuwe partners, en oude bedrijven zijn vaak minder goed in orde dan nieuwe.

Eén van de belangrijkste resultaten luidt dat bedrijven die meerdere keren een M-audit kregen de tweede en derde keer beter in orde waren dan de eerste keer, wat suggereert dat in deze context monitoring zijn vruchten heeft afgeworpen.

3.7.2 Controlevariabelen en instrumenten (Levitt, 1997)

Doelstelling

De auteur wil het causaal verband identificeren tussen politiemacht en criminaliteit. Hiertoe gebruikt de auteur controlevariabelen en een instrumentele variabele.

Context

De auteur stelt vast dat in tijden van gemeente- en provincieraadsverkiezingen in de Verenigde Staten, de politiemacht zich uitbreidt. In tijden van verkiezingen neemt de politiemacht in grote steden toe met ongeveer 2% per jaar, terwijl deze groei in niet-verkiezingsjaren onbestaande is. Omdat een verkiezingsjaar aanleiding geeft tot een grotere politiemacht, en omdat de auteur ervan uitgaat dat het verkiezingsjaar geen determinant is van het plegen van criminele feiten, gebruikt deze een indicator voor verkiezingsjaar als instrumentele variabele. Uiteraard kan dit instrument nog heel wat andere dingen oppikken die wél een invloed hebben op criminaliteit. Immers, in een verkiezingsjaar kan er nog heel wat anders veranderen dan de grootte van de politiemacht. De auteur gebruikt daarom ook controlevariabelen zoals de werkloosheidsgraad en uitgaven voor onderwijs en welvaartsprogramma's. De auteur construeert een paneeldataset voor 59 Amerikaanse steden en observeert deze voor de periode 1970 tot 1992. Tijdens de analyses worden verschillende indicatoren voor criminaliteit gebruikt als afhankelijke variabele.

Resultaat

De auteur vindt een positieve correlatie tussen het aantal delicten en de grootte van de politiemacht. Echter, wanneer in de analyses controlevariabelen worden toegevoegd aan het model en de instrumentele variabele wordt gebruikt, krijgt de geschatte coëfficiënt een ander teken. De auteur besluit dan ook dat er een negatief oorzakelijk verband bestaat tussen de grootte van een politiemacht en het aantal gepleegde criminele feiten.

Deze studie is echter niet gespaard gebleven van kritiek. Niet enkel stonden onderzoekers vaak sceptisch tegenover het instrument omwille van conceptuele gronden, maar later bleek ook dat er een berekeningsfout in het model was geslopen die leidde tot de negatieve coëfficiënt: bij correctie van de fout verdwijnt de evidentie voor een negatieve causale relatie.

3.7.3 Gerandomiseerd experiment (Olken, 2007)

Doel

De auteur wil via een casestudy voor een ontwikkelingsland (Indonesië) meer inzicht verwerven in welke inspectiemethodes het meest effectief zijn om corruptie terug te dringen. Enerzijds zijn er overheidsinspecties, maar het is niet ondenkbeeldig dat ambtenaren in deze landen zelf corrupt zijn en dat de verduistering van fondsen niet vermindert maar enkel de verdeling ervan gaat veranderen. Een andere optie is om de burgerbevolking via participatie in vergaderingen in te schakelen om goed beheer van middelen te monitoren, maar dit vraagt vaak een inspanning van de burgers die ze niet altijd kunnen en willen opbrengen omdat de directe baten van een individuele inspanning voor henzelf vrij klein zijn.

De auteur zet een gerandomiseerd experiment op in een wegebouwproject om na te gaan welke strategie de financiële lekkage kan tegengaan.

Context

Het experiment omvat 608 dorpen in Indonesië, die bij de start van het experiment aan het begin van een wegebouwproject staan. Er werden drie treatmentgroepen gevormd, en de interventies waren als volgt:

1. Een groep van dorpen werd steevast onderworpen aan een audit door een hogere overheid, waardoor de kans op audit steeg van 4% in de controlegroep naar 100%. In principe kunnen vastgestelde onregelmatigheden leiden tot strafrechtelijke vervolging, maar dat gebeurt zelden. Belangrijker echter is dat de resultaten door de auditoren worden voorgelezen in een publieke dorpsvergadering, en indien er fraude wordt vastgesteld kan dat dus leiden tot ernstige sociale sancties.
2. Er werden in de tweede treatmentgroep van dorpen honderden uitnodigingen verstuurd om mensen aan te moedigen deel te nemen aan vergaderingen waar projectverantwoordelijken rekenschap moesten geven over het project? dit om de burgerparticipatie bij het monitoren te verhogen.
3. In een derde groep van treatmentdorpen gingen deze uitnodigingen vergezeld met een evaluatieformulier, dat men anoniem kon invullen en in een verzegelde box moest droppen, blanco of ingevuld. Dit diende te gebeuren enkele dagen voor de

start van de vergadering, zodat de resultaten van de formulieren konden worden voorgelezen op de vergadering.

De verschillende interventies werden aangekondigd in de dorpen, nadat het project was toegekend maar vooraleer het van start was gegaan.

Een andere uitdaging is het zoeken naar een goede maatstaf van corruptie. Na het project werden onafhankelijke ingenieurs op pad gestuurd om per dorp proefboringen uit te voeren en de samenstelling van het wegdek te achterhalen. Lokale leveranciers werden geïnterviewd om kosten van materiaal te kennen, en de plaatselijke bevolking werd geïnterviewd om de gangbare lonen te achterhalen. Op die manier kon men een gedetailleerde schatting maken van de kosten van het project. Dit onafhankelijke veldonderzoek werd niet gecommuniceerd naar de auditeurs van de overheid.

Het verschil tussen de opgegeven kosten van de projectverantwoordelijken en de kostenraming aan de hand van veldonderzoek na het beëindigen van het wegenbouwproject vormt dan de basis voor de berekening van het percentage lekkage. Uiteraard werd ook rekening gehouden met natuurlijk lekkage, d.i., er kunnen bij de aanleg van een weg altijd wat grondstoffen verloren gaan.

Resultaten

- De audits door de overheid waren het meest effectief om lekkage terug te dringen, namelijk met 8 procentpunt ten opzichte van de controlegroep.
- Onregelmatigheden vastgesteld door auditeurs waren sterk gecorreleerd met het percentage lekkage (gebaseerd op de schattingen van de onafhankelijke ingenieurs). Dit suggereert dat, in een omgeving waar corruptie erg wijdverspreid is, het werk van de auditeurs toch niet geheel onbetrouwbaar is.
- Het uitsturen van uitnodigingen werkte inderdaad om de deelname aan de publieke vergaderingen te doen toenemen. Over het algemeen was het effect op lekkage van een verhoogde participatie van de bevolking in het monitoringproces erg gering. Toch dienen enkele nuances gemaakt te worden. Terwijl monitoring door het publiek vaak niet leidde tot detectie van fraude inzake materiaal, was monitoring door het publiek wel effectief om fraude met arbeids- en loonkosten vast te stellen.
- De mogelijkheid om de burgers anoniem opmerkingen te laten maken op formulieren had enkel een effect op het reduceren van corruptie indien deze formulieren uitsluitend via dorpsscholen werden verspreid. Anders hadden de verantwoordelijke ambtenaren immers de mogelijkheid om ervoor te zorgen dat de formulieren in handen kwamen van een elite, hun supporters.
- Hoewel een verhoogde kans op audit door hogere instanties zorgde voor een reductie in lekkage, merkte de auteur wel op dat in deze treatment groep meer van de projectgerelateerde jobs gaan naar familieleden van projectverantwoordelijken. De ene vorm van corruptie wordt dus in zekere mate gesubstitueerd door de andere.

3.7.4 Gerandomiseerde experiment (Fellner et al., 2012)

Doel

De auteurs hebben tot doel empirisch te onderzoeken welke mechanismen leiden tot het nakomen van belastingverplichtingen. Als case study kozen zij het kijk- en luistergeld in Oostenrijk, en onderzochten de factoren die leiden tot correcte betaling van deze belasting aan de hand van een gerandomiseerd experiment.

Context

Het kijk- en luistergeld in Oostenrijk is een aanzienlijk bedrag, en varieerde in 2005 tussen € 206 en € 263 op jaarbasis. De hoogte van het kijk- en luistergeld is onafhankelijk van het aantal gezinsleden, maar varieert tussen deelstaten. 94% van de gezinnen betaalt kijk- en luistergeld, en schattingen geven aan dat 99% van de gezinnen een televisie of radio hebben. Het niet-betalen van de belasting is dus een groot knipperlicht voor non-conform handelen, omdat naar schatting vier vijfde van deze gezinnen in de fout gaan. Op het ontduiken van kijk- en luistergeld staan echter strenge straffen. Wanneer bij een audit wordt vastgesteld dat een gezin geen kijk- en luistergeld betaalt maar wel een radio of televisie heeft, kan men het kijk- en luistergeld vorderen met terugwerkende kracht (tot verschillende maanden voor de audit) en kan men daarenboven een boete opleggen tot € 2.180.

De (vereenvoudigde) opzet van het experiment

De overheid wilde de niet-betalende gezinnen via een mailing sensibiliseren, en de onderzoekers kregen de mogelijkheid te experimenteren met de brief en de inhoud ervan. Alle 50.000 gezinnen die geen kijk- en luistergeld betaalden werden "at random" opgedeeld in vijf verschillende groepen:

- Vijf procent van de observaties behelzen de controlegroep. Deze gezinnen kregen geen brief;
- Treatment groep 1 kreeg een standaardbrief van de overheid;
- De tweede treatmentgroep kreeg een brief die expliciet herinnerde aan de strenge straffen die staan op het ontduiken van kijk- en luistergeld;
- Een derde treatmentgroep kreeg een brief die beroep deed op het moreel appel: het is niet eerlijk tegenover de belastingbetalende medeburger;
- Een vierde en laatste treatmentgroep kreeg een brief met sociale informatie, namelijk het hoge percentage van gezinnen die kijk- en luistergeld betalen.

De data hebben de structuur van een cross-sectionele dataset. De afhankelijke variabele is een dummy (of 0/1-variabele) die aangeeft of een gezin 50 dagen na de mailing belasting heeft betaald. De onafhankelijke variabelen zijn indicatoren die aangeven tot welke groep het gezin behoort.

Resultaten

Het effect van een brief is immens. Na 50 dagen is de kans dat een gezin uit treatment groep 1 kijk- en luistergeld betaald heeft tienmaal hoger dan in de controlegroep. Onderlinge vergelijking van treatmentgroepen leert dat gezinnen gevoelig zijn voor de dreiging met sancties. In treatment groep 2 had zo'n tien procent meer gezinnen kijk- en luistergeld betaald dan in treatmentgroep 1.

Verrassend genoeg heeft beroep doen op het moreel appel geen of eerder een negatief effect. Beroep doen op iemands moreel appel kan immers de indruk geven dat de staat niet bij machte is om de regels af te dwingen. Uiteraard mag men een averechts effect van moreel appel niet veralgemenen tot de gehele populatie. Immers, de steekproef is slechts representatief voor een zeer klein gedeelte van de bevolking, namelijk dat gedeelte dat geen kijk- en luistergeld betaalt, en het is mogelijk dat deze groep gemiddeld gezien asocialer is dan de rest van de bevolking.

3.7.5 Natuurlijk experiment

Voorbeeld 1 (Drago et al., 2009)

Doel

De auteurs hebben tot doel empirisch te onderzoeken of er een causaal verband bestaat tussen de zwaarte van de straf (gevangenisstraf) die men boven het hoofd hangt en de kans op plegen van delicten. Met andere woorden: Hebben zware sancties een afschrikkingseffect?

Een causale analyse in deze materie is niet voor de hand liggend. Uiteraard varieert de zwaarte van sancties voor eenzelfde misdrijf van land tot land, maar een analyse over de landen heen is problematisch, omdat landen die zwaardere sancties opleggen misschien reageren op een hoge graad van non-conform handelen. Men kan ook een vergelijking maken van individuen binnen hetzelfde land, die verschillende straffen boven het hoofd hangen voor hetzelfde feit (bv. omwille van recidivisme). Zo zijn er in de Verenigde Staten een aantal lichtere misdrijven (bijvoorbeeld alcoholgebruik onder 21 jaar) die na driemaal leiden tot een gevangenisstraf. Men zou er dus voor kunnen opteren na te gaan of de kans op een tweede inbreuk na de eerste inbreuk groter is dan de kans op een derde inbreuk na de tweede inbreuk. Het is echter niet vanzelfsprekend dat individuen die op een bepaald tijdstip twee inbreuken hebben gepleegd vergelijkbaar zijn met individuen die op datzelfde tijdstip slechts één inbreuk hebben gepleegd. De auteurs maken gebruik van een natuurlijk experiment in Italië, veroorzaakt door een wetswijziging.

Context

In Italië werd in 2006 een collectief gratiedecreet van kracht. Dit decreet hield in dat iedere gevangene die vóór 2 mei 2006 een straf kreeg, kon genieten van een strafvermindering van maximum 36 maanden. Het kwijtgescholden deel van de straf werd dan omgezet in een voorwaardelijke straf. Deze voorwaardelijke straf moest uitgezeten worden indien men binnen de vijf jaar na vrijlating opnieuw een misdrijf pleegt.

Door het decreet kwamen op 1 augustus 2006 40% van de gevangenen vrij. De voorwaardelijke straffen van deze vervroegd vrijgelaten gevangenen varieerden tussen de 1 en 36 maanden. Als we dan individuen vergelijken die dezelfde initiële straf kregen, of als we met andere woorden in onze analyse controleren voor de initiële straf, wordt de hoogte van de voorwaardelijke straf enkel bepaald door de datum van opsluiting, en het lijkt erg aannemelijk dat deze variatie in voorwaardelijke straf geen verband houdt met persoonskenmerken. Dit vermoeden wordt nog

versterkt wanneer de auteurs geen correlatie vinden tussen de lengte van de voorwaardelijke straf enerzijds, en in de data geobserveerde socioeconomische karakteristieken van de vrijgelaten gevangenen anderzijds.

De auteurs construeren een enkelvoudige cross-sectionele dataset op individueel niveau. De afhankelijke variabele is een indicator voor recidivisme, en de onafhankelijke variabele meet de lengte van de voorwaardelijke straf, en een constant terugkerende controlevariabele is de initiële lengte van de gevangenisstraf.

Resultaat

De studie vindt dat een hogere voorwaardelijke straf leidt tot lagere kans op recidivisme. Hoewel de auteurs zeer creatief te werk zijn gegaan, is deze studie niet helemaal immuun voor kritiek. Immers, de vrijgelaten gevangenen met eenzelfde initiële straf verschillen wél op een belangrijk vlak. Zij met een kortere voorwaardelijke straf hebben immers langer in de gevangenis gezeten, en een langdurige gevangenisstraf kan dan weer leiden tot vervreemding van de maatschappij.

Voorbeeld 2 (Di Tella en Schargrodsky, 2004)

Doel

De auteurs willen empirisch onderzoeken of meer blauw op straat leidt tot minder criminaliteit. Zoals eerder aangegeven wordt er vaak een positieve correlatie gevonden tussen de grootte van de politiemacht en criminaliteit, maar er bestaat een enorm causaliteitsprobleem. De auteurs vonden een natuurlijk experiment in Argentinië, namelijk een bomaanslag in een Joods centrum en daaropvolgende veiligheidsmaatregelen.

Context

Op 18 juli 1994 werd in de hoofdstad van Argentinië, Buenos Aires, een aanslag gepleegd op het Joodse centrum. Een week later besliste de overheid om alle religieuze centra in het land (meer dan 270 in aantal) constante politiebewaking te bieden. Men gaat ervan uit dat de geografische distributie van deze gebouwen exogeen is in een criminaliteitsregressie, dus de verdeling van de gebouwen over wijken hangt niet samen met andere factoren die criminaliteit determineren. Door deze verhoogde beveiliging van religieuze gebouwen hebben bepaalde wijken meer politiebescherming dan andere.

De auteurs analyseren data van drie niet aan elkaar grenzende buurten in Buenos Aires. Deze buurten bestaan dan weer uit wijken, de eenheid van observatie in de data. De afhankelijke variabele is het maandelijks aantal aangegeven diefstallen van motorvoertuigen. Individuen die het slachtoffer worden van diefstal van een motorvoertuig hebben twee goede redenen om dit aan te geven:

- Ten eerste is aangifte noodzakelijk om een vergoeding te kunnen ontvangen van de verzekering;
- Ten tweede worden gestolen voertuigen vaak gebruikt voor andere misdrijven, en aangifte van diefstal voorkomt dat de eigenaar in beschuldiging wordt gesteld;

- De afhankelijke variabele wordt geobserveerd op maandelijkse basis tijdens de periode 1 april 1994 tot 31 december 1994: men beschikt dus over paneeldata. De onafhankelijke variabelen zijn indicatoren die aangeven of:
 - Een wijk een godsdienstig centrum bevat;
 - Een wijk aanpalend is aan een wijk met een godsdienstig centrum;
 - Een wijk twee wijken verwijderd is van een wijk met godsdienstig centrum;
 - Indicatoren voor elke maand.

Resultaten

De paneeldata bevatten data vóór en ná de aanslag. Dit heeft als voordeel dat men enkele extra checks kan uitvoeren om na te gaan of inderdaad de religieuze gebouwen willekeurig (“at random”) zijn verdeeld. Het al dan niet aanwezig zijn van een religieus gebouw in een wijk lijkt inderdaad niet te zijn gecorreleerd met socio-economische karakteristieken van die wijk, en vóór de interventie lijkt er gemiddeld gezien geen verschil te bestaan in aantal diefstallen (per duizend inwoners) in wijken met en zonder religieus gebouw. Na de aanslag echter en het verhoogde politietoezicht in wijken met religieus gebouw, observeert men een sterke reductie in diefstallen van motorvoertuigen in wijken met een religieus centrum. Het effect is echter eerder lokaal. Het effect is nauwelijks merkbaar in aanpalende wijken, en is onbestaande in niet-aanpalende wijken.

3.8 Aanbevelingen

- Bij het uitvoeren van onderzoek naar de impact van inspecties (of evaluaties) dienen ten eerste de doelen van de inspecties te zijn vastgesteld.
- Er bestaat echter geen “gouden standaard” om het oorzakelijke verband tussen het inspectieoptreden en het behalen van de doelen te meten. Dat wil zeggen dat men in elke context moet afwegen welke methode toepasbaar is en de meest zuivere resultaten oplevert.
- Bij elke voorgestelde onderzoeksopzet moet duidelijk aangegeven worden welke aannames vereist zijn, zodat beleidmakers een goed kader hebben om te beslissen of het verantwoord is om een bepaalde impactstudie (volgens een bepaalde methode) uit te voeren.
- Tevens dient men bewust te zijn van de soorten variabelen (onafhankelijk, afhankelijk, instrumenteel, controle, etc.) en de structuur van datasets bij analyses van impactmeting. Voor de onderzoeksopzet kan het veel uitmaken of bestanden micro (individueel of onderneming) of macro (geaggregeerd over personen of ondernemingen) zijn, of er sprake is van (herhaalde) cross-sectie, (pseudo-)paneel, of tijdreeksen.
- Dit geldt ook voor de hier besproken valkuilen van simultaneïteit en causaliteit, en de aanwezigheid van onobserveerbare variabelen waarvan de effecten onbekend kunnen zijn.
- Strategieën bij de analyses voor de impactmeting om met valkuilen om te gaan zijn het opnemen van controlevariabelen en instrumentele variabelen, het uitvoeren van gerandomiseerde experimenten of het gebruik maken van natuurlijke experimenten.

- Men dient zich ervan bewust te zijn dat elke beleidswijziging in potentie een natuurlijk experiment is, dat mits goed voorbereid, belangrijke onderzoeksresultaten kan geven voor de impactmeting.
- Studies die gebruik maken van gerandomiseerde experimenten hebben, mits goed uitgevoerd, vaak een hoge interne validiteit. De context is vaak zeer specifiek waardoor de externe validiteit beperkt kan zijn. Dit vereist goede kennis van de context en inkadering in de literatuur. Aanvullend meer kwalitatief onderzoek kan hierbij zeer verrijkend zijn.
- Men dient te beseffen dat de zoektocht naar impact of causale verbanden kan leiden tot woord en wederwoord, intense en zelfs bitsige discussies naargelang het onderwerp en het maatschappelijke belang hiervan. Derhalve dienen onderzoekers/opdrachtgevers zich bewust te zijn van de sterktes maar ook zwaktes van de impactstudie. Duidelijke doch genuanceerde conclusies zijn belangrijk opdat de impactstudie kan bijdragen tot een breed en sereen maatschappelijk debat.

4 Praktijkervaringen in België en Nederland

4.1 Inleiding

In dit hoofdstuk wordt aangegeven hoe andere diensten in België en Nederland omgaan met hun inspectiewerkzaamheden. In de volgende paragraaf wordt gekeken naar de optimalisatie van de risicoanalyse bij de Sociale Inspectie in België. In paragraaf 4.3 wordt gekeken naar twee voorbeelden in Nederland. Het gaat daarbij om de werkzaamheden van de Inspectieraad, als overkoepelende organisatie van de verschillende rijksinspectiediensten in Nederland, en om de Inspectie van het Ministerie van Sociale Zaken en Werkgelegenheid, die vergelijkbare werkzaamheden heeft als de Inspectie WSE in Vlaanderen, maar niettemin een veel ruimer takenpakket. In paragraaf 4.4 volgen nog enige aanbevelingen op grond van de beschreven praktijkervaringen in dit hoofdstuk.

4.2 België: Sociale Inspectie

OASIS (Organisation Anti-fraude des Services d'Inspection Sociale) is een databank die de vier federale inspecties moet helpen om fraude beter op te sporen. De data wordt aangeleverd via de Kruispuntbank Sociale Zekerheid, een instituut dat is opgericht in 1991 om datastromen van de verschillende instanties binnen het land samen te kunnen brengen en te coördineren. We geven hier een voorbeeld van hoe de Sociale Inspectie met behulp van data-analyse probeert zijn risicoanalyse te optimaliseren.

Doelstelling

De Sociale Inspectie is op zoek naar de meest optimale risicoanalyse. Zoals vermeld in hoofdstuk 2 is een risico een aggregatie van meerdere dimensies, en een van de dimensies daarvan is de kans dat een kwalijke gebeurtenis (een inbreuk) zich voordoet. Bij een goede risicoanalyse dienen dus ondernemingen met een hogere kans op inbreuk hoger op de prioriteitenlijst te staan.

Methodologie

Bij de zoektocht naar een geschikt model zou men kunnen kiezen voor een gerandomiseerd experiment, waarbij men de poel van ondernemingen willekeurig ("at random") opdeelt in verschillende groepen en elke groep onderwerpt aan een verschillend model van risicoanalyse. De groep ondernemingen waarbij de ratio van inbreuken en inspecties het hoogst is, kan dan worden beschouwd als die groep die werd onderworpen aan de meest effectieve risicoanalyse. Het opzetten van een dergelijk gerandomiseerd experiment is in de praktijk echter niet altijd eenvoudig te realiseren. De Sociale Inspectie heeft daarom gekozen voor een retrospectieve analyse.

De steekproef van de analyse betreft inspecties uit het verleden, toen er nog geen formele risicoanalyse werd gebruikt. De steekproef wordt vervolgens “at random” opgesplitst in twee groepen, *Groep A* en *Groep B*. De ondernemingen in *Groep A* worden onderworpen aan datamining. De afhankelijke variabele is een indicator die aangeeft of er in een bepaalde materie al dan niet een inbreuk werd vastgesteld. De onafhankelijke variabelen worden geput uit het eerder genoemde data-warehouse OASIS. De software van het data-warehouse probeert routinematig duizenden modellen uit, met allerhande verschillende combinaties van onafhankelijke variabelen, met getransformeerde variabelen en interactietermen. Daarna geeft het systeem een rangorde van de modellen die de kans op het plegen van een inbreuk het best verklaren.

Er schuilt in deze manier van datamining een enorm gevaar. Immers, bij een statistische schatting bestaat er telkens een kans dat men een verband vindt dat er in werkelijkheid niet is. Als men een hele reeks modellen schat en er het model met de beste performance uitpikt, is de kans des te groter dat dit model enkel “per toeval” een goed resultaat oplevert, maar dat het in werkelijkheid helemaal niet effectief is. De analisten van de Sociale Inspectie zijn zich daar van bewust en nemen daarom volgende maatregelen:

- Het model dat de data het best verklaart, wordt door de Sociale Inspectie niet noodzakelijk als “winnaar” aangeduid. Een model moet ook interpreteerbaar zijn. Een model dat strookt met het gezond verstand / met hypothesen afgeleid uit een theorie, is minder waarschijnlijk een toevalstreffer dan een ingewikkeld model met tal van interactietermen en polynomen (i.e. kwadraten en hogere machten van de verklarende variabelen). Daarenboven laat het gebruik van een interpreteerbaar model toe om aan inspecteurs te communiceren waarom een bepaalde onderneming op de lijst staat. Door een overzicht te krijgen van de kenmerken die gezamenlijk ertoe geleid hebben dat het bedrijf werd geselecteerd, kunnen de inspecteurs, dankzij hun vaak jarenlange ervaring, beter inschatten wat er loos is en doelgericht inspecteren;
- Om de kans op de selectie van een model dat “per toeval” goede resultaten oplevert verder te verkleinen, wordt het uitgekozen model getest op de tweede helft van de steekproef, *groep B*. Als het model ook in de tweede helft van de steekproef goede resultaten oplevert, mag men ervan uitgaan dat het model inderdaad kan gebruikt worden om het doelgericht opsporen van fraude te verbeteren.

4.3 Nederland: Inspectieraad en Inspectie SZW

Inspectieraad

In Nederland overkoepelt de Inspectieraad de verschillende rijksinspectiediensten. Sinds enkele jaren heeft men de intentie om kwantitatief de impact van het inspectie-optreden te meten. Daartoe werd in 2009 een meerjarenprogramma opgestart met als titel “Effecten van toezicht”, dat eind 2012 werd afgerond. De doelstellingen van het programma waren ambitieus. Men maakte zich sterk dat er binnen een

tijdsspanne van 4 jaar bij elke inspectiedienst een systematische effectmeting zou plaatsvinden, die mogelijk zou worden gemaakt door intense samenwerking en kennisdeling tussen de verschillende inspectiediensten. Eind 2010 werden de doelstellingen enigszins bijgesteld, en werd afgesproken dat elke inspectiedienst afzonderlijk verantwoordelijk werd voor de implementatie van een impactmeting en het aantonen van impact. Het meerjarenprogramma diende de inspectiediensten hiertoe te ondersteunen en te stimuleren.

Een belangrijke verwezenlijking van het meerjarenprogramma is de publicatie van "Effecten van Toezicht en Handhaving Meten: Een Handreiking", een coproductie in samenwerking met het Centrum voor Criminaliteitspreventie en Veiligheid. De inhoud van deze handreiking loopt in grote lijnen parallel met hoofdstuk 3 van dit rapport, en reikt enkele algemene kadermethodieken aan. Tijdens het meerjarenprogramma werd ook een jaarlijks congres georganiseerd waar de voortgang werd besproken, en ook werden workshops georganiseerd waarin de inhoud van de handreiking werd gedoceerd.

Ook werd tijdens het programma een quickscan ontwikkeld en uitgezet bij de verschillende inspectiediensten. Momenteel worden de laatste scans afgewerkt en misschien volgt er nog een algemene rapportage. Het doel van de quickscans is: "Het in kaart brengen van de wijze waarop binnen de rijksinspecties, die aangesloten zijn bij de Inspectieraad, wordt omgegaan". Tijdens enkele quickscans worden interviews afgenomen, en ook wordt gebruik gemaakt van zowel interne als openbare documenten (zoals jaarplannen, draaiboeken etc.). De Inspectieraad maakt geen resultaten per inspectiedienst bekend, maar tijdens een presentatie van voorlopige resultaten werden de volgende algemene bemerkingen gemaakt:

- De meeste inspectiediensten besteden aandacht aan impactmeting in de jaarplannen, en de doelen verschillen per inspectie;
- Veelal zijn echter de procedures vaag, expertise is versnipperd en vluchtig, en impactstudies worden zelden of nooit in de begroting opgenomen;
- Op directieniveau is er nauwelijks aandacht voor impactmeting.

Inspectie SZW

De tegenhanger van de Inspectie Werk en Sociale Economie in Nederland is de Inspectie Sociale Zaken en Werkgelegenheid (SZW). Deze inspectie is nog maar in voege sinds 1 januari 2012 en is ontstaan uit een fusie van de Arbeidsinspectie, de Inspectie Werk en Inkomen, en de Sociale Inlichtingen- en Opsporingsdienst van het ministerie SZW. De Inspectie SZW heeft dus een veel ruimer takenpakket dan de Inspectie WSE in Vlaanderen. De Inspectie SZW dient de naleving van wetgeving op volgende domeinen te controleren:

- Arbeidsomstandigheden en arbeidstijden;
- Risico's zware ongevallen, en risico-inventarisatie en -evaluatie;
- Verder dient de inspectie ook grootschalige sociale zekerheidsfraude op te sporen alsook arbeidsuitbuiting en mensensmokkel;

- Tot slot detecteert zij nieuwe trends in het domein van Sociale Zekerheid en werkgelegenheid en deelt deze mee aan de betrokken partijen in het jaarplan.

Ten einde deze doelen te kunnen realiseren gebruikt de Inspectie SZW het volgende instrumentarium:

- Preventieve campagnes, waarbij betrokken partijen worden gewezen op rechten en plichten;
- Inspecties;
- Sanctionering (zoals het opleggen van geldboetes of strafrechtelijke vervolging).

Nederland is wat betreft structuren een veel eenvoudiger land dan België. Desalniettemin is het opvallend hoezeer in het jaarplan 2012 van de Inspectie SZW de nadruk wordt gelegd op het belang van samenwerking. Eerst en vooral is er de internationale samenwerking, omdat de Inspectie SZW te maken krijgt met heel wat Europese instanties en wetgeving. Daarna echter wordt nationale samenwerking benadrukt, met andere rijksinspectiediensten overkoepeld door de inspectieraad. Zoals hierboven aangegeven hebben de inspectiediensten onder de auspiciën van de inspectieraad gezamenlijk gebrainstormd over effectmeting van het optreden van inspectiediensten. Ook echter werken inspectiediensten nauw samen om daadwerkelijk de effectiviteit van hun optreden te verhogen. Zo hebben de rijksinspectiediensten het concept van "inspectievakantie" ontwikkeld. Bedrijven die na inspectie blijken aan de regelgeving te voldoen, worden gedurende de volgende paar jaren met rust gelaten. De "rotte appels", daarentegen, worden de jaren daarna aan een verhoogd toezicht onderworpen. Een systeem wordt daarom momenteel geoptimaliseerd ten einde goede informatiedoorstroming mogelijk te maken. Inspectiediensten moeten kunnen zien of een bedrijf al dan niet een inbreuk heeft gepleegd op een domein waar de inspectiedienst in kwestie niet voor bevoegd is, zodat de inspectiedienst in kwestie de mogelijkheid heeft deze informatie te gebruiken als alarmsignalen.

Ook in Nederland lijkt men het echter erg moeilijk te hebben om impactstudies omtrent inspectieoptreden op te starten. Sinds het einde van de jaren tachtig zijn er nochtans al heel wat studies verschenen over impactmeting van inspectieoptreden, waarin veel aandacht werd besteed aan (algemene) methodologische aspecten van het leggen van causale verbanden tussen het optreden van de inspectie enerzijds, en normconform handelen anderzijds. Sommige van deze studies zijn ook evaluatief, dat wil zeggen ze gaan na of de implementatie van impactstudies bij inspecties reeds daadwerkelijk plaatsvindt. De algemene conclusie van studies met zo'n evaluatieve component is vaak echter dat er nog nauwelijks aan impactmeting wordt gedaan, en dat impactmeting nog niet geïntegreerd is in het bedrijfsproces (Bogaerts, 2011; Werkgroep Effectmeting Inspectie SZW, 2011).

De methodologieën zijn bekend, en reeds vele jaren worden er handreikingen uitgebracht die kadermethodieken op een educatieve manier proberen uit te leggen. Men kan zich dan terecht afvragen hoe het komt dat impactmeting zo moeizaam op gang komt. Dit komt volgens Bogaerts (2012) omdat een goede kennis van

methodologieën en kwantitatieve methodes niet het enige nodige is om een impactstudie succesvol te kunnen uitvoeren. Er zijn immers andere drempels die overwonnen moeten worden. Ten eerste kan de bedrijfscultuur tegenwerken, omdat niet iedereen is opgezet met kwantitatieve prestatiemeting. Bovendien lijken algemene kadermethodieken vaak erg eenvoudig en helder, maar er is vaak heel wat creativiteit nodig om ze toe te passen op een concrete situatie. Tot slot vergt een impactstudie middelen die niet altijd vrijgemaakt kunnen worden, en bovendien kunnen andere prioriteiten in het gedrang komen. Om bijvoorbeeld niveaumetingen van non-conform handelen te kunnen uitvoeren is het belangrijk dat men de steekproeven aselekt samenstelt (Pranger en Dagevos, 2001). Als men zich immers enkel gaat richten op de ratio van inbreuken en inspectieoptredens, krijgt men wellicht een enorme overschatting van non-conform handelen, omdat inspectiediensten via risicoanalyse de meest verdachte gevallen eruit lichten. Evenzeer hoeft een positieve trend van het aantal vastgestelde inbreuken per inspectieoptreden niet noodzakelijk te betekenen dat het non-conform handelen toeneemt over de tijd: de inspectiedienst kan ook efficiënter geworden zijn. Het spanningsveld tussen het toepassen van een gedegen geachte risicoanalyse enerzijds, en een impactevaluatie anderzijds, wordt dan ook meteen duidelijk. Het is niet altijd evident voor beleidsmakers om een deel van de winst die ze halen uit een goede risicoanalyse te herinvesteren in het aselekt uitvoeren van inspecties om zo het domein opnieuw te verkennen. Het tijdelijk loslaten van een risicoanalyse om een goede impactstudie mogelijk te maken is in Nederland dan ook een heikel punt, zo blijkt uit onze gesprekken met rechers van de inspectiediensten.

Voorlopig lijkt men zich bij de Inspectie SZW vooral te focussen op het meten van een fenomeen, en de trend van dat fenomeen over de tijd, vooraleer over te gaan naar het analyseren van een causaal verband tussen het inspectieoptreden enerzijds, en het normconform handelen anderzijds. Zo wil de regering erg inzetten op het bestrijden van illegale asbestsaneerders, en wil men dat de inspectie kan aantonen dat het fenomeen wordt teruggedrongen en wordt herleid tot nul. Het is echter niet evident om te meten hoeveel malafide asbestsaneerders er actief zijn, laat staan er zich van te vergewissen dat het fenomeen volledig verdwenen is, daar een malafide asbestsaneerder per definitie nergens geregistreerd staat. Daarom gaat men op zoek naar proxies, dus variabelen die het fenomeen niet rechtstreeks meten maar er wel mee gecorreleerd zijn. Zo bestaat er een webstek www.werkspot.nl, waar klussers en opdrachtgevers elkaar kunnen vinden. Men kan op deze webstek zijn diensten aanbieden zonder zijn echte identiteit bekend te maken, en vele malafide asbestsaneerders gebruiken deze site dan ook om te adverteren. De inspectie kan van deze site handig gebruik maken om het fenomeen in kaart te brengen: daalt het aantal aanmeldingen op werkspot.nl, dan gaat men ervan uit dat het fenomeen afneemt.

Hoewel op microniveau impactstudies nog vrij beperkt zijn in Nederland, heeft men recentelijk op macroniveau toch enkele erg markante vorderingen gemaakt. Zo is recentelijk gedetailleerd de tevredenheid van geïnspecteerden met de verschillende inspectiediensten in kaart gebracht. Het is immers belangrijk dat de geïnspecteerden de inspectie als vakkundig, behulpzaam en transparant ervaren, want men kan

verwachten dat een gerespecteerde inspectiedienst het normconform handelen bevordert. Daarom gaf de Inspectieraad het programma InternetSpiegel -- een programma van het ministerie van Binnenlandse Zaken en Koninkrijksrelaties -- de opdracht om een klanttevredenheidsonderzoek uit te voeren voor de negen rijksinspecties. Voor de enquêtering liet de arbeidsinspectie in augustus 2011 die personen benaderen die gedurende de laatste twaalf maanden als contactpersoon hadden gefungeerd bij communicatie met de inspectie (zie Inspectie SZW, 2011).

In het jaarplan 2012 werd reeds aangegeven dat de Inspectie SZW in dat jaar zou starten met een meerjarenprogramma rond effectmeting, aansluitend op het net beëindigde programma "Effecten van Toezicht" van de Inspectieraad. In het jaarplan 2013 werd hieromtrent gemeld dat de inspectiedienst samen met de andere diensten aan het bouwen is aan een 'gereedschapskist' met diverse toezichtinstrumenten. Bovendien wil de Inspectie SZW meerjarige proefprojecten implementeren en verderzetten zoals omtrent:

- impactmeting van de aanpak van malafide uitzendbureaus,
- impactmeting van de aanpak malafide asbestsaneerders, en
- impactmeting van het toezicht op grote-risico-bedrijven.

Men hoopt in de loop van 2013 interimresultaten van deze impactstudies bekend te maken.

Een klanttevredenheidsonderzoek bevat vooral vragen waarop men antwoorden kan geven op een vijfpuntschaal (zoals 1 = zeer ontevreden en 5 = zeer tevreden), maar de enquête bevatte ook ja/nee vragen en zelfs een aantal open vragen. Daarenboven peilt de enquête ook naar de kennis over de arbeidsinspectie. De enquête bestond uit verschillende modules waarvan hieronder enkele voorbeelden worden gegeven.

- Inwinnen van informatie en vragen van advies: betrouwbaarheid, duidelijkheid van gevraagd advies alsook de tevredenheid over de snelheid waarmee een vraag werd behandeld, tevredenheid over de manier waarop men informatie kon inwinnen etc.
- Melding doen: Werd een melding adequaat behandeld?
- Bezoek van de arbeidsinspectie: Heeft de inspectie het bedrijfsproces verstoord? Richtte de inspecteur zich op de belangrijkste risico's?
- Rapport of verslag: Was het rapport helder, correct?
- Maatregel of sanctie: Was de sanctie redelijk/onredelijk? Heeft de maatregel of sanctie geholpen om de kwaliteit van uw organisatie of product te verbeteren?
Bezwaar: Was u tevreden met duidelijkheid rond procedure indienen bezwaar, manier waarop bezwaar kon ingediend worden?
- Imago en vertrouwen: Hoe groot is de kans dat u zich positief uitlaat over arbeidsinspectie tegenover anderen?
- Indienen van een klacht: Bent u op de hoogte van een klachtenprocedure, bent u tevreden met de manier waarop klachten kunnen worden ingediend?

Over het algemeen waren de respondenten zeer tevreden over de arbeidsinspectie. Dit leverde de arbeidsinspectie – als eerste inspectiedienst – een “bewijs van goede dienst” op. Dit bewijs werd afgeleverd op 7 december 2012 en is geldig tot 7 december 2014. Naast dit bewijs werd ook meteen een verbeterplan opgesteld, zodat de inspectie kan werken aan de punten waarop iets minder goed werd gescoord.

In het kader van impact op macroniveau vermelden we ook nog de webstek www.zelfinspectie.nl. Op deze site kan men als bedrijfsleider op een interactieve wijze nagaan of het bedrijf voldoet aan de regels, en welke sancties boven het hoofd hangen bij een bepaalde inbreuk. Deze webstek moet het bonafide ondernemers dan ook gemakkelijker maken om zich te schikken naar een complexe wetgeving. Een dergelijke webstek zou ook in Vlaanderen tegemoet kunnen komen aan de verzuchtingen van het middenveld inzake complexe en onduidelijke wetgeving

4.4 Aanbevelingen

- Om de risicoanalyse van de IWSE te optimaliseren biedt de data-analyse van de Sociale Inspectie in België een interessante werkwijze, waarbij gebruik wordt gemaakt van de OASIS databank en de Kruispuntbank Sociale Zekerheid.
- Om de impact van de inspectiediensten in Nederland te kunnen meten zijn er tussen de Inspectieraad en de inspectiediensten afspraken gemaakt over het opstellen van meerjarenprogramma's, waarin kennisdeling en systematisch impactmeting geprioriteerd werden. Tevens werd er op initiatief van de Inspectieraad een handreiking voor impactmeting samengesteld.
- De Nederlandse Inspectie Sociale Zaken en Werkgelegenheid (SZW) is vergelijkbaar met de IWSE in Vlaanderen, maar heeft een breder takenpakket, en probeert door preventieve campagnes, inspecties en sanctionering het non-conform handelen terug te dringen.
- Hoewel ook de Nederlandse inspectie grote moeite heeft om impactstudies uit te voeren, is er nu toch een aantal proefprojecten op dit terrein gelanceerd. Ook zijn er andere interessante initiatieven ontplooid. Zo is het concept “inspectievakantie” ingevuld, bestaat er een webstek waar klussers en opdrachtgevers elkaar kunnen vinden om zicht te krijgen op de omvang van malafide ondernemingen zonder deze overigens te kunnen opsporen, worden er tevredenheidsonderzoeken gehouden onder contactpersonen van de inspectie, en kunnen ondernemers zichzelf inspecteren zodat ze zich gemakkelijker kunnen schikken naar de complexe wetgeving.

5 Handreikingen

5.1 Inleiding

Dit laatste hoofdstuk van het rapport omvat handreikingen, dat wil zeggen enkele concrete procedures om in een aantal materies van de Inspectie Werk en Sociale Economie impactstudies in de stijgbeugels te zetten. Sommige van deze impactstudies kunnen gevolgen inhouden voor het debat over het handelen van de Inspectie Werk en Sociale Economie. Ook kunnen de uitkomsten van impactstudies erg nuttig blijken in de dialoog met andere inspectiediensten, andere partners en in het bijzonder de overheid. Immers, de resultaten kunnen een leidraad zijn bij het herzien van de wetgeving. Wanneer het optreden van de inspectie geen invloed lijkt te hebben op het normconform handelen van de geïnspecteerde in een bepaalde materie, kan men opperen dat de wetgeving dient te worden aangepast teneinde de inspectie meer munitie te geven om zich van haar taak te kwijten.

In dit hoofdstuk wordt in de volgende paragraaf ingegaan op de toegevoegde waarde van administratieve databestanden om impactstudies mee uit te voeren. Vervolgens wordt in paragraaf 5.3 uitvoerig ingegaan op hoe de IWSE zijn eigen rapportagebestanden kan creëren ten behoeve van impactanalyses. Er wordt daarbij uitgebreid stilgestaan bij twee voorbeelden van het genereren van kwantitatieve rapportagebestanden voor impactstudies, namelijk bij de regeling tewerkstellingspremie 50+ en de regeling professionele integratie van personen met een arbeidshandicap (PIPA). Het ROA zou bij deze voorbeelden verder invulling kunnen geven aan de uitvoering van de impactstudies. In paragraaf 5.4 wordt nog verder ingegaan op enkele aanvullende concrete handreikingen bij het uitvoeren van impactstudies op microniveau. Tot slot wordt nog ingegaan op enkele concrete handreikingen op macroniveau.

5.2 Administratieve bestanden als basisbouwstenen voor impactstudies

De rapportages op ondernemingsniveau dienen te kunnen worden gefuseerd met andere rijke databronnen, teneinde het aantal analyse-mogelijkheden te vergroten. In het eerste deel van deze paragraaf geven we een overzicht van databronnen die IWSE momenteel kan raadplegen. Daarna vermelden we enkele andere nuttige databronnen en platvormen die in België operationeel zijn en die worden gebruikt door federale inspecties. Het is echter niet evident voor een kleine inspectie om deze allemaal te raadplegen voor velerlei redenen, en daarom komen we in het derde deel met enkele voorstellen om dit probleem op te lossen.

1. Administratieve databronnen raadpleegbaar door IWSE

In België zijn heel wat regionale en nationale databanken opgericht die toelaten informatie over sociaal verzekeren of ondernemingen op te zoeken, en bij uitbreiding de informatie te gebruiken voor statistische analyses. Sinds kort kan ook

de IWSE (beperkt) gegevens opvragen uit een aantal van hierna genoemde databanken.

- DIMONA (Déclaration Immédiate/ Onmiddellijke aangifte): de werkgever is verplicht om een indiensttreding onmiddellijk te melden in de DIMONA-attribank.

Deze databank bevat het KBO-nummer, de begin- en einddatum van de tewerkstelling, het paritair comité en adres van tewerkstelling.

- Het rijksregister.
- Personeelsbestand van de werkgevers ingeschreven bij de RSZ of RSZPPO.
- Werkgeversrepertorium: bevat informatie over werkgevers die voor het publiek toegankelijk zijn (zoals het adres van de hoofdzetel) alsook afgeschermd gegevens (met informatie over het sociaal secretariaat waarop men beroep doet, de historiek van werkgeverscategorien etc.).
- LIMOSA (Landenoverschrijdend Informatiesysteem ten behoeve van MigratieOnderzoek) bij de Sociale Administratie: een databank waarin de buitenlandse werknemers geregistreerd staan.

2. Andere data en analysemogelijkheden ter beschikking van federale inspecties

De Inspectie WSE kan voor een bepaald bedrijf of persoon weliswaar (een beperkt aantal) gegevens opvragen die in de databanken zijn opgeslagen, maar men heeft hiermee allerminst de mogelijkheid om over te gaan tot statistische analyses.

De federale inspecties hebben wel meer mogelijkheden tot het uitvoeren van statistische analyses via het eerder vermelde data-warehouse OASIS. Naast data uit bronnen die eerder zijn genoemd bevat het data-warehouse een selectie van gegevens uit de hieronder genoemde databronnen.

- DMFA (Déclaration multifonctionnelle/ multifunctionele Aangifte): In deze databank voert de werkgever de loon- en arbeidstijdsgegevens van de werknemers in.
- Bestand van werven en onderaannemers: Via de toepassing Unieke Werfmelding op de webstek van de RSZ dienen aannemers een aantal activiteiten te melden die in deze databank worden opgenomen.
- Bestand van rekeningen en inningen RSZ: In dit bestand werkt de RSZ de rekeningen met werkgevers bij, zodat zij kan nagaan hoeveel er nog moet geïnd worden aan socialezekerheidsbijdragen.
- Werklozenbestand RVA: Dit bestand bevat informatie over de werkzoekenden die een uitkering ontvangen van de RVA.
- Driemaandelijke BTW-aangifte en BTW-listings van klanten.

De Kruispuntattribank Sociale Zekerheid (KSZ) verzamelt deze gegevens van drie verschillende instanties (RVA, RSZ en BTW-administratie) die periodiek data aanleveren voor het data-warehouse OASIS. Data uit de DMFA, DIMONA en het RVA werklozenbestand vallen onder de wetgeving van de bescherming van persoonlijke gegevens en zij worden dan ook door de kruispuntbank geanonimiseerd vooraleer ze worden doorgestuurd naar het data-warehouse. Indien inspecteurs van

de inspectiediensten voor hun opdracht meer specifieke persoonsinformatie nodig hebben uit één van deze drie databanken die vallen onder de bescherming van persoonlijke gegevens, kunnen zij hiervoor een gemotiveerd verzoekschrift richten aan de kruispunt-databank. De inspecteur stuurt een listing van geëncrypteerde Identificatienummers van Sociale Zekerheid (INSZ) door naar de Kruispunt-databank. Indien het Sectoraal Comité van de Sociale Zekerheid (SCSZ) hen hiertoe de machtiging verleent, kan de kruispunt-databank extra informatie opvragen bij een van de instanties die de data aanlevert, en de data doorsturen naar OASIS. De inspecteur krijgt de data uiteraard terug met een geëncrypteerd INSZ zodat hij geen additionele persoonsgevoelige informatie kan gaan opzoeken in andere tot diens beschikking zijnde databanken.

Ook kan andere data opgevraagd worden via de KSZ, die niet afkomstig zijn van deze drie partners. Hiervoor zal echter wel een intense procedure moeten gevolgd worden, en de SCSZ zal moeten oordelen of er voldoende aanwijzingen zijn dat de data veilig beheerd worden en inderdaad noodzakelijk zijn voor het uitvoeren van de taken van de inspectiedienst. Voor het verkrijgen van informatie over BTW-saldi (zie ook hoofdstuk 3 voor een voorbeeld over negatieve BTW-saldi als proxy-variabele voor non-conform handelen) zal de IWSE een samenwerkingsakkoord moeten sluiten met de federale overheidsdienst Financiën of dient men dit te comprimeren via de jaarrekeningen van de ondernemingen.

Verder beschikken de federale inspecties ook nog over een aantal andere databronnen en -platformen. Zo is er bijvoorbeeld GENESIS, dat in tegenstelling tot OASIS toegankelijk is voor inspecteurs (niet enkel voor een paar data-analisten) en waarin men gegevens kan opzoeken uit een aantal van de bovenvermelde databanken teneinde op een efficiënte manier een inspectie voor te bereiden. GENESIS bevat ook een module waarin korte inspectieverslagen te vinden zijn, wat de inspectiediensten helpt om acties op elkaar af te stemmen. Het GENESIS-platform heeft als doel informatie op te kunnen zoeken, maar kan niet gebruikt worden voor datamining. Eind vorig jaar is er een gelijkaardig systeem gelanceerd door de RSZ genaamd DOLSIS. Dit systeem wordt ook opengesteld voor andere instellingen dan de federale inspecties. Deze andere instanties (zoals regionale inspecties, politiediensten) kunnen toegang krijgen tot die gegevens waarvoor de SCSZ hen een machtiging verleend heeft.

Sommige inspecties beschikken ook over commerciële data van Graydon of EURODB. Deze worden slechts in uitzonderlijke omstandigheden gebruikt. Een troef van zulke data kan bijvoorbeeld zijn dat ze de mandatarissen van elke onderneming registreren, waardoor men gemakkelijker bedrijven kan opsporen die gelieerd zijn met een mandataris van een malafide onderneming.

3. Data-analyse voor de Inspectie WSE: hoe verder te werk gaan?

De vraag is nu hoe een relatief kleine inspectie als de IWSE data-analyse ten behoeve van bijvoorbeeld de impactmeting kan laten verrichten. Regionale inspecties zoals de IWSE voeren al veel inspecties uit in opdracht van instanties

zoals het Vlaams Subsidieagentschap Werk en Sociale Economie (VSAWSE), het arbeidsauditoraat en in samenwerking met de federale inspecties gecoördineerd door het centrale sturingsorgaan Sociale Inlichtingen en Opsporingsdienst (SIOD). In deze samenwerking zou door de IWSE een discussie kunnen worden gestart om samen de andere instanties na te denken over de evaluatie van de impactmeting, waarbij ook zij een bijdrage kunnen leveren aan de data-analyse.

Daarnaast voert de Inspectie WSE een heel aantal inspecties uit op eigen initiatief. Juist bij deze inspecties kan zij initiatieven nemen voor het optimaliseren van een risicoanalyse op materieniveau teneinde de effectiviteit van de inspectiedienst te vergroten. Ook kan zij hierbij de mogelijkheid creëren om zelf een kwantitatieve impactanalyse uit te voeren van haar inspectieoptreden. Indien de inspectie hierin slaagt, zal zij immers een groter gewicht kunnen aantonen in de dialoog met het arbeidsauditoraat, het VSAWSE, de federale inspecties en de overheid.

Teneinde deze evaluaties te kunnen uitvoeren kan men er niet om heen dat de inspectie nood heeft aan meer data en analysemogelijkheden. Het lijkt echter niet erg realistisch dat de Inspectie WSE, naar het voorbeeld van haar federale collega-inspectiediensten, in de toekomst zelf beschikt over een omvangrijk data-warehouse als OASIS. Zo'n structuur brengt allerlei hoge vaste kosten met zich mee, zodat de gemiddelde kost per inspectie voor een kleine dienst absurd hoog kan liggen. Zo zal de structuur van de inspectie gewijzigd moeten worden, en moet iemand verantwoordelijk worden voor het correct gebruik van de data. Men zal moeten aantonen aan het SCSZ dat de data veilig opgeslagen zijn. Bovendien zal minstens één van de medewerkers intensief omgeschoold moeten worden om het data-warehouse optimaal te kunnen gebruiken. Dit geldt ook voor andere kleine regionale inspecties.

Tijdens een verhelderend gesprek met een data-analist van een federale inspectie kwamen de twee volgende denkpistes naar voren, een eerder theoretische en een pragmatische.

- De theoretische denkpiste houdt in dat men op termijn een nieuw orgaan/een nieuwe structuur kan opzetten, waarbij de regionale inspecties de krachten bundelen om samen een (uitgebreide) versie van OASIS tot stand te brengen. De regionale inspecties zouden gezamenlijk een aantal analisten kunnen aanstellen. Misschien kan het Managementondersteuning Vlaamse Instellingen en het Inspectie-netwerk" (MOVI-netwerk) een goede opportuniteit bieden om zo'n initiatief op gang te trekken.
- Een meer pragmatische oplossing, die wellicht op kortere termijn haalbaar is, bestaat eruit om via de SIOD te negotiëren met de federale inspecties met de vraag enkele analyses uit te voeren. De Inspectie WSE zou een kwantitatief rapportagebestand (zoals verderop beschreven in dit hoofdstuk) kunnen aanleveren aan de KSZ. De KSZ zou het bestand dan kunnen doorsturen naar OASIS zodat één van de federale inspecties ermee aan de slag kan.

Kortom, het lijkt erop dat meer autonomie voor de IWSE op korte termijn vooral ligt in een nog intensievere samenwerking met de federale inspecties en andere partners. In de volgende paragraaf, waarin we enkele eerste ruwe schetsen geven van impactstudies, zal uitgegaan worden van het laatste (pragmatische) scenario waarbij de regionale inspectie samenwerkt met de federale inspecties voor het uitvoeren van de kwantitatieve analyses.

5.3 Kwantitatieve rapportagebestanden als basisbouwstenen voor impactstudies

Voor een kwantitatieve impactanalyse is het belangrijk dat de vaststellingen ook op een kwantitatieve manier worden gerapporteerd, en dat er elk jaar (of periode) door de IWSE een *kwantitatief rapportagebestand* (per materie) wordt opgeleverd. Deels gebeurt een kwantitatieve rapportage reeds via het elektronisch inspectiedossier, wat resulteert in tabellen in het jaarrapport die het aantal controles, inbreuken etc. weergeven en die zelfs vergelijkingen met cijfers van vorige jaren mogelijk maken. Deze gegevens zijn echter niet toereikend om een kwantitatieve impactanalyse op te starten. Immers, als de inspectie doelgericht is gaan inspecteren (als de risicoanalyse dus verbeterd is), en men een opwaartse trend ziet in het aantal vastgestelde inbreuken per controle, zou men tot de verkeerde conclusie kunnen komen dat het optreden van de inspectie geen impact heeft. Het is wel degelijk mogelijk dat het inspectieoptreden een impact heeft gehad op het normconform handelen, maar de gegevens zoals verzameld via het elektronisch inspectiedossier en gerapporteerd in het jaarrapport laten niet toe hier een dieper inzicht in te krijgen. We stellen daarom voor om, naast het invullen van het elektronisch inspectiedossier, de resultaten ook op een andere manier te rapporteren.

Het bijhouden van een *kwantitatieve rapportagebestand*, wat wij hieronder zullen toelichten, zal normalerwijze niet al te tijdrovend zijn voor de inspecteurs maar zal wel heel wat meer mogelijkheden bieden om beter inzicht te krijgen in de impact van Inspectie WSE op het normconform handelen. Voor de constructie van een bruikbaar *kwantitatief rapportagebestand* dienen de volgende stappen gevolgd te worden:

- De rapportages dienen te worden bijgehouden in een tabelformaat, waarbij elke rij in de tabel een observatie (bijv. inspectieoptreden) voorstelt;
- De kolommen in de tabel stellen variabelen voor die voor deze materie relevant zijn. Enkele uitgewerkte voorbeelden in de volgende sectie moeten dit duidelijk maken;
- Het rapportagebestand kan na verloop van tijd uitgroeien tot een paneeldatabestand (zie hoofdstuk 3), waarin een geïnspecteerde op verschillende tijdstippen wordt geobserveerd. De uitbouw van een paneeldatabestand kan versneld worden over de tijd, als inspecteurs tijdens het vooronderzoek gegevens van eerdere controles invoeren in het rapportagebestand. Indien men bijvoorbeeld in 2013 een inspectie wil uitvoeren bij een onderneming die in 2011 en 2012 ook reeds werd geïnspecteerd, kan men de bevindingen van 2013 registreren in het databestand alsook de bevindingen die men deed in 2011 en 2012. Deze werkwijze, die weliswaar wat meer administratief werk met zich meebrengt

in het jaar waarop het kwantitatief rapportagebestand wordt ingevoerd, laat toe om sneller aan de slag te kunnen met de impactanalyse;

- Het is ook belangrijk dat er bij elk kwantitatief rapportagebestand een handleiding en codebook (beschrijving van de variabelen) worden gevoegd, zodat de data-analist er mee aan de slag kan;
- Er bestaan velerlei types variabelen die in het databestand zouden kunnen voorkomen. In principe kunnen dit ook string variabelen zijn, dus variabelen waarvan de waardes niet-numeriek zijn maar wel bestaan uit cijfer- en lettertekens. Men kan er de voorkeur aan geven om dit soort variabelen te vermijden, omdat deze niet rechtstreeks kunnen gebruikt worden voor een kwantitatieve analyse. Het is daarom soms wenselijk om bv. namen van geïnspecteerden, ondernemingen etc. te coderen en bij te houden in numeriek variabelen. Als men er toch voor kiest te werken met string-variabelen, kunnen deze later in een statistisch programma zoals STATA of SAS gecodeerd worden, omgezet worden naar numerieke waarden. Het statistisch programma zal dan de string-waardes als waardelabels of "label values" opslaan, zodat het voor de analist eenvoudig na te gaan is waarvoor een numerieke waarde staat;
- Vaak wordt er in grote databestanden voor gezorgd dat numerieke variabelen altijd 0 of positief zijn, en niet negatief. Dit wordt gedaan omdat men dan enkele negatieve waardes kan voorbehouden voor speciale codes, die omschrijven waarom een waarde voor een variabele niet beschikbaar is. In de Swiss Household Panel vindt men voor elke variabele consequent de volgende codering terug:
 - 8 other error
 - 7 filter error
 - 3 inapplicable
 - 2 no answer
 - 1 does not know;
- Indien men toch een fenomeen wil meten dat zowel negatieve als positieve waardes kan aannemen (bv. Saldo op rekening) kan men de variabele in kwestie eenvoudig herdefiniëren. In het geval van rekeningssaldo kan men de variabelen "positief saldo" en "negatief saldo" aanmaken.

5.4 Impact op microniveau: twee concrete voorbeelden met kwantitatieve rapportagebestanden

In deze sectie geven we enkele concrete voorbeelden van impactanalyses op microniveau, dus op het niveau van de geïnspecteerde. We laten zien hoe de Inspectie WSE aan de hand van kwantitatieve rapportagebestanden reeds op relatief korte termijn kan starten met impactanalyses voor de tewerkstellingspremie 50+. De twee uitgewerkte voorbeelden hieronder moeten duidelijk maken hoe een rapportagebestand kan worden opgebouwd en gebruikt. Bij de bespreking van de materies zal de nadruk liggen op de werkwijze en de kadermethodieken eerder dan op het uitgebreid herhalen van alle bepalingen die in de betreffende wetten, decreten en regeringsbesluiten kunnen worden teruggevonden.

Voorbeeld 1: Tewerkstellingspremie 50+

Situering

Naar aanleiding van het Besluit van de Vlaamse regering van 28/04/2006 werd een tewerkstellingspremie 50+ ingevoerd teneinde niet-werkende werkzoekende 50-plussers meer kansen te geven op de arbeidsmarkt. De tewerkstellingspremie 50+ behelst een loonsubsidie voor de werkgever. Sinds de invoering van de premie werd de regelgeving enkele malen lichtjes gewijzigd, de laatste maal in januari 2013. In dit uitgewerkte voorbeeld echter, gaan we uit van de situatie voor de wijzigingen die van kracht werden sinds januari 2013. Immers, de meeste inspecties die nu nog lopen hebben betrekking op dossiers waarvan de aanvraag werd ingediend voor 1 januari 2013.

De hoogte van de premie kan oplopen tot 50% van het brutoloon, en wordt berekend aan de hand van de loonschaal waarin het brutoloon (per kwartaal) van de werknemer valt. De premie bedraagt:

- € 1.200 in de loonschaal € 2.400 - € 4.200
- € 2.100 in de loonschaal € 4.200 - € 6.000
- € 3.000 in de loonschaal € 6.000 - € 10.500
- € 45.000 in de loonschaal boven € 10.500

Naast de leeftijdsgrens van de aangeworven niet-werkende werkzoekende zijn er echter nog enkele andere belangrijke criteria waaraan voldaan moet worden opdat de werkgever recht heeft op de premie. Een aantal van deze voorwaarden staan hieronder vermeld.

- In het jaar voor de aanvraag van de premie mag de som van federale en Vlaamse subsidies niet hoger zijn dan 50% van de loonkosten van de betreffende onderneming.
- Men mag geen werknemer ontslaan teneinde een gesubsidieerde werknemer in dienst te kunnen nemen.
- De tewerkstellingspremie 50+ kan niet toegekend worden wanneer de werknemer in een periode van 6 maanden voor de aanvraag voor de onderneming heeft gewerkt.
- De werknemer moet een contract van onbepaalde duur krijgen, en de premie dient te worden terugbetaald als de werknemer minder dan 5 kwartalen in dienst bleef, tenzij de werknemer zelf ontslag nam of werd ontslagen om dwingende redenen.
- De tewerkstellingspremie 50+ kan niet gecumuleerd worden met dienstencheques.
- Voor de periode 1 oktober 2010 tot 31 december 2011 kan de tewerkstellingspremie 50+ niet gecumuleerd worden met een gelijkaardige federale winwin-premie.
- De premies worden uitbetaald door de VDAB. Via de kruispuntbank krijgt de VDAB DMFA-loongegevens doorgestuurd. Het duurt echter altijd even vooraleer deze loongegevens beschikbaar zijn, en daarom zal de VDAB

zich voor de eerste twee kwartalen baseren op het brutoloon zoals opgegeven in de aanvraag.

Opbouw rapportagebestand tewerkstellingspremie 50+

De rijen stellen verschillende observaties voor, en in dit geval is de eenheid van observatie een inspectieoptreden. Tijdens een inspectieoptreden kunnen er uiteraard meerdere dossiers worden geïnspecteerd. De kolommen stellen de variabelen voor, en hoofdingen van de kolommen hebben zeer beknopte namen, omdat dit gemakkelijk is bij het bewerken en analyseren van de data. Deze beknopte nomenclatuur maakt echter meteen duidelijk waarom het cruciaal is dat een databestand vergezeld gaat met een goed codeboek. Het codeboek voor het rapportagebestand tewerkstellingspremie 50+ is hieronder weergegeven.

- KBO: Het bedrijfsnummer van de gecontroleerde onderneming zoals terug te vinden in de Kruispuntbank van Ondernemingen.
- NACE: de NACE code (Nomenclature statistique des activités économiques dans la Communauté européenne) ook terug te vinden in de Kruispuntbank van Ondernemingen.
- jaar: Het jaar waarin de inspectie plaats vond.
- materie: de code van de materie waarover de observatie handelt. De code van 50+ = 1. Het is nuttig zo'n variabele toe te voegen aan het bestand, omdat zo, indien gewenst, de fusie van databestanden op materieniveau wordt mogelijk gemaakt.
- aant_cont: Het aantal geïnspecteerde dossiers in de betreffende onderneming.
- alg_inbreuk: Het aantal dossiers waarin één of meerdere inbreuken inzake deze materie werden vastgesteld.
- oversubs: Aantal dossiers bij dewelke in het jaar voor aanvraag voor een premie de totale som van alle federale en Vlaamse subsidies meer dan 50% bedroeg van de loonkosten. In de praktijk zal dit opgaan voor ofwel geen enkel dossier, ofwel alle geïnspecteerde dossiers, tenzij zeer recentelijk de ratio van subsidies en lonen toe- of afnam.
- brugperiode: Het aantal dossiers waarbij de gesubsidieerde werknemer reeds eerder in dienst was van het bedrijf én de arbeidsovereenkomst werd beëindigd minder dan 6 maanden voor de hertewerkstelling.
- ontslag: Bestaat er een vermoeden dat werknemers werden ontslagen om gesubsidieerde werknemers te kunnen aannemen? 0 = Neen; 1 = Ja.
- cheque: Het aantal geïnspecteerde dossiers waarbij men vaststelde dat de gesubsidieerde werknemer werd uitbetaald met dienstencheques.
- discr_bruto: Het aantal dossiers waarbij gedurende de eerste twee kwartalen te veel subsidies werden uitbetaald aan de onderneming omwille van een discrepantie tussen het brutoloon opgegeven in de aanvraag en het werkelijke brutoloon.

Een illustratie van een kwantitatief rapportagebestand volgt hieronder.

Tabel 5.1
Fictief rapportagebestand van werkgevers voor de tewerkstellingspremie 50+

KBO	NACE	jaar	materie	aant_ cont	alg_ inbreuk	over- subs	brug- periode	ontslag	cheque	discr_ bruto
100	41	2010	1	7	3	0	2	0	0	3
100	41	2011	1	9	2	0	0	0	0	2
100	41	2012	1	1	0	0	0	0	0	0

In dit rapportagebestand staat voorlopig één (fictieve) onderneming geregistreerd met KBO-nummer 100. Er zijn echter wel meerdere rijen data in het bestand, omdat het bestand data bevat van controles op drie verschillende tijdstippen, nl. 2010, 2011 en 2012. Het aantal dossiers die werden geïnspecteerd bedroegen 7 in 2010, 9 in 2011 en 1 in 2012. In 2010 werden in 3 dossiers problemen opgemerkt. In 2011 waren dit 2 dossiers en in 2012 was het geïnspecteerde dossier volledig in orde. Kijken we naar de deelaspecten, zien we dat het bedrijf nooit te veel subsidies ontving van de Vlaamse en federale overheid samen om voor een bepaald dossier niet in aanmerking te kunnen komen voor de tewerkstellingspremie 50+. In het jaar 2010 kwam het tweemaal voor dat een 50-plusser reeds een tijdje gewerkt had voor de onderneming tijdens de zes maanden voorafgaand aan de huidige aanstelling. IN geen enkel geïnspecteerd dossier werden de gesubsidieerden (deels) vergoed met dienstencheques. In alle dossiers die niet in orde waren (3 in 2010 en 2 in 2011) werd vastgesteld dat de eerste twee kwartalen van de aanstelling te veel loon werd uitbetaald, omdat het brutoloon opgegeven bij de aanvraag in een hogere schaal lag dan het werkelijk uitbetaalde loon.

Voorbeeld 2: Professionele integratie van personen met een arbeidshandicap (PIPA)

Situering

De Vlaamse overheid kent heel wat activeringsmaatregelen om personen met een Arbeidshandicap te ondersteunen in het reguliere arbeidsproces. Deze “Professionele integratie van personen met een arbeidshandicap” (PIPA) maatregelen worden geregeld in het Besluit van de Vlaamse Regering van juli 2008. De formele voorwaarden om als persoon met een handicap in aanmerking te komen voor deze steunmaatregelen, alsook de types werkgevers die er beroep kunnen op doen, staan ook duidelijk omschreven op de webstek van de VDAB.

In dit voorbeeld volgt eerst een korte beschrijving van de vele bestaande PIPA-maatregelen en de geldende voorwaarden. Daarna wordt ingegaan hoe de informatie, die tijdens inspectierondes wordt verzameld, kan worden verwerkt in rapportagebestanden die kunnen helpen de impact van het inspectieoptreden te meten.

1. De Vlaamse Ondersteuningspremie (VOP)

De Vlaamse ondersteuningspremie is, net zoals de premie 50+, een loonsubsidie ten voordele van de werkgever. In principe loopt de premie gedurende 5 jaar, waarbij de premie geleidelijk aan uitdooft over de tijd. Een werkgever kan echter wel een aanvraag indienen voor een verhoging van de subsidie alsook een verlenging in de tijd. Men moet zich wel houden aan enkele regels.

- De werkgever mag geen werknemer ontslaan om een werknemer met een VOP (of een hogere VOP) in dienst te nemen.
- De werkgever mag geen werknemer, waarvan de VOP is uitgedoofd, ontslaan om hem daarna weer in dienst te nemen met het oogmerk opnieuw te kunnen genieten van een VOP.
- Het referteloon (d.i. het loon waarop de premie wordt berekend) kan niet hoger zijn dan tweemaal het gemiddeld gewaarborgd bruto minimumloon.

Naast een VOP voor werkgevers bestaat er ook een VOP voor zelfstandigen. Een zelfstandige kan echter enkel in aanmerking komen wanneer er voldoende bedrijfsactiviteit wordt vastgesteld, nl. wanneer het belastbaar netto-bedrijfsinkomen hoger is dan het gemiddeld gewaarborgd minimumloon.

2. Aanpassingen van en aan de arbeidsomgeving

De arbeidsgehandicapte of zijn werkgever kan een tegemoetkoming krijgen voor arbeidskledij of -gereedschap, alsook aanpassingen van de arbeidspost. De VDAB bepaalt of de aanvraag gerechtvaardigd is. Wanneer een werknemer zelf een tegemoetkoming vraagt voor materiaal of kledij, zal de VDAB volgende punten nagaan bij het behandelen van de aanvraag:

- Wordt het aangevraagde materiaal courant gebruikt op de werkvloer, en is het specifiek voor de arbeidshandicap? De VDAB mag enkel de meerkosten vergoeden die de arbeidsgehandicapte ondervindt door diens handicap, dus als een aanpassing nodig is voor gereedschap dat ook andere werknemers zich moeten aanschaffen mag de VDAB enkel de aanpassing vergoeden.
- De arbeidsgehandicapte dient een verklaring voor te leggen dat de werkgever zelf de kost niet zal dragen, en de arbeidsgehandicapte ook niet zal vergoeden met een tegenwaarde in speciën.
- Er moet ook een minimale kosten-batenanalyse gemaakt worden. De aanpassing of aanschaf van materiaal en kledij dient in functie te zijn van de arbeidshandicap, er dient frequent gebruik van gemaakt te worden en er is voldoende doelmatigheid/rendementstoename om de kosten te rechtvaardigen.

Wanneer de werkgever zelf een aanvraag doet voor aanpassingen aan de werkplek, dient de VDAB bij de behandeling van de aanvraag de gelijkaardige punten in acht te nemen:

- Men moet kunnen aantonen dat de aanpassing van de arbeidspost niet gebruikelijk is in de betreffende beroepstak, en dat de aanpassing waarvoor een tegemoetkoming wordt gevraagd in functie staat van de arbeidshandicap van de werknemer of persoon in opleiding. Ook hier geldt

het principe dat enkel de meerkosten worden vergoed die een werkgever ondervindt tegenover een valide werknemer.

- Het bedrag van de aanpassing waarvoor een tegemoetkoming werd ontvangen mag niet worden ingebracht als bedrijfskosten bij de belastingafrekening.
- De werkgever moet ook enkele verbintenissen willen aangaan, zijnde:
- De arbeidsgehandicapte voor wie de aanpassingen gebeuren minstens 6 maanden in dienst houden.
- In de toekomst moet de werkgever prioriteit geven aan arbeidsgehandicapten voor de arbeidspost die met behulp van een tegemoetkoming werd aangepast.

3. Tegemoetkoming in de verplaatsingskosten en verblijfskosten voor personen met een arbeidshandicap met mobiliteitsproblemen

Ook hier gelden grosso modo dezelfde principes als onder 2. De arbeidsgehandicapte dient te voldoen aan volgende voorwaarden om een tegemoetkoming te kunnen ontvangen voor verplaatsings- of verblijfskosten:

- De aard en ernst van de handicap moeten de tussenkomst rechtvaardigen.
- De verminderde mobiliteit moet kunnen worden bewezen met een attest van een arts-specialist.
- Alle andere toelagen voor begeleiding die de werknemer ontvangt worden in mindering gebracht. Dubbele financiering is immers uit den boze.
- De kosten per kwartaal moeten hoger zijn dan € 15 per kwartaal.
- De arbeidsgehandicapte moet bewijsstukken kunnen voorleggen die de werkelijkheid van de gemaakte kosten staven.

Verder kan de VDAB ook tussenkomen in vervoerskosten voor woonwerkverkeer. Indien de aard en ernst van de handicap dit rechtvaardigt kan de VDAB de kosten van openbaar vervoer van een begeleider voor zijn rekening nemen. Indien het door de handicap niet aangewezen is om het openbaar vervoer te gebruiken, kan de VDAB tussenkomen in de vervoerskosten gemaakt door derden die de arbeidsgehandicapte vervoeren, de kosten van gespecialiseerd vervoer of de kosten gemaakt door het gebruik van een eigen aangepast gemotoriseerd voertuig. De hoogte van deze vergoeding is wettelijk bepaald. Ook kunnen arbeidsgehandicapten die een beroepsopleiding of stage volgen onder bepaalde voorwaarden recht hebben op een tegemoetkoming van de VDAB voor vervoersonkosten.

Opbouw van rapportagebestand PIPA

Zowel arbeidshandicaps alsook de taken die arbeidsgehandicapten uitvoeren zijn erg divers, en er is dan ook een heel arsenaal aan verschillende maatregelen beschikbaar om aan deze diversiteit een antwoord te bieden.

Zoals het bovenstaande illustreert, moeten er per materie ook heel wat aspecten afgetoetst worden. Indien men een heldere impactstudie wil maken van de inspectie op het PIPA-beleid, zal men hoe dan ook een aantal van die aspecten moeten aggregeren tijdens de analyse. Deze aggregatie kan plaatsvinden op verschillende

ogenblikken van de onderzoekscyclus. men kan aggregeren ex post, dat is, tijdens de analyse van het rapportagebestand kan men bekijken welke van de variabelen men wil aggregeren tot 1 variabele. Men zou echter ook kunnen aggregeren ex ante. Bij het opstellen van het rapportagebestand gaat men dan niet voor elk te controleren/controleerbaar deelaspect een kolom aanmaken, maar men gaat dan reeds een aantal facetten bundelen in 1 variabele. aggregatie ex ante heeft als nadeel dat er iets minder flexibiliteit bestaat bij de analyse van het bestand. Het voordeel van een (beperkte) mate van aggregatie ex ante is echter dat het een tijdsbesparing betekent voor de inspecteur die de tabel moet invullen. Bovendien wordt het bestand compacter (minder kolommen) en dus overzichtelijker, en is er minder kans op het invoeren van foute gegevens. Ook zijn minder computervaardigheden nodig om achteraf een heldere analyse te maken.

In onderstaande illustraties stellen we daarom reeds een beperkte mate van aggregatie ex ante voor. Bij deze oefening van aggregeren en structureren merken we eerst op dat er drie types van geïnspecteerden in het spel zijn:

- De VDAB die de aanvragen voor tegemoetkoming behandelt. De VDAB speelt een cruciale rol in alle hierboven besproken PIPA-maatregelen.
- De werkgever die arbeidsgehandicapten in dienst neemt. De volgende maatregelen dienen rechtstreeks ter ondersteuning van de werkgever en bijgevolg zal de werkgever logischerwijze het voorwerp van inspectie uitmaken:
 - a) Tegemoetkoming in de kosten voor de aanpassing van de arbeidspost van werknemers met een arbeidshandicap en van personen met een arbeidshandicap die een geïntegreerde individuele beroepsopleiding (GIBO) volgen.
 - b) De VOP voor Werkgevers.
- De werknemer met een arbeidshandicap. De volgende maatregelen ondersteunen de werknemer rechtstreeks en bijgevolg is deze logischerwijze het object van inspectie:
 - c) Tegemoetkoming in de kosten van arbeidsgereedschap en -kledij.
 - d) Tegemoetkoming in de verplaatsingskosten en verblijfkosten voor personen met een arbeidshandicap met mobiliteitsproblemen.
 - e) De VOP voor een zelfstandige.

Het lijkt ons aangewezen om de PIPA-maatregelen op te delen in twee rapportagebestanden volgens het geïnspecteerde object. Maatregelen a) en b) kunnen dan worden gerapporteerd in rapportagebestand A, terwijl c), d) en e) kunnen worden ondergebracht in rapportagebestand B.

Rapportagebestand A: Inspectie PIPA-maatregelen werkgever

Teneinde een beeld te krijgen van hoe we op een zinvolle en logische wijze kunnen aggregeren merken we op dat de te inspecteren aspecten als volgt kunnen gecategoriseerd worden.

- Fiscale fraude: Het bedrag van kosten (zoals aanpassingen van een werkplek) die werden gedekt door een tegemoetkoming worden desalniettemin aangegeven als bedrijfskosten bij de belastingaangifte.
- Fraude bij uitzendkantoren: het uitzendkantoor stort de helft van de VOP, zoals voorgeschreven, niet door aan de gebruiker.
- Overbodige of onterechte ondersteuning: de aanpassing van de werkplek is couranter dan ingeschat door de VDAB of aangegeven door de werkgever bij de aanvraag. De aanpassing lijkt niet in functie te zijn van, of in verhouding te staan tot, de arbeidshandicap van de werknemer. De aanpassing lijkt overbodig, het aangepaste materiaal wordt niet of nauwelijks gebruikt of het rendement staat niet in verhouding tot de kosten. Hier is de grens tussen conform en non-conform handelen vrij vaag. Het gaat hier vaak immers om interpretatie en meningsverschillen, en vaststellingen zullen dan vaak ook dienen om terug te koppelen met de VDAB om vuistregels gebruikt bij het behandelen van aanvragen aan te scherpen.
- Manipulatie aansprakelijkheidscriteria: Werknemers werden ontslagen teneinde iemand met recht op een VOP of hogere VOP te kunnen aannemen. Een werknemer werd eerst ontslagen teneinde hem later opnieuw te kunnen aannemen met een (hogere) VOP.
- Overige: Na de aanpassing van de werkplek werd de werknemer na minder dan zes maanden ontslagen. de werkgever lijkt geen inspanning te doen om een arbeidsgehandicapte aan te stellen op een vrijgekomen arbeidspost die met een tegemoetkoming van de VDAB werd aangepast.

De structuur van het rapportagebestand ziet er vrij gelijkaardig uit als dat van de tewerkstellingspremie 50+. Elke rij in het bestand stelt een inspectie voor. Per inspectie kunnen uiteraard meerdere dossiers geïnspecteerd worden, en als bedrijven op verschillende tijdstippen werden geïnspecteerd, zijn er meerdere rijen in het bestand gewijd aan één onderneming, en hebben we dus een paneeldatabestand.

Het codebook is als volgt.

- KBO: Het bedrijfsnummer van de geïnspecteerde onderneming zoals terug te vinden in de Kruispuntbank van Ondernemingen.
- NACE: de NACE code ook terug te vinden in de Kruispuntbank van Ondernemingen.
- jaar: Het jaar waarin de inspectie plaats vond.
- materie: de code van de materie waarover de observatie handelt. De code van inspectie PIPA werkgevers =2.
- fraude: aantal dossiers waarin sprake is van fiscale fraude, of waarbij een interimkantoor niet de helft van de VOP doorstort aan de gebruiker.
- fraud_bed: totaal bedrag fraude.
- buitprop: Het aantal dossiers waarin werd vastgesteld dat er overbodige of onterechte ondersteuning werd geleverd.
- manipul: Een variabele die aangeeft voor hoeveel dossiers de aansprakelijkheidscriteria zijn gemanipuleerd (het ontslaan van werknemers

om dezelfde of een andere werknemer met meer subsidiëring in dienst te kunnen nemen).

- overig: het aantal dossiers waarin andere inbreuken werden vastgesteld.

Hieronder volgt een illustratie van een rapportagebestand voor enkele fictieve werkgevers.

Tabel 5.2

Fictief rapportagebestand van werkgevers voor PIPA-maatregelen

KBO	NACE	jaar	materie	aant_cont	alg_inbreuk	fraude	fraud_bed	buitprop	manipul	overig
100	41	2011	2	11	2	0	-1	2	0	0
100	41	2012	2	10	0	-1	-1	0	0	0
101	41	2011	2	3	3	3	12.000	0	0	0
101	41	2012	2	3	0	0	0	0	0	0

Code -1 = "niet van toepassing"

In het fictieve rapportagebestand zijn gegevens opgenomen over twee ondernemingen. Beide ondernemingen werden geïnspecteerd in de kalenderjaren 2011 en 2012. De variabele "materie" neemt vanzelfsprekend de waarde 2 aan voor beide bedrijven, omdat de materie die in dit bestand wordt geregistreerd gaat over de PIPA-maatregelen voor werkgevers. In onderneming met KBO-nummer 100 werden 11 dossiers geïnspecteerd in 2011, en 10 in 2012. In 2011 werden in 2 dossiers problemen vastgesteld. Er werd geen financiële fraude vastgesteld, maar in beide dossiers was de inspectie van oordeel dat bepaalde aanpassingen en materiaal onterecht waren gefinancierd door de VDAB. Verder werden er dat jaar voor onderneming met KBO-nummer 100 geen problemen vastgesteld. In 2012 waren alle dossiers voor deze onderneming in orde.

Bij onderneming met KBO-nummer 101 was in 2011 meer aan de hand. Er werden 3 dossiers geïnspecteerd, en in alle drie de dossiers was er sprake van financiële fraude. Bepaalde onkosten die door de VDAB waren gesubsidieerd, werden toch aangegeven in de belastingaangifte als bedrijfskosten. Het ging in totaal om een bedrag van € 12.000. Verder was alles in orde. In 2012 werd de onderneming opnieuw geïnspecteerd en toen werden geen inbreuken meer vastgesteld.

Rapportagebestand B: Controle van de werknemer

Ook hier kunnen de te controleren aspecten worden opgedeeld in een aantal categorieën.

- Fraude: Men kan nagaan of de werknemer, in tegenstelling tot wat hij verklaarde bij de aanvraag van materiaal of kledij, ook een tegemoetkoming heeft ontvangen voor dezelfde kosten waarin de VDAB is tussengekomen.
- Overbodige of onredelijke subsidiëring van materiaal/kledij: vergoede hulpmiddelen of kledij is couranter dan ingeschat door de VDAB of aangegeven door de werknemer. Het materiaal wordt niet frequent gebruikt, de baten staan niet in verhouding tot de kosten, of men heeft te maken met een zelfstandige met VOP bij wie te weinig bedrijfsactiviteit werd vastgesteld.

- Overig: Men kan denken aan het te veel uitbetalen van vervoerskosten. Een begeleider kan bijvoorbeeld niet het goedkoopste biljet genomen hebben terwijl het volledige bedrag toch is teruggevorderd van de VDAB.

De structuur van het rapportagebestand is nu enigszins anders dan voorheen. Immers, in dit rapportagebestand komen de rijen niet overeen met een inspectieopdracht, maar wel met een geïnspecteerd dossier. Elk geïnspecteerd dossier (i.e. van een werknemer) krijgt dus een rij in de tabel. De kolommen zijn wederom de variabelen, en het codeboek ziet er als volgt uit.

- KBO: Het bedrijfsnummer van de geïnspecteerde onderneming zoals terug te vinden in de Kruispuntbank van Ondernemingen.
- NACE: de NACE code ook terug te vinden in de Kruispuntbank van Ondernemingen.
- jaar: Het jaar waarin de inspectie plaats vond.
- materie: de code van de materie. De code van inspectie PIPA werknemers is 3.
- RRN: Rijksregisternummer van de gecontroleerde arbeidsgehandicapte.
- fraud_bed: Dit is gelijk aan 0 wanneer er geen financiële fraude is vastgesteld. In het andere geval is deze variabele gelijk aan de som die dubbel werd gefinancierd.
- buitprop: Een 0/1 variabele die 1 aanneemt wanneer er onterechte of onredelijke tegemoetkomingen werden opgemerkt.
- vopdrempel: Deze variabele is 0 indien er geen voorstel tot terugvordering van een VOP voor zelfstandigen wordt gedaan. Deze variabele is gelijk aan het bedrag waarvoor een terugvordering wordt voorgesteld van een VOP voor zelfstandigen, bv. omdat de minimum bedrijfsactiviteit niet werd gehaald. In de gevallen waarin het dossier niet handelt over een zelfstandige met een VOP, neemt deze variabele de waarde "-1" aan.
- overig: Een 0/1 variabele die 1 aanneemt wanneer andere problemen werden vastgesteld.

Een illustratie van een rapportagebestand voor fictieve personen werkende in fictieve ondernemingen wordt hieronder weergegeven.

Tabel 5.3

Fictief rapportagebestand van werknemers voor PIPA-maatregelen

KBO	NACE	jaar	materie	RRN	fraud_bed	buitprop	vopdrempel
100	41	2011	3	260119772541231	0	0	-1
100	41	2011	3	1511198821211	3.500	0	-1
214	41	2012	3	1511198821211	0	0	-1

Code -1 = "niet van toepassing"

Dit bestand bevat informatie over twee werknemers (zie RRN), twee ondernemingen (zie KBO) en twee verschillende kalenderjaren. De code voor de variabele "materie" is steeds 3, inspectie PIPA-maatregelen werknemers. De persoon met rijksregisternummer 260119772541231 werkte in 2011 bij onderneming met KBO-nummer 100. Bij inspectie bleken er geen problemen te zijn, en deze werd in 2012 niet meer opnieuw geïnspecteerd. Bij de persoon met rijksregisternummer

1511198821211 is meer aan de hand. In 2011 werkte deze persoon voor de onderneming met KBO-nummer 100. Er werd financiële fraude vastgesteld. Hij had namelijk materiaal aangekocht en vergoed gekregen door de werkgever, en had hetzelfde materiaal ook volledig vergoed gekregen door de VDAB. In totaal werd een bedrag van € 3.500 dubbel gefinancierd. De werknemer werd op staande voet ontslagen maar vond spoedig werk in een andere onderneming in een andere provincie. Ook nu maakte hij opnieuw gebruik van PIPA-maatregelen en de inspectie WSE vond het raadzaam om in 2012 dit dossier opnieuw aan een inspectie te onderwerpen. In 2012 werden er bij deze persoon geen inbreuken meer vastgesteld.

5.5 Impact op microniveau: Samenwerking met derde instanties voor evaluatie en bijsturing risicoanalyse

In deze sectie wordt geïllustreerd hoe de inspectie, door samen te werken met andere instanties, respectievelijk haar risicoanalyse kan optimaliseren en evalueren. Ook hierbij is een kwantitatief rapportagebestand het uitgangspunt. In het jaarrapport 2011 van IWSE wordt een risicoanalyse geschetst op macroniveau. Dit wil zeggen dat de verschillende materies waarvoor IWSE bevoegd is een label krijgen van extreem ernstig tot neutraal. In een volgende stadium wil men echter een risicoanalyse uitwerken op materieniveau. Binnen elke materie gaat men dus kijken welke ondernemingen een groter risico op non-conform handelen betekenen dan andere. Om doelgerichte controles te kunnen uitvoeren is het daarom belangrijk om verschillende risicomodellen tegenover elkaar te zetten en na te gaan welk model het beste werkt.

Men zou het voorbeeld kunnen volgen van de federale inspecties, zoals eerder beschreven, en via een retrospectieve analyse nagaan welk model het best non-conform handelen voorspelt.

Werkwijze:

- De Inspectie WSE levert aan de KSZ de kwantitatieve rapportagebestanden van 2011 en 2012. Dit zijn jaartallen waarin de materiegerichte risicoanalyse nog niet in voege was.
- De KSZ stuurt het bestand door naar de federale inspectie waarmee een overeenkomst werd gesloten.
- De federale inspectiedienst laadt de kwantitatieve rapportagebestanden in OASIS.
- De observaties (geïnspecteerde ondernemingen) worden in twee groepen verdeeld, *Groep A* en *Groep B*.
- Op *Groep A* wordt dan datamining uitgevoerd. Het computersysteem zal een gamma aan modellen uittesten en een lijst opleveren van die modellen die het al dan niet plegen van een inbreuk goed voorspellen. Dit proces kan apart worden uitgevoerd voor verschillende (groepen van) inbreuken.
- Inspectie WSE kiest samen met de analisten van de federale inspectiedienst het model uit dat de voorkeur wegdraagt. Zoals uitgelegd in vorig hoofdstuk hoeft dit niet meteen het model te zijn met de grootste

voorspellingskracht. Ook interpreteerbaarheid van een model speelt een grote rol in de keuze.

- Het gekozen model wordt dan uitgetest op de observaties van *Groep B*.
- Indien ook nu het model een goede voorspelling oplevert van het al dan niet plegen van een inbreuk, kan men het gaan gebruiken voor de risicoanalyse. De federale inspectie kan nu, aan de hand van het gekozen model, een lijst opleveren van risicovolle ondernemingen alsook een lijst van parameters die ertoe geleid hebben dat de onderneming werd geselecteerd als risicovol.

Uiteraard moet men er zich van bewust zijn dat fraude verandert over de tijd. Ondernemingen leren bij en zullen hun gedrag aanpassen om minder snel gedetecteerd te worden, doch daarom niet om (volledig) normconform te handelen. Het is dus belangrijk om zo'n retrospectieve risicoanalyse periodiek uit te voeren, en om in sommige jaren het risicomodel los te laten en meer lukrake controles uit te voeren. De *winst* die men haalt uit een risicoanalyse, moet men deels herinvesteren in het verkennen van het terrein.

5.6 Impact op microniveau: Samenwerking met derde instanties voor succesvolle implementatie en evaluatie van preventiecampagnes

Voor materies met een lage dekkingsgraad, of met een regelgeving die door de potentieel geïnspecteerden niet voldoende gekend is, kan het nuttig zijn om het normconform handelen niet enkel trachten te beïnvloeden door inspecties uit te voeren, maar ook door preventiecampagnes op te zetten. Hieronder illustreren we hoe de Inspectie WSE, in samenwerking met ondernemingen, preventiecampagnes kan opzetten én de impact ervan kan evalueren om de tewerkstelling van illegale arbeiders en het aantal malafide uitzendkantoren terug te dringen.

De dekkingsgraad van de controles in deze materies is zeer laag gegeven het enorm aantal ondernemingen. Daarom stellen we een preventieve campagne voor waarbij ondernemingen worden aangeschreven en worden herinnerd aan de sancties die hen boven het hoofd hangen wanneer non-conform handelen wordt vastgesteld. Zo'n campagne kan de mate van het normconform handelen verhogen maar kan ons eveneens bruikbare informatie verschaffen voor de bijsturing van het beleid.

Wanneer de preventie-campagne werkt, heeft de IWSE op een vrij goedkope manier de impact op het normconform handelen kunnen vergroten. Wanneer de preventiecampagne niet heeft gewerkt, is er een discrepantie met de internationale literatuur. Een illustratie in hoofdstuk 3 leerde ons immers dat individuen nogal gevoelig zijn voor een mailing waarin wordt uitgelegd wat de sancties zijn voor non-conform handelen. Heeft de mailing dus geen impact, dan kan deze informatie misschien meegedeeld worden in een parlementaire zitting of als antwoord op een parlementaire vraag, en kunnen politici in beweging gezet worden om de wetgeving omtrent sanctionering van de misdrijven aan te passen.

Werkwijze:

- De inspectie stelt een poel van ondernemingen samen die ze graag zou inspecteren, maar die te groot is om volledig af te werken.
- De poel van bedrijven wordt willekeurig ("at random") opgedeeld in twee groepen, *Groep A* en *Groep B*. Om de randomisatie goed te laten verlopen, zou men bv. ondernemingen met een even KBO-nummer kunnen onderbrengen in *Groep A* en ondernemingen met een oneven KBO-nummer in *Groep B*.
- Ondernemingen uit *Groep A* krijgen een brief toegestuurd met de post, waarin hen wordt gewezen op hun plichten inzake tewerkstelling van vreemdelingen en/of inzake het werken met uitzendkantoren. Eveneens kunnen ook de sancties vermeld worden die worden opgelegd bij een vastgestelde overtreding.
- Een bestand met KBO-nummer en een dummy-variabele die aangeeft of een bedrijf tot *Groep A* dan wel *Groep B* behoort wordt verstuurd naar de KSZ.
- De KSZ stuurt het bestand door naar een federale inspectiedienst waarmee afspraken werden gemaakt.
- Na koppeling met OASIS kan opnieuw via computerkracht datamining worden uitgevoerd, waarbij de afhankelijke variabele van de modellen de indicator is die aangeeft tot welke groep een onderneming behoort. De datamining kan ons leren of in de maanden na de mailing ondernemingen die een brief ontvingen zich anders gaan gedragen dan de ondernemingen die geen brief ontvingen. Daalt de technische werkloosheid, worden er meer arbeidskaarten aangevraagd?
- IWSE selecteert ondernemingen uit de poel voor een controle, en kan vaststellen of ondernemingen die een brief ontvingen minder vaak betrapt worden op inbreuken dan andere.

Uiteraard leent de opzet van het onderzoek zich voor verdere analyse door de informatie over aan welke ondernemingen welke soort brief is verstuurd, op te nemen in de eerder genoemde kwantitatieve rapportagebestanden.

Een aanzet voor een brief die kan verstuurd worden volgt hieronder.

"Geachte Heer/Mevrouw,

Het Ministerie van Werk en Sociale Economie heeft besloten, samen met haar inspectiedienst, een hoge prioriteit te maken van het bestrijden van illegale uitzendarbeid.

In de regel is het uitlenen van arbeid immers verboden (zie Artikel 31, § 1, eerste lid, van de Wet Terbeschikkingstelling). Uitzendkantoren die een erkenning hebben verkregen van de bevoegde overheid kunnen echter wel uitzendkrachten ter beschikking stellen van gebruikers. Zij dienen te handelen conform het decreet van 10 december 2010 betreffende de private arbeidsbemiddeling en haar uitvoeringsbesluiten waarop Inspectie WSE het toezicht uitoefent. Omdat een niet-correcte naleving van de regelgeving vaak leidt tot uitbuiting van werknemers, het

ontduiken van sociale zekerheidsbijdragen en andere financiële verplichtingen en dus tot oneerlijke concurrentie, neemt de overheid inbreuken hieromtrent zeer ernstig.

Gebruikers mogen enkel een beroep doen op erkende uitzendkantoren. In bijlage alsook op de website <http://www.werk.be/online-diensten/bureaus-private-arbeidsbemiddeling/erkenning-van-uitzendbureaus> vindt u een helder overzicht van de regelgeving, van de erkende bureaus alsook van de sancties waaraan een gebruiker zich mag verwachten indien een overtreding wordt vastgesteld.

Wij hopen dat deze informatie u kan helpen om samen met ons een correcte toepassing van de regelgeving omtrent private arbeidsbemiddeling te verwezenlijken.

Met vriendelijke groeten,

X”

5.7 Impact op macroniveau: het meten van maatschappelijke impact

Tot nog toe hebben we gesproken over impactmeting op het normconform handelen van de (potentieel) geïnspecteerden. Het is echter ook interessant te kijken naar de maatschappelijke impact. Om een beeld te krijgen van de impact van IWSE op de politiek en op de maatschappij zijn de volgende metingen denkbaar.

1. De frequentie van “Vlaamse Arbeidsinspectie” of “Inspectie WSE” in de media

Het analyseren van patronen in krantenartikelen om maatschappelijke fenomenen te kunnen meten en onderzoeken is reeds een tijd ingeburgerd in de verschillende takken van de sociale wetenschappen. De huidige technologie maakt het trouwens eenvoudig om een beeld te krijgen welke uitdrukkingen hoe vaak worden gebruikt per tijdseenheid. Er kunnen dus tijdstrends geconstrueerd worden, die aangeven of Inspectie WSE er in slaagt vaker de media te halen dan voorheen. Ook kan men kijken naar pieken in de patronen, die aangeven of inspectie WSE er regelmatig in slaagt een belangrijk maatschappelijk probleem onder de aandacht te brengen van het grote publiek. In Vlaanderen kunnen de betalende diensten op de webstek van mediargus (www.mediargus.be), het digitale platform van de Vlaamse Dagbladpers, helpen om zo'n analyse uit te voeren. Immers, alle artikels van de Vlaamse dagbladen worden op deze webstek gearchiveerd, en de tekst wordt uit de mediabestanden geëxtraheerd zodat de krantenartikelen met een zoekfunctie doorzocht kunnen worden.

2. Het scannen van kamerstukken

Debatten in het parlement reflecteren wat er leeft in de maatschappij. De inspectie WSE zou kunnen bijhouden hoe vaak zij wordt gevraagd om antwoorden voor te bereiden op parlementaire vragen. Ook kan men parlementaire stukken scannen op woordgroepen zoals “Inspectie WSE” “Arbeidsinspectie” of op de materies waar de inspectie zich mee inlaat.

3. *Het meten van trends in attitudes en trends in normconform handelen via enquêtes*

Het optreden van de inspectie heeft op korte termijn in het beste geval een impact op het normconform handelen van de geïnspecteerde. Echter, men mag hopen (ook op basis van internationaal onderzoek) dat een goede handhaving zal leiden tot een verandering van normen en waarden in de maatschappij, waardoor het normconform handelen wordt geïnternaliseerd en niet meer moet worden afgedwongen met een constante dreiging van inspectie en sanctionering.

Om de maatschappelijke impact in kaart te brengen kan het erg nuttig zijn om de bevolking op actieve leeftijd op regelmatige basis te bevragen over attitudes omtrent illegale arbeid, discriminatie etc. Het zou erg interessant zijn mochten zulke vragen geïntegreerd kunnen worden in een survey die zeer frequent wordt afgenomen van een grote steekproef, zoals de enquête naar de arbeidskrachten (EAK) die wordt gecoördineerd door de Federale Overheidsdienst Economie. Navraag bij de FOD Economie leert ons dat een verzoek tot opnemen van een aantal vragen zal moeten gericht worden aan de studiedienst van de Vlaamse regering. Indien deze positief oordeelt over de relevantie van de gevraagde gegevens, zal zij een producent van statistieken contacteren om te zien wat mogelijk is, bijv. de Algemene Directie voor de Statistiek en Economische Informatie (ADSEI), het vroegere Nationale Instituut voor de Statistiek (NIS).

Men zou zelfs verder kunnen gaan, en deze enquêtes kunnen gebruiken niet enkel voor het meten van attitudes tegenover non-conform gedrag, maar ook om een globale trend te meten in het normconform handelen zelf. Men zou kunnen vragen aan individuen:

- of zij weet hebben van discriminatie op de werkplek,
- of er illegale arbeiders zijn op de werkplek,
- of er gewerkt wordt met malafide uitzendkantoren,
- of er wordt gefraudeerd met overheidssubsidies, etc.

De vraag kan gesteld worden of antwoorden op vragen die peilen naar normconform handelen of attitudes daaromtrent betrouwbare resultaten opleveren. Immers, hoewel de enquêteur anonimiteit verzekert, is de respondent wellicht terughoudend om toe te geven dat hij (of zijn bedrijfsleider) zich schuldig maakt aan door de wet verboden activiteiten. Deze terughoudendheid kan niet enkel gevolgen hebben voor de desbetreffende vragen, maar kan een respondent er ook toe aanzetten om volledig af te zien van het invullen van de survey, waardoor de nonresponse ratio omhooggaat en de kwaliteit van de data afneemt. Bij het opstellen van een aanvraag om data toe te voegen aan een enquête zal het cruciaal zijn om de producent van de statistieken ervan te overtuigen dat het toevoegen van de extra vragen de kwaliteit van de data niet in gevaar brengt.

In de literatuur worden enkele technieken voorgesteld om de waarheidsgetrouwheid van antwoorden op gevoelige vragen te verhogen, en om de nonresponse ratio die deze vragen kunnen veroorzaken te verkleinen. In een recent overzichtspaper vatten

Pacolet en de Wispelaere (2012) een aantal surveytechnieken samen met het toepassingsgebied van het meten van sociale en fiscale fraude indachtig.

De Randomized Response Techniek in enquêtes

Men kan gebruik maken van de "randomized response techniek", die reeds veelvuldig is gebruikt in de literatuur om normconform handelen te meten. Er bestaan vele varianten van deze techniek, maar het principe is als volgt.

- De interviewer stelt een vraag, bv. "Wordt er bij u op het werk gediscrimineerd bij de selectie van nieuwe werknemers? Ja/nee".
- De respondent gooit met een dobbelsteen en enkel hij ziet het resultaat van de worp (dus niet de enquêteur).
- Indien het aantal ogen oneven is, geeft de respondent een eerlijk antwoord. Indien het aantal ogen even is, zegt hij altijd ja. Op die manier weet de enquêteur niet of de respondent "ja" antwoordt omdat er gediscrimineerd wordt, dan wel omdat hij een even aantal ogen heeft gegooid.
- De wet van de grote aantallen leert ons dat bij grote steekproeven het aantal worpen met even ogen ongeveer 50% zal bedragen. De methode laat ons dus nog wel toe een beeld te krijgen van het niveau van normconform handelen, en de evolutie van dit niveau over de tijd.
- Om het percentage nee-antwoorden te kennen dient men het percentage respondenten die "nee" antwoorden te verdubbelen.

Verscheidene varianten van de Randomized Response Techniek werden voorgesteld in de literatuur, maar de nadelen van deze techniek over de verschillende varianten heen zijn grosso modo de volgende:

- Het is niet altijd eenvoudig en bijgevolg tijdrovend om het concept uit te leggen aan de respondent;
- Hierbij aansluitend vertrouwt de respondent niet altijd dat de methode anonimiteit garandeert, wat dan weer de nonresponse ratio en het aantal valse ontkenningen doet toenemen.

De Unmatched Count Techniek of List Randomization in enquêtes

De "unmatched count techniek" (zie o.a. Coutts en Jann, 2011; McKenzie en Siegel, 2013; Raghavarao en Federer, 1979) is ietwat meer immuun tegen de kritieken die opgaan voor de Randomized Response Techniek. Het principe is als volgt.

- Men deelt de respondenten "at random" op in twee groepen van gelijke grootte, Groep A en Groep B.
- Tijdens het interview krijgt Groep A een vraag met een lijst van vrij neutrale stellingen, bijv.: "Ik ben tijdens het voorbije jaar verhuisd.
Ik ben gedurende het voorbije jaar op reis geweest naar Spanje.
Ik heb het voorbije jaar een nieuwe fiets gekocht.
Ik verjaar in de lente."
De respondenten worden gevraagd om met "ja" te antwoorden wanneer één van de stellingen opgaat, en met "nee" in het andere geval.

- De respondenten in Groep B krijgen tijdens het interview dezelfde vraag, maar aan de lijst van stellingen wordt dan een gevoelige stelling toegevoegd zoals:
"Bij mij op het werk wordt er gediscrimineerd op grond van afkomst."
- Aan de hand van de verschillen in ja-antwoorden tussen Groep B en Groep A kan men dan trachten te berekenen wat het percentage ja-antwoorden is op de gevoelige stelling.

Er dienen hierbij echter twee belangrijke opmerkingen gemaakt te worden. Ten eerste wordt in de toegepaste literatuur vaak het percentage ja-antwoorden op de gevoelige stelling gelijkgesteld aan het verschil in ja-antwoorden tussen Groep B en Groep A. Dit lijkt ons echter niet helemaal correct, en lijkt te leiden tot een onderschatting van het percentage ja-antwoorden op de gevoelige stelling. Immers, er zijn respondenten die "ja" antwoorden omdat voor hen een neutrale stelling opgaat alsook de gevoelige stelling.

Definiëren we X als een dummyvariabele die 1 aanneemt wanneer minstens één neutrale stelling opgaat, en Y een dummy-variabele die de waarde 1 aanneemt wanneer de gevoelige stelling opgaat. Groep A laat ons toe om $Pr(X)$ te schatten, de kans dat X gelijk is aan 1. In ons voorbeeld behelst dit de kans dat een respondent "ja" antwoordt op de vraag waarin de vier neutrale stellingen worden opgesomd.

Groep B laat ons toe om de kans te schatten dat X gelijk is aan 1 of Y gelijk is aan 1 of zowel X als Y gelijk zijn aan 1. Laten we deze kans benoemen als $Pr(Z)$. In ons voorbeeld is dit de kans dat een respondent "ja" antwoordt op de vraag waarin zowel de vier neutrale stellingen als de gevoelige stelling worden opgesomd.

Regels van het kansrekenen leren ons om $Pr(Z)$ te schrijven als:

$$Pr(Z) = Pr(X) + Pr(Y) - Pr(X) * Pr(Y)$$

Hieruit volgt:

$$Pr(Y) = (Pr(Z) - Pr(X))(1 - Pr(X))$$

Stel bij wijze van voorbeeld dat in Groep A de ratio van ja- en nee-antwoorden gelijk is aan 1/3, en dat deze in Groep B gelijk is aan 1/2. Als de beide groepen groot genoeg zijn, kan men ervan uitgaan dat $Pr(X)$ gelijk is aan 1/3 en $Pr(Z)$ gelijk aan 1/2. De ratio van respondenten waarvoor de gevoelige stelling opgaat is dan gelijk aan $(1/2 - 1/3) / (1 - 1/3) = 1/4$. In dit voorbeeld gaat de gevoelige stelling dus op voor ca. 25% van de populatie.

Er is echter een tweede kanttekening die we moeten maken, namelijk dat bovenstaande berekening enkel opgaat als X en Y onafhankelijk zijn van elkaar. Het is immers niet ondenkbeeldig dat het al dan niet opgaan van een neutrale stelling verband houdt met het al dan niet opgaan van de gevoelige stelling. Immers, normconform handelen is vaak gerelateerd aan iemands socioeconomische

achtergrond. Als de veronderstelling van onafhankelijkheid van X en Y incorrect is, dan zijn onze berekeningen vertekend.

5.8 Aanbevelingen

- Probeer verslagen van inspecties zo veel mogelijk te kwantificeren en bij te houden in een database.
- De inspectie hoeft niet te wachten op allerhande samenwerkingsverdragen voor een eerste aanzet tot impactanalyse. Op korte termijn kan via een eenvoudig kwantitatief rapportagebestand voor de materies met een voldoende hoge dekkingsgraad reeds een impactmeting worden uitgevoerd.
- Op middellange termijn is samenwerking met andere (overheids)instanties erg aangewezen. Analisten, ICT-personeel en -materiaal kunnen gedeeld worden, wat de vaste kosten enorm kan drukken.
- Samenwerking dient niet enkel te gebeuren op regionaal niveau, maar ook dient men op termijn de samenwerking met federale diensten te versterken, én te institutionaliseren.
- Het is ook belangrijk om goede contacten te houden met de academische wereld. Enkele eerste verkennende analyses zouden een welgekomen onderwerp kunnen zijn voor een student die een bachelor- of masterscriptie dient te schrijven. Indien onderzoek interessante resultaten kan opleveren voor een breder (internationaal) publiek, bestaan er mogelijk opportuniteiten om samen te werken met academische onderzoekers en doctoraatstudenten.
- Voor onderzoek dat van waarde is voor een breder academisch publiek kan men trachten andere financieringsbronnen aan te boren dan VIONA. Vele verstrekkers van subsidies voor academisch onderzoek komen sneller over de brug wanneer onderzoek gebeurt in samenwerking met, en eventueel gefinancierd door, een instantie buiten de academische wereld.

Mogelijke initiatieven door IWSE op korte, halflange en lange termijn op basis van voorliggende rapportage

Termijn	Initiatieven t.a.v. de materie/geïnspecteerde en/of beleids- & maatschappelijke impact			
	Inhoudelijk	Op het vlak van ICT	Samenwerking met andere instanties	Organisatorisch op afdelingsniveau
Korte termijn (binnen nu tot 1 jaar)	<ul style="list-style-type: none"> * Beleids- & maatschappelijke impact: 1. Tellen parlementaire vragen 2. Turven van het aantal X dat Inspectie WSE in de pers voorkomt * De uitkomsten van een risicoanalyse: labels aanpassen aan de hand van prioriteit i.p.v. ernst 		<ul style="list-style-type: none"> * Onderhoud goede contacten met de academische wereld, bijvoorbeeld door studenten stages te laten lopen of scripties te laten schrijven op het terrein van impact meting 	<ul style="list-style-type: none"> * Plan van aanpak uitschrijven (K.T., half L.T. en L.T.) n.a.v. rapport met aangeven van beleidsmatige prioritering * Op korte termijn kan via een eenvoudig kwantitatief rapportagebestand voor de materies met een voldoende hoge dekingsgraad reeds een impactmeting worden uitgevoerd.
Halflange termijn (binnen 1 tot 2 jaar)	<ul style="list-style-type: none"> * Let bij impactonderzoek goed op aannames, variabelen, strategieën om doelen te bereiken * Let op dat zoektocht naar impact of causale verbanden niet evident is, en tegelijkertijd politiek gevoelig kan liggen * Bewustwording dat elke beleidswijziging in potentie een natuurlijk experiment is, mits goed voorbereid, van groot belang voor impactmeting 	<ul style="list-style-type: none"> * Stel ondernemers via website in staat zichzelf te inspecteren * Heroriënteren van de verplichte registratiegegevens in het Geïntegreerde Inspectie-applicatie (GINA) om effectmetingen mogelijk te maken (opbouw kwantitatief rapportage bestand) 	<ul style="list-style-type: none"> * Uitbreiding juridische ondersteuning aan inspecteurs * Voor onderzoek dat van waarde is voor een breder academisch publiek kan men trachten andere financieringsbronnen aan te boren dan VIONA, want subsidiegevers voor academisch onderzoek zijn op zoek naar valorisatie (i.e. maatschappelijke relevantie) door samenwerking met veldpartijen, zoals de IWSE 	<ul style="list-style-type: none"> * Houd tevredenheidsonderzoeken onder contactpersonen van de inspectie * Onderzoeken in hoeverre trends in attitudes en in normconform handelen via enquêtes in Vlaanderen te achterhalen zijn (cf. Nederlandse arbeidsinspectie) * Preventieve acties kunnen inbreuken voorkomen, en kunnen de bekendheid en het imago van de IWSE vergroten
Lange termijn (binnen 2 tot 4 jaar)	<ul style="list-style-type: none"> * Vul het begrip "inspectievakantie" in, bijv. zoals bij de Nederlandse arbeidsinspectie 	<ul style="list-style-type: none"> * Ontsluiting en raadplegen van databanken van federale sociale inspectiediensten (Opgelet: impact 6^o staatshervorming) * Richt een webstek in waar klussers en opdrachtgevers elkaar kunnen vinden om zicht te krijgen op de omvang van malafide ondernemingen zonder deze overigens op te sporen 	<ul style="list-style-type: none"> * Federale sociale inspectiediensten. (Opgelet: impact 6^o staatshervorming!) 	

Literatuur

- Angrist, J. and S. Pischke (2008) *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press; first edition.
- Bartel, A. en L. Thomas. (1985) Direct and Indirect Effects of Regulation: A New Look at OSHA's Impact. *Journal of Law and Economics* 28, 1-25.
- Bertrand, M., E. Duflo en S. Mullainathan (2004) How Much Should We Trust Differences-in-Differences Estimates? *The Quarterly Journal of Economics* 119, 249-275.
- Björkman, M. en J. Svensson (2009) Power to the People: Evidence from a Randomized Field Experiment on Community-Based Monitoring in Uganda. *The Quarterly Journal of Economics* 124, 735-769.
- Bogaerts, A. (2011) *Review Literatuur Effectmeting: omgevingscan in het Kader van de Vorming Inspectie SZW*. Inspectie SZW, juni.
- Bogaerts, A. (2012) *Welke Factoren Beïnvloeden de Implementatie van Effectmeting bij Rijksinspecties?* Afstudeerscriptie Faculteit Sociale Wetenschappen. Vrije Universiteit Amsterdam.
- Cameron, A. en P. Trivedi (2005) *Microeconometrics: Methods and Applications*. Cambridge University Press.
- Coutts, E. en B. Jann (2011) Sensitive Questions in Online Surveys: Experimental Results for the Randomized Response Technique (RRT) and the Unmatched Count Technique (UCT). *Sociological Methods & Research* 40, 169-193.
- Departement WSE (2011), *Het Departement Werk en Sociale Economie Verruimt je Blik op Werk*, Brussel.
- Departement WSE (2013), *Jaarrapport Inspectie Werk en Sociale Economie 2012*, Brussel.
- Di Tella, R. and E. Schargrodsy (2004) Do Police Reduce Crime? Estimates Using the Allocation of Police Forces after a Terrorist Attack. *American Economic Review* 94, 115-133.
- Drago, F., R. Galbiati en P. Vertova (2009) The Deterrent Effects of Prison: Evidence from a Natural Experiment. *Journal of Political Economy* 117, 257-280.
- Duflo, E., R.N. Hanna en S.P. Ryan. Incentives Work: Getting Teachers to Come to School. *American Economic Review*, Forthcoming.
- Fellner, G., R. Sausgruber and C. Traxler (Forthcoming) Testing Enforcement Strategies in the Field: Threat, Moral Appeal and Social Information. *Journal of the European Economic Association*.
- Gray, W.B. en C.A. Jones (1991) Longitudinal Patterns of Compliance with Occupational Safety and Health Administration: Health and Safety Regulations in the Manufacturing Sector. *Journal of Human Resources* 26, 623-653.
- Imbens, G. en J. Wooldridge (2009) Recent Developments in the Econometrics of Program Evaluation. *Journal of Economic Literature* 47, 5-86.

- Inspectie SZW (2011) *Evaluatie-onderzoek Arbeidsinspectie: Onderzoek onder Geïnspecteerden naar de Beleving van het Contact met Rijksinspecties*. Hoofdrapport, september.
- Locke, R., F. Qin en A. Brause (2007) Does Monitoring Improve Labor Standards? Lessons from Nike. *Industrial and Labor Relations Review* 61, 3-31.
- Mckenzie, D. en M. Siegel (2013) *Eliciting Illegal Migration Rates through List Randomization*. IZA Discussion Papers 7401. Institute for the Study of Labor (IZA), Bonn.
- Mendeloff, J. (2005) The Declining Effects of OSHA Inspections on Manufacturing Injuries, 1979-1998. *Industrial and Labor Relations Review* 58, 571-587.
- Olken, B.A. (2007) Monitoring Corruption: Evidence from a Field Experiment in Indonesia. *Journal of Political Economy* 115, 200-249.
- Pacolet, J. en F. De Wispelaere (2008) *Ontwikkelen van een Conceptueel Methodologisch Kader voor Risicoanalyse binnen de Entiteit Inspectie Werk en Sociale Economie*. Een Onderzoek in Opdracht van de Minister van Werk, Onderwijs en Vorming in het Kader van het VIONA-Onderzoeksprogramma.
- Pacolet, J. en F. De Wispelaere (2012) *Social and Fiscal Fraud in Belgium: Designing an Appropriate Survey Methodology to Reveal Social and Fiscal Fraud*. Working Paper. Hiva, Leuven.
- Pranger, R. en J. Dagevos (2001) *Haalbaarheidsstudie Instrument Effectiviteitsmeting Arbeidsinspectie*. IVA. Tilburg.
- Raghavarao, D. en W. Federer (1979) Block Total Response as an Alternative to the Randomized Response Method in Surveys. *Journal of the Royal Statistical Society Series B* 41, 40-45.
- Raudenbusch, S., A. Martinez en J. Spybrook (2007) Strategies for Improving Precision in Group-Randomized Experiments. *Educational Evaluation and Policy Analysis* 29, 5-29.
- Rossie, T. (2012) *Risicoanalyse Macroanalyse 2012*. Departement Werk en Sociale Economie, Brussel.
- United Nations (2004) *Handbook on Practical Anti-Corruption Measures for Prosecutors and Investigators*, Vienna.
- Warner, S. (1965) Randomized Response: A Survey Technique for Eliminating Evasive Answer Bias. *Journal of the American Statistical Association* 60, 63-69.
- Weber, E. en C. Hsee (1998) Cross-cultural Differences in Risk Perception but Cross-cultural Similarities in Attitudes towards Risk. *Management Science* 44, 1205-1217.
- Weil, D. (2005) Public Enforcement/Private Monitoring: Evaluating a New Approach to Regulating the Minimum Wage. *Industrial and Labor Relations Review* 58, 238-257.
- Weil, D. (2007) Crafting a Progressive Workplace Regulatory Policy: Why Enforcement Matters. *Comparative Labor Law and Policy Journal* 28, 125-154.
- Weil, D. (2008) A Strategic Approach to Labour Inspection. *International Labour Review* 147, 349-375.

Werkgroep Effectmeting IG-Beraad (2005) *Leidraad Effectmeting bij Inspecties*, Den Haag.

Werkgroep Effectmeting Inspectie SZW (2011) *Visie op Effectmeting bij Inspectie SZW*. Werkdocument, december.

WODC, Ministerie van Justitie (2010) *Toezicht en Effectmeting: Het Kan. Researchsynthese van Onderzoek Naar de Effectiviteit van Toezichtsinterventies*. Den Haag.

Websites:

www.mediargus.be

www.rva.be

www.vdab.be

www.syntra.be

www.rsz.be

www.werk.be

www.zelfinspectie.nl

Lijst van afkortingen

ADSEI: Algemene Directie voor de Statistiek en Economische Informatie
DIMONA: Déclaration Immédiate/ Onmiddellijke aangifte
DMFA: Déclaration multifonctionelle/ multifunctionele Aangifte
EAK: Enquête naar de Arbeidskrachten
ECHP: European Community Household Panel
ESF: Europees Sociaal Fonds
FOD: Federale Overheidsdienst
GIBO: geïntegreerde individuele beroepsopleiding
INSZ: Identificatienummers van Sociale Zekerheid
IWSE: Inspectie Werk en Sociale Economie
KBO: Kruispuntbank van Ondernemingen
KSZ: Kruispuntdatabank Sociale Zekerheid
LIMOSA: Landenoverschrijdend Informatiesysteem ten behoeve van Migratie Onderzoek
MOVI-netwerk: Managementondersteuning Vlaamse Instellingen en het Inspectie-netwerk
NACE: Nomenclature statistique des activités économiques dans la Communauté européenne
NIS: Nationaal Instituut voor de Statistiek
OASIS: Organisation Anti-fraude des Services d'Inspection Sociale
PIPA: Professionele Integratie van Personen met een Arbeidshandicap
PSBH: Paneelstudie van Belgische Huishoudens
ROA : Researchcentrum voor Onderwijs en Arbeidsmarkt
RSZ: Rijksdienst voor Sociale Zekerheid
RVA: Rijksdienst voor Arbeidsvoorziening
SCSA: Sectoraal Comité van de Sociale Zekerheid
SILC: statistics on Income and Living Conditions
SZW: Sociale Zaken en Werk
TSW: Toezicht Sociale Wetten
VDAB: Vlaamse Dienst voor Arbeidsbemiddeling en Beroepsopleiding
VOP: Vlaamse Ondersteuningspremie
VSAWSE: Vlaams Subsidieagentschap Werk en Sociale Economie
WSE: Werk en Sociale Economie

Glossarium

Afhankelijke variabele: De afhankelijke of verklaarde variabele is een variabele in het model die afhangt, wordt bepaald, door andere factoren.

Categorische variabele: een variabele die een beperkt aantal waarden kan aannemen en beschrijft categorische data die van nature geordend of ongeordend kunnen zijn.

Ceteris paribus: geijkte uitdrukking voor “als al het overige hetzelfde blijft”.

Continue variabele: een variabele die waarden aanneemt uit de verzameling van de reële getallen.

Cross-sectionele data: De data bevatten informatie van een steekproef voor slechts één tijdsperiode.

Cross-sectionele tijdreeksen: Men observeert meerdere subjecten voor een heel aantal periodes. De structuur van cross-sectionele tijdreeksdata is een spiegeling van die van paneeldata. Bij paneeldata observeert men relatief veel subjecten en relatief weinig tijdsperiodes, en bij cross-sectionele tijdreeksen is dit omgekeerd.

Discrete variabele: een variabele die waarden aanneemt uit de verzameling van de natuurlijke getallen (1, 2, 3 ...).

Dummy-variabele: een variabele die de waarde 0 of 1 aanneemt.

Effect (eerste betekenis): uitwerking, gevolg.

Effectiviteit: doeltreffendheid

Herhaalde cross-sectionele data: De data bevatten metingen van verschillende steekproeven op verschillende tijdstippen.

Impact: kracht die van iets uitgaat. Synoniemen: invloed, inwerking, draagwijdte.

Instrumentele variabele: een variabele waarvan men veronderstelt dat deze, vanuit conceptueel standpunt, geen impact kan hebben op de verklaarde variabele, maar wel op de verklarende variabele.

Onafhankelijke of verklarende variabele: zie afhankelijke variabele

Paneeldata: Paneeldata bevatten gegevens van een populatie (of steekproef van deze populatie) voor meer dan één periode. Het verschil tussen paneeldata en Herhaalde cross-secties bestaat erin dat in paneeldata men tracht dezelfde steekproef op te volgen in elke periode waarin data worden verzameld.

Parameter: kwantificeert de relatie tussen een onafhankelijke en afhankelijke variabele.

Populatie: een groep van subjecten (individuen, bedrijven, steden...) waarop de studie betrekking heeft.

Populatiemodel: een model dat een bepaalde wetmatigheid beschrijft, de causale relatie tussen een aantal onafhankelijke variabelen enerzijds, en de afhankelijke variabele anderzijds.

Proxy variabelen: variabelen waarvan men verwacht dat ze in hoge mate gecorreleerd zijn met het te meten fenomeen.

Pseudo paneeldata: paneeldata die worden geconstrueerd door informatie van herhaalde cross-secties te aggregeren.

Regressie(analyse): men probeert aan de hand van beschikbare data het populatiemodel te schatten.

Steekproef: een (al dan niet) willekeurige trekking van subjecten uit de populatie.

Tijdreeks: Men observeert een variabele van één subject voor een heel aantal periodes. Bij meervoudige tijdreeksen observeert men meerdere variabelen voor deze subject.

Variabele: een kenmerk dat kan veranderen tussen subjecten van een populatie, of dat kan veranderen voor eenzelfde subject over de tijd.