



Data assessment Toegepaste Onderzoek Organisaties (TO2), 2012-2019

Tommy Span

CBS Den Haag
Henri Faasdreef 312
2492 JP Den Haag
Postbus 24500
2490 HA Den Haag
+31 70 337 38 00
www.cbs.nl

projectnummer

000486
DBD
7 augustus 2020

1. Inleiding

Het ministerie van Economische Zaken en Klimaat (EZK) heeft het Centraal Bureau voor de Statistiek (CBS) gevraagd een data assessment op te stellen ten behoeve van de evaluatie van Toegepaste Onderzoek Organisaties (TO2) over de periode 2016-2019. Dit data assessment maakt inzichtelijk welke microdata bij het CBS beschikbaar zijn en benut kunnen worden voor de evaluatie.

1.1 Aanleiding

De organisaties voor toegepast onderzoek spelen een centrale rol in de ontwikkeling, verspreiding en benutting van kennis in binnen- en buitenland. Daartoe werken zij samen met bedrijven, overheden en andere kennisinstellingen, onder meer via publiek-private samenwerking, en voeren zij ook onderzoeks- en adviesopdrachten uit voor bedrijven en publieke organisaties.

Het ministerie van EZK heeft bij evaluaties op het terrein van bedrijvenbeleid een traditie opgebouwd met het zo veel mogelijk laten uitvoeren van micro-econometrisch onderzoek met behandel- en controlegroepen voor het meten van de effectiviteit van de beleidsinstrumenten. Het advies van de Commissie Theeuwes vormt hiervoor de grondslag¹. Het is de wens om ook bij de evaluatie van de TO2-instellingen, in aanvulling op het kwalitatieve deel dat het grootste deel van de evaluatie uitmaakt, te proberen te werken met een econometrische onderbouwing.

1.2 Strekking van het data assessment

Het uitgangspunt van dit data assessment is de aanvraag van het ministerie van EZK. Die aanvraag bevat een lijst met de gewenste indicatoren en vormt de basis voor de inventarisatie. Het data assessment heeft tot doel om de voorbereiding van kwantitatieve analyses in de evaluatie soepel te laten verlopen door een overzichtelijk en realistisch beeld te schetsen van welke data binnen het CBS beschikbaar zijn. Tegelijkertijd is een uitvraag bij de TO2-instellingen gedaan om te bezien in welke mate er data van deze organisaties beschikbaar kunnen worden gesteld voor een econometrische analyse op microniveau.

Het data assessment helpt onderzoeksbureaus bij het uitwerken van een onderzoeksaanpak voor de econometrische analyse door een goed inzicht te bieden in de beschikbare data bij het CBS. Voorts dient het data assessment om tijdig in beeld te krijgen welke acties bij het ministerie, het CBS en de TO2-instellingen nodig zijn om te voorzien in een adequate databeschikbaarheid voor de evaluatie.

1.3 Leeswijzer

Dit data assessment is als volgt opgezet. In hoofdstuk 2 worden indicatoren besproken die gebruikt kunnen worden om verschillende effecten van het gebruik van diensten van de TO2-instellingen bij bedrijven te bepalen. Er wordt aangegeven hoe die indicatoren gemeten kunnen worden. Dit hoofdstuk gaat ook in op de datamogelijkheden om controlegroepen samen te stellen voor de evaluatie. Hoofdstuk 3 beschrijft de bronnen van de indicatoren. De bijlage bevat een compleet overzicht van de bronnen en hun kenmerken en frequentietellingen.

¹ Zie het eindrapport van de expertwerkgroep effectmeting, [Durf te Meten](#) uit 2012.

2. Additionaliteit en de vorming van controlegroepen

Databeschikbaarheid is cruciaal voor goede beleidsevaluaties. Dat is één van de voornaamste conclusies van de Expertgroep Effectmeting in haar in 2012 gepubliceerde eindrapportage. Door middel van een gedegen evaluatie kan het effect van een interventie zoals het gebruik van de diensten van TO2-instellingen gemeten worden. Een interventie is als een cyclus waarin meerdere stappen worden doorlopen, van input en de interventie zelf tot de uiteindelijke effecten. In iedere stap van die cyclus kunnen indicatoren gemeten worden, waardoor het mogelijk wordt na te gaan wat de resultaten zijn.

Om de effecten van de interventie en overige omstandigheden van elkaar te scheiden, zijn twee ingrediënten onmisbaar. Ten eerste een **controlegroep** die vergelijkbaar is met de experimentgroep, buiten het daadwerkelijke gebruik van de diensten van de TO2-instellingen. Ten tweede moeten er **data** van voldoende kwaliteit beschikbaar zijn om verschillen tussen deze twee groepen in kaart te brengen. Deze data dienen veelal in tijdreeksen voorhanden te zijn.

Het CBS beschikt over veel data van bedrijven of personen die gebruikt kunnen worden voor beleidsevaluaties. In dit hoofdstuk worden alleen de mogelijkheden om bepaalde indicatoren te meten besproken. Kenmerken van de databronnen, zoals databeschikbaarheid en compleetheid zijn onderdeel van het volgende hoofdstuk en staan opgenomen in de overzichtstabel in bijlage 1.

2.1 Onderzoekspopulatie en eenheden

Deze evaluatie richt zich op de prestaties van de organisaties die gebruik maken van de diensten van de TO2-instellingen en niet op de prestaties van die TO2-instellingen zelf. Het gaat om het doorsijpeleffect van de inspanningen van die TO2-instellingen op bedrijven als samenwerkingspartners en afnemers van onderzoeks- en adviesdiensten².

Afbakening onderzoekspopulatie

Het CBS beschikt niet over een lijst met bedrijven die gebruik hebben gemaakt van de diensten van de TO2-instellingen. De partij die de evaluatie uitvoert, zal deze lijsten zelf beschikbaar moeten stellen in de Remote Access (RA) omgeving van het CBS. Dat kan op twee manieren:

1. Een lijst met Kamer van Koophandel (KVK) nummers of fiscaalnummers die door het CBS wordt gepseudonimiseerd en verder onveranderd in de RA-omgeving van het project wordt geplaatst;
2. Een lijst met KVK- of fiscaalnummers, bedrijfsnaam, adres en postcode die door het CBS wordt gekoppeld aan het ABR en vervolgens in de RA-omgeving van het project wordt geplaatst.

In optie (1) voert het onderzoeksbureau zelf de vertaling van juridische- naar bedrijfseenheid uit, in optie (2) doet het CBS dat. Optie (2) is analoog aan de gestandaardiseerde koppelprocedure die het CBS onder meer jaarlijks hanteert voor de WBSO. Het ministerie van EZK dient de koppeling in dat geval ook te bekostigen en aan te vragen.

² De regels die zijn verbonden aan het gebruik van de microdata laten uitspraken over de vijf TO2-instellingen ook niet toe. De outputrichtlijnen van het CBS vereisen dat geëxporteerde uitkomsten van analyses geaggregeerd zijn en op ten minste tien waarnemingen betrekking hebben. Het is ook niet zonder meer toegestaan om uitspraken over het personeel van de TO2-instellingen te doen omdat het nog steeds het personeelsbestand van minder dan tien organisaties betreft.

2.2 Additionaliteit meten

In het meten van de additionaliteit (ook wel de doeltreffendheid en doelmatigheid) van het beleid worden de bereikte effecten gerelateerd aan de ingezette middelen. Dat kunnen effecten zijn die vrijwel direct volgen op de interventie of effecten die juist daar weer op volgen en dus van een hogere orde zijn. Het kan bijvoorbeeld zijn dat bedrijven die samenwerken met een TO2-instelling vervolgens hogere uitgaven aan R&D doen en dat zij daardoor innovatiever en productiever worden.

Tabel 2.1 geeft een overzicht van de kwantitatieve effectindicatoren die relevant kunnen zijn voor de evaluatie van de TO2-instellingen. Per indicator wordt in deze paragraaf beschreven hoe deze gemeten zou kunnen worden met behulp van microdata die beschikbaar zijn (of komen) binnen het CBS. Relevante microdata die beschikbaar zijn bij de TO2-instellingen blijven in dit data assessment buiten beschouwing en zijn onderdeel van een apart data assessment door die instellingen.

Tabel 2.1 Outputindicatoren organisaties die gebruik maken van TO2-instellingen naar orde van het effect

Eerste-orde-effect

- Hogere eigen R&D-uitgaven
- Meer eigen R&D-uren
- Hogere R&D-uitgaven aan derden
- Meer samenwerking in R&D

Tweede-orde-effect

- Meer innovativiteit

Derde-orde-effect

- Hogere toegevoegde waarde
- Hogere omzet
- Hogere productiviteit
- Hogere werkgelegenheid

Bron: Aanvraag van data assessment bij CBS ten behoeve van de in 2020-2021 uit te voeren evaluatie van TO2-instellingen van het ministerie van EZK. De indeling naar orden van effect is toegevoegd door het CBS.

Eerste-orde-effecten

Bedrijven kunnen op drie manieren een beroep doen op de diensten van de TO2-instellingen:

1. Contractonderzoek in opdracht van bedrijven (100 procent financiering door bedrijf);
2. Publiek-private samenwerking (PPS) in precompetitief onderzoek (gedeeltelijke financiering door bedrijf);
3. Adviesprojecten in opdracht van bedrijven (100 procent financiering door bedrijf), vooral op basis van reeds beschikbare kennis bij de TO2-instelling.

Gemeenschappelijk bij deze drie manieren van dienstverlening door een TO2-instelling is dat er kennis wordt gegenereerd die door bedrijven is te benutten bij innovatie. De gegenereerde kennis is als een eerste-orde-effect te beschouwen dat niet met bepaalde indicatoren te meten is. Aan het genereren van de betreffende kennis zijn wel deels R&D-inspanningen verbonden. Tevens kan de benutting van de gegenereerde kennis in innovatieprojecten van bedrijven extra impulsen geven aan de R&D-inspanningen van bedrijven. Extra R&D-inspanningen die langs directe en indirecte weg voortvloeien uit het genereren van kennis kunnen eveneens beschouwd worden als eerste-orde-effecten van de dienstverlening van de TO2-instellingen aan bedrijven.

Bij contractonderzoek door TO2-instellingen in opdracht van bedrijven is sprake van een initiële impuls bij de uitgaven van bedrijven aan R&D door derden. Bij publiek-private samenwerking van bedrijven met TO2-instellingen vindt een initiële impuls plaats bij de uitgaven aan eigen R&D bij bedrijven. Bij adviesprojecten van TO2-instellingen in opdracht van bedrijven is geen sprake van een initiële impuls bij de uitgaven van bedrijven aan R&D door derden of de uitgaven van bedrijven aan eigen R&D. Terwijl initiële impulsen aan de uitgaven van bedrijven aan eigen R&D slechts bij een van de drie onderscheiden manieren van dienstverlening van TO2-instellingen zijn te benoemen, geldt bij alle drie de manieren dat kennis die resulteert uit de dienstverlening van TO2-instellingen op indirecte wijze positieve impulsen kan geven aan eigen R&D-activiteiten van bedrijven, via de benutting van die kennis in eigen R&D-projecten van bedrijven.

Op basis van het bovenstaande kunnen vanuit het perspectief van bedrijven twee eerste-orde-effecten op R&D-uitgaven worden onderscheiden:

1. een effect op de uitgaven van bedrijven aan **eigen R&D**;
2. een effect op de uitgaven van bedrijven aan **R&D door derden**.

Daarnaast is nog apart te rekening te houden met het feit dat bij publiek-private samenwerking van bedrijven met TO2-instellingen **samenwerking in R&D** van bedrijven een impuls krijgt. In dat licht is ook een effect op R&D-samenwerking door bedrijven als eerste-orde-effect te beschouwen van de TO2-dienstverlening aan bedrijven, waarbij desgewenst een verbijzondering plaats kan vinden naar samenwerking met onderzoeksinstellingen.

Hogere eigen R&D-uitgaven

De uitgaven aan R&D kunnen met behulp van drie bronnen in beeld gebracht worden:

1. De **R&D-enquête** (RTD) bevat variabelen die de uitgaven aan eigen R&D-activiteiten meten en die aan R&D-activiteiten van derden. Die uitgaven worden bovendien verbijzonderd naar zaken als R&D-personeel, investeringen voor R&D, zoals in gebouwen, laboratoria en grond, en typen onderzoek;
2. De **Community Innovation Survey** (CIS) bevat variabelen die aangeven óf er R&D-onderzoek met eigen personeel (doorlopend of incidenteel) of uitbestede R&D heeft plaatsgevonden. Dit zijn dummyvariabelen die niet de omvang van de uitgaven meten en dus geen bedragen;
3. Bedrijven die aan R&D doen, kunnen hun kosten van ontwikkelings- en onderzoeksprojecten verlagen met behulp van de fiscale stimuleringsregeling **WBSO** (Wet Bevordering Speur- en Ontwikkelingswerk) en voorheen de Research & Development Aftrek (**RDA**). Het speur- en ontwikkelingswerk (S&O) dat in aanmerking komt voor de WBSO (en eerder in aanmerking kwam voor WBSO in combinatie met RDA) betreft technisch wetenschappelijk onderzoek en de ontwikkeling van technisch nieuwe producten, productieprocessen en programmatuur. Deze definitie is smaller dan de gangbare definitie van R&D. In de WBSO-bestanden³ die beschikbaar zijn bij het CBS zijn S&O-uren en bedragen voor de S&O-loonkosten én niet-loonkosten opgenomen⁴.

³ Hier en in de rest van dit stuk geldt dat met deze term eigenlijk WBSO- én RDA-bestanden bedoeld worden. Gebruik van deze regelingen is voor de periode 2012-2015 steeds in één koppelbestand gevangen. In 2016 is de RDA geïntegreerd in de WBSO.

⁴ Niet-loonkosten zijn pas onderdeel van een fiscale stimuleringsmaatregel vanaf 2012 en pas vanaf 2016 geïntegreerd in de WBSO. In de beschrijving van deze databron in hoofdstuk 3 wordt in detail stilgestaan bij de kenmerken.

Hierbij dient in acht genomen te worden dat de RTD en CIS op beperkte steekproeven zijn gebaseerd en dat het microbedrijf (bedrijven met minder dan 10 werkzame personen) buiten de onderzoekspopulatie van deze enquêteonderzoeken vallen. Koppeling met de gebruikers van de TO2-instellingen kan om die reden in de praktijk beperkt resultaat opleveren. Zie bijlage 2 voor een overzicht van de aantallen waarnemingen in die bronnen.

Meer samenwerking in R&D en hogere uitgaven aan R&D door derden

Samenwerking en uitgaven aan R&D door derden kunnen in beeld gebracht worden met behulp van drie bronnen die beschikbaar zijn of komen binnen het CBS:

1. De **RTD** bevat variabelen die aangeven of bedrijven R&D-uitgaven aan derden of inkomsten uit R&D-activiteiten aan derden hebben. Dit kan dienen als een proxy voor samenwerking (manier 2 waarop bedrijven een beroep kunnen doen op TO2-instellingen), maar omvat ook ingekochte R&D of R&D verleend buiten een samenwerkingsverband om (manier 1). Er wordt een onderscheid gemaakt tussen binnen- en buitenlandse partijen;
2. De **CIS** bevat variabelen die samenwerking in beeld brengen, met een onderscheid naar verschillende typen bedrijven en kennisinstellingen en partijen uit binnen- en buitenland. Het gaat hier om samenwerking in technologische innovatie. Dat vormt een benadering van samenwerking in R&D;
3. Privaat-publieke samenwerkingsverbanden en Topconsortia voor Kennis en Innovatie (TKI's) kunnen bij RVO.nl een aanvraag indienen voor **PPS-projecttoeslag**. Die toeslag bestaat uit een opslag van 30 procent door het ministerie van EZK bovenop de bedragen die bedrijven in onderzoeksorganisaties investeren. In de uitvoeringsadministratie van de PPS-toeslag staan de bedrijven en onderzoeksorganisaties per samenwerking geregistreerd. Vanzelfsprekend bevat deze alleen de bedrijven en onderzoeksorganisaties die gebruik hebben gemaakt van de toeslag en alleen samenwerkingsverbanden waar ten minste één onderzoeksorganisatie aan deelneemt. Samenwerkingsverbanden met daarin alléén bedrijven ontbreken dus.

De hoeveelheid en kwaliteit van de samenwerking zijn lastig te meten met de hierboven beschreven data. De RTD bevat wel bedragen en een procentuele verdeling van de uitgaven of inkomsten aan activiteiten voor derden. De CIS bevat alleen een dummyvariabele die aangeeft of samenwerking heeft plaatsgevonden en met wat voor partij. De administratie van de PPS-toeslag bevat wel het budget, de omvang van de toegekende subsidie en de totale omvang van de projectkosten.

Tweede-orde-effecten

Extra uitgaven aan en samenwerking in R&D als gevolg van de dienstverlening door TO2-instellingen kunnen op hun beurt tot meer innovatie bij bedrijven leiden. Dit kan rechtstreeks lopen via de benutting van kennis die de dienstverlening door TO2-instellingen oplevert. Deels kan het ook lopen via extra uitgaven van bedrijven aan eigen R&D bij de benutting van de kennis.

Meer innovatie

De voornaamste bron van statistieken over innovatie binnen het CBS is de CIS. Deze bevat meerdere maatstaven van innovatie die ook de ontwikkeling van de inspanningen in en resultaten van innovatie meten, waaronder:

- Goederen die nieuw zijn voor het bedrijf;
- Diensten die nieuw zijn voor het bedrijf;
- De wijze van ontwikkeling van goederen- en diensteninnovaties (zelf, samen, uitbesteed);

- De mate van innovativiteit (nieuw voor bedrijf, nieuw voor de markt, beperkte verandering, omzet uit nieuwe en verbeterde producten);
- Type procesinnovatie (productie, logistiek, ondersteunende activiteiten);
- Uitgaven aan innovatie.

Een andere maatstaf van de innovativiteit van bedrijven en instellingen zijn octrooien. De CIS meet of respondenten een aanvraag hebben gedaan voor patenten, een industrieel ontwerp of een handelsmerk hebben gedeponeerd. Dit is eveneens een dummyvariabele die dus alleen aangeeft óf een organisatie dit gedaan heeft. Het CBS beschikt ook over een niet-actuele reeks microdata met aanvragen van patenten bij het European Patent Office (EPO) en Agentschap NL (nu RVO), divisie NL Octrooiencentrum (NLOC).

Derde-orde-effecten

Als instellingen meer en beter aan R&D en innovatie doen, zou dit uiteindelijk ook door kunnen werken in de productiviteit en werkgelegenheid van organisaties die deze activiteiten ondernemen.

Hogere toegevoegde waarde en omzet

De Productiestatistieken (PS) bevatten gegevens over de opbrengsten, kosten en werkgelegenheid van het bedrijfsleven in Nederland en kunnen gebruikt worden om de ontwikkeling van toegevoegde waarde en omzet te meten. Voor kleinere bedrijven zijn de PS gebaseerd op een procentueel steeds kleinere steekproef. Waar de PS tekortschieten, kan ook de Statistiek Financiën niet-Financiële Ondernemingen (SFO) gebruikt worden. De SFO omvat alleen ondernemingen met rechtspersoonlijkheid. De omzet van bedrijven kent een betere vulling in het CBS-interne bestand Directe Ramingen Totaal (DRT), waarin de omzet uit drie verschillende bronnen wordt gecombineerd. De DRT zijn niet beschikbaar voor extern gebruik, maar derden kunnen wel een maatwerkverzoek voor het samenstellen van een onderzoeksbestand doen.

Hogere productiviteit

Om de arbeidsproductiviteit te berekenen zijn twee indicatoren per bedrijf nodig: de toegevoegde waarde (of de productiewaarde) en de werkgelegenheid in fte. Voor de arbeidsproductiviteit van werknemers kunnen de volgende bronnen gebruikt worden:

1. Werkgelegenheid van werknemers (inclusief directeur-grotaandeelhouders, dga's) in fte. Deze komt uit de Polisadministratie;
2. De toegevoegde waarde komt voor in PS. Voor kleinere bedrijven zijn die gebaseerd op een steekproef. Waar de PS tekortschieten, kan de toegevoegde waarde aangevuld worden met afleidingen van de balans- en resultatenrekening uit de SFO.

Deze aanpak laat de arbeidsproductiviteit van zelfstandige ondernemers buiten beschouwing. Het CBS beschikt niet over een bron die de inzet van zelfstandige ondernemers in fte meet. Dat betekent dat de arbeidsproductiviteit die voortkomt uit het combineren van de twee bovenstaande bronnen een incompleet beeld van die productiviteit geeft. Het leidt bovendien tot een overschatting, omdat de toegevoegde waarde niet aan alle fte gerelateerd wordt.

2.3 Vorming van controlegroepen

Om een goede effectmeting uit te voeren is het van groot belang het effect van een interventie toe te kunnen schrijven aan die specifieke interventie en los te weken van andere invloeden. Daarvoor moet de groep bedrijven en kennisinstellingen die beïnvloed wordt door een interventie (interventie- of experimentgroep) vergeleken kunnen worden met een groep die zo vergelijkbaar mogelijk is, maar de

interventie niet gebruikt hebben (controlegroep). De interventiegroep bestaat in dit geval uit bedrijven die gebruik hebben gemaakt van de diensten van de TO2-instellingen en de controlegroep uit vergelijkbare bedrijven die dat niet hebben gedaan.

Selectiviteit in de vergelijking

Er zijn verschillende manieren om een controlegroep samen te stellen. Ten eerste is het mogelijk alle bedrijven te selecteren die tot een bepaalde doelgroep behoren maar het instrument niet gebruikt hebben. Deze aanpak kan selectiviteit introduceren. Het is mogelijk immers niet voor niets dat bedrijven of instellingen geen beroep hebben gedaan op de TO2-instellingen. In dat geval kan dit leiden tot een onderschatting van het effect als deze organisaties toch al wel samenwerking in R&D of hoge uitgaven aan R&D hebben zonder hulp van de TO2-instellingen.

Daarnaast geldt dat het lastig is om een doelgroep af te bakenen van bedrijven voor wie TO2-diensten relevant zijn en die vervolgens met beschikbare microdata te identificeren. Als het om R&D-uitvoerende bedrijven zou gaan, zouden bedrijven die gebruikmaken van de WBSO tot de doelgroep gerekend kunnen worden. Die afbakening zou te beperkend zijn. Eerder is een bredere groep van potentieel innovatieve bedrijven als relevante doelgroep te beschouwen. Feitelijke innovatie wordt wel gemeten met de CIS, maar alleen bij bedrijven met ten minste 10 werkzame personen en met een beperkte steekproefdekking van bedrijven in hogere grootteklassen.

Door binnen de econometrische schattingen rekening te houden met *fixed effects* kan relatief eenvoudig worden gecontroleerd voor statische verschillen tussen bedrijven. Daarmee kan in statisch opzicht het probleem van selectiviteit worden ondervangen. Of er dan nog verschillen in dynamisch perspectief zijn, hangt af van de vraag of een *common trend* van toepassing is⁵.

Technieken om een willekeurige verdeling na te bootsen

Ten tweede is het mogelijk om natuurlijke experimenttechnieken te gebruiken om een controlegroep samen te stellen. Voorbeelden zijn propensity score matching (PSM) en het regression discontinuity design (RDD). PSM vormt koppeltjes van organisaties op basis van hoe vergelijkbaar zij zijn op een set indicatoren die bepalend is voor het gebruik van de instrumenten en de uitkomsten van dat gebruik. Organisaties die op de gehanteerde indicatoren niet vergelijkbaar zijn met de interventiegroep worden zo uitgesloten van de effectmeting. RDD beperkt de controlegroep tot bedrijven die zich net onder een zeker afkappunt bevinden. Het idee is dat organisaties die *nét* niet gebruik mogen maken van één van de instrumenten heel vergelijkbaar zijn met organisaties die net wel mee mogen doen. De technieken introduceren een vergelijkbaarheid tussen de bedrijven die de *common trend* aanname kan helpen rechtvaardigen.

Voor RDD is een schaalvariabele nodig waarop een punt valt aan te wijzen waarop organisaties nog net wel of net niet gebruik maken van de diensten van de TO2-instellingen. Bij de aanvraag van een subsidie zou dit bijvoorbeeld de score op de aanvraag kunnen zijn. Het afkappunt moet dan door een externe factor bepaald worden, bijvoorbeeld de uitputting van het budget. Van deze constructie is geen sprake bij de TO2-instellingen. In lijn met eerder onderzoek naar de effectiviteit van een TO2-instelling (TNO, 2019) lijkt PSM daarom meer kansrijk⁶.

⁵ Zie in dit verband het rapport *Durf te meten* van de Commissie Theeuwes; <https://zoek.officiëlebeelden.nl/blg-195283>, blz. 14-15.

⁶ Dat onderzoek is gepubliceerd in ESB (H. Boonman e.a., 2019), *Samenwerken met TNO heeft positieve impact op bedrijven*, ESB, 104(4779) en in een achterliggend onderzoeksrapport van TNO (E. Poliakov e.a. (2019) *A microeconomic assessment of RTO's impact on firms output: the case of TNO*, TNO Working Paper 2019-01.

Databehoeften PSM

Propensity Score Matching kan ingezet worden om binnen een brede groep bedrijven tot goed vergelijkbare controlegroepen te komen. Per bedrijf in de interventiegroep wordt een bedrijf (of een groepje bedrijven) gevonden dat op alle relevante en meetbare kenmerken zo vergelijkbaar mogelijk is. Matching gebeurt aan de hand van propensity scores: scores die de kans vertegenwoordigen dat een bedrijf gebruik maakt van een TO2-instelling, los van het feit of het dit ook daadwerkelijk doet. Die scores brengen alle dimensies waarop bedrijven vergeleken worden terug tot één dimensie (de propensity score) waardoor bedrijven gemakkelijker met elkaar vergeleken kunnen worden. De propensity scores worden geschat voor ieder bedrijf waarvoor voldoende data voor de diverse in aanmerking te nemen indicatoren aanwezig zijn in de te gebruiken databases: inclusief de bedrijven die gebruik hebben gemaakt van de TO2-instellingen. De data die gebruikt worden moeten aan twee voorwaarden voldoen:

1. Ze moeten van invloed zijn op de kans dat een bedrijf gebruik maakt van een TO2-instelling én op de uitkomstvariabelen van het onderzoek;
2. Ze mogen niet beïnvloed zijn door de interventie.

Als er geen eenduidige doelgroep aangewezen kan worden waarde controlegroep van kan worden afgeleid, bestaat het risico dat PSM tot een te brede controlegroep leidt. De keuze voor variabelen die gebruikt worden om koppeltjes te maken, is dan extra belangrijk. Dit en bovenstaande leiden tot de volgende databehoeften voor het gebruik van Propensity Score Matching:

- a. Goed meetbare indicatoren die van invloed zijn op het gebruik van een TO2-instelling en de uitkomstvariabelen. Als dit voor een belangrijk deel bepaald wordt door kwalitatieve indicatoren, is de controlegroep nooit echt vergelijkbaar. Omdat bij het gebruik van TO2-instellingen geen eenduidige doelgroep is aan te wijzen, zouden in de ideale situatie ook indicatoren opgenomen worden die de bereidheid tot R&D en innovatie van bedrijven meten;
- b. Deze variabelen moeten voor een groot deel van de bedrijven in de experimentgroep en de daarbij aansluitende potentiële controlegroep beschikbaar zijn. Bedrijven waarvoor ze dat niet zijn, blijven in de schatting van de propensity scores buiten beschouwing;
- c. Variabelen die beschikbaar zijn voor een verslagjaar voorafgaand aan het jaar waarin gebruik wordt gemaakt van de TO2-instellingen;
- d. Onder andere om voorgenoemde reden zijn gegevens nodig die verder teruggaan dan de evaluatieperiode. Een langere datareeks kan bovendien helpen de *common trend* aanname te verantwoorden door de ontwikkeling van kenmerken en prestaties te vergelijken vóór het gebruik van een TO2-instelling. In de effectmeting kan het een vereiste zijn om effecten over een langere periode te bepalen;
- e. Idealiter is ook voor de jaren voorafgaand aan de evaluatieperiode bekend of bedrijven gebruik hebben gemaakt van een TO2-instelling. Zo kan voorkomen worden dat de invloed van eerder gebruik van een TO2-instelling het resultaat beïnvloedt.

Matchingvariabelen

De kwaliteit van de controlegroep wordt mede bepaald door de variabelen die gebruikt worden om de propensity score te berekenen. Daarvoor kunnen gegevens over bedrijfskenmerken, bedrijfsprocessen en bedrijfsprestaties worden gebruikt. Voor het gebruik van continue variabelen, zoals de kapitaalintensiteit, wordt het aangeraden om ook een kwadratische term op te nemen. Op die manier kan ook voor niet-lineaire verbanden gecontroleerd worden. Het is ook mogelijk om “hard” te matchen op enkele variabelen, bijvoorbeeld om te zorgen dat bedrijven in de industrie alleen aan andere industriële bedrijven worden gekoppeld en niet aan een dienstverlenend bedrijf.

Tabel 2.2 geeft een niet-uitputtend overzicht van variabelen die gebruikt zouden kunnen worden, onder andere gebaseerd op het eerdergenoemde onderzoek van TNO. Als matchingvariabelen kunnen ook de (vertraagde) outputindicatoren gebruikt worden (zie tabel 2.1). Daarbij geldt dat variabelen uit steekproefbronnen een beperkende factor kunnen zijn, omdat de propensity score alleen berekend kan worden voor bedrijven die op alle matchingvariabelen een waarneming hebben.

Tabel 2.2 Selectie van matchingindicatoren

Bedrijfskenmerken	
• Bedrijfstak	• Grootteklasse
• Leeftijd	• Zelfstandig mkb
Bedrijfsproces en -prestaties	
• Investerings in vaste activa	• Kapitaalgoederenvoorraad
• Arbeidsproductiviteit	• S&O-(niet)-loonkosten en -uren
• Innovativiteit	

Bron: CBS en Aanvraag van data assessment bij CBS ten behoeve van de in 2020-2021 uit te voeren evaluatie van TO2-instellingen van het ministerie van EZK.

3. Databronnen

In dit hoofdstuk wordt beschreven welke data bruikbaar en beschikbaar zijn om de indicatoren die in hoofdstuk 2 genoemd zijn te meten. Er wordt steeds ook aangegeven voor welke jaren de gegevens beschikbaar zijn. De ruggengraat van de bedrijvenstatistiek van het CBS wordt gevormd door het Algemeen Bedrijven Register (ABR).

Tabel 3.1 Bronnenoverzicht indicatoren evaluatie TO2-instellingen, 2012-2019

Indicator	Bron	Periode (zonder breuken)
Eerste-orde-effecten		
Uitgaven aan R&D	RTD	2014-2018*
	CIS	2010-2018**
Samenwerking in R&D	Koppelbestanden WBSO	2004-2019***
	R&D-enquête (RTD)	2013-2018*
	Community Innovation Survey (CIS)	2010-2018**
	Koppelbestanden PPS-toeslag	2013-2019****
Tweede-orde-effecten		
Innovatie	CIS	2010-2018**
	Patentaanvragen	2000-2010
Derde-orde-effecten		
Toegevoegde waarde	Productiestatistiek (PS)	1993-2018
	Statistiek Financiën niet-financiële Ondernemingen (SFO)	2000-2018
Omzet	Aangifte omzetbelasting (btw)	2007-2017
	Directe Raming Totalen (DRT)	2014-2020
	PS	1993-2018
	SFO	2000-2018
Werkgelegenheid (fte)	Polisadministratie	2006-2019
Werkgelegenheidsgroei	Bedrijfsdemografisch Kader (BDK)	2007-2019
Arbeidsproductiviteit	Polisadministratie	2006-2019
	PS	1993-2018
	SFO	2000-2018
Matchingvariabelen		
Bedrijfstak, grootteklasse	ABR	2006-2019
Leeftijd, zelfstandig mkb	BDK	2007-2019
Investerings vaste activa	Investerings in (im)materiële vaste activa bij niet-financiële bedrijven	2006-2018
Kapitaalgoederenvoorraad	SFO	2000-2018

* Er heeft recent een herziening van de RTD plaatsgevonden vanwege aanpassingen aan de Frascati-richtlijnen voor het produceren van R&D-statistiek. Dat leidt onder meer tot een afname van R&D-bedrijven, omdat personen die minder dan 10 procent van hun tijd aan R&D besteden, niet meer worden meegeteld. Een andere aanpassing is dat ingeleend R&D-personeel ook meegenomen moet worden, waardoor de eigen uitgaven aan R&D zullen toenemen. Het microdatabestand voor 2018 zal naar verwachting in oktober 2020 beschikbaar komen in de RA-faciliteit.

** Alleen even jaren. Er zijn ook data beschikbaar vanaf 1996, maar met ingang van verslagjaar 2010 zijn de cijfers ingrijpend veranderd ten opzichte van voorgaande jaren. Het gebruik van innovatiedata van vóór 2010 is dan ook af te raden voor tijdreeksanalyses. Het microdatabestand voor 2018 zal naar verwachting in oktober 2020 beschikbaar komen in de RA-faciliteit.

*** Het koppelbestand voor 2019 komt naar verwachting in oktober beschikbaar in de RA-faciliteit.

**** Het samenstellen van deze koppelbestanden vindt op dit moment nog plaats. Naar verwachting komen zij in oktober beschikbaar in de RA-faciliteit.

Per indicator zijn in de tabel één of meerdere mogelijke bronnen aangegeven. Deze bronnen worden hier stuk voor stuk nader uitgewerkt aan de hand van de beschikbare tijdreeks, een korte beschrijving van de populatie, of het om enquêtegegevens of registers gaat en of de data toegankelijk zijn voor externe onderzoekers. De kenmerken van de CBS-bronnen zijn opgenomen in bijlage 1. In bijlage 2 zijn frequentietellingen opgenomen voor enkele belangrijke variabelen in de beschreven bronnen.

Aangifte omzetbelasting (btw)

De bestanden Aangifte omzetbelasting bevatten gegevens over alle bedrijven (op het niveau van de bedrijfseenheid, BEID) in Nederland die omzetbelasting betalen. Naast de betaalde omzetbelasting bevatten de bestanden ook de omzet waarover deze belasting betaald wordt. De bestanden bevatten echter ook witte vlekken: activiteiten waarvoor in de omzetbelasting een vrijstelling bestaat, zijn uitgesloten. Zo is de dekking van de bestanden in sectoren zoals de landbouw, zorg en onderwijs bijvoorbeeld slecht. Bovendien zijn de bestanden niet bruikbaar voor het bepalen van de financiële positie van bedrijven buiten de omzet. Deze 'btw-bestanden' zijn vanaf 2007 beschikbaar.

Algemeen Bedrijven Register (ABR)

Het Algemeen Bedrijven Register (ABR) vormt voor het CBS de ruggengraat van het statistisch proces voor economische statistieken. Het ABR is een systeem waarin identificerende gegevens en structuurgegevens over alle bedrijven en instellingen (inclusief zelfstandigen) zijn geregistreerd. Hieruit worden de statistische eenheden bedrijfseenheid, ondernemingengroep en lokale bedrijfseenheid afgeleid. Het ABR bevat informatie over de economische activiteit en het aantal werkzame personen. Daarnaast bevat het ABR ook informatie over bepaalde 'events'. Een event geeft een gebeurtenis of wijziging weer binnen het ABR: bijvoorbeeld de oprichting, overname of opheffing van een bedrijf. De Regionale eenheden base (ABR-regio) is een verrijking van het ABR, met informatie over de vestigingsplaats.

Bedrijfsdemografisch kader (BDK)

Het BDK is een verrijking van het ABR met meer bedrijfskenmerken en samenhang tussen bedrijven, waardoor bedrijven beter door de tijd gevolgd kunnen worden.

Community Innovation Survey (CIS) - Innovatie statistiek

De Community Innovation Survey is een enquête over innovatieve inspanningen en innovatieresultaten van bedrijven, die iedere twee jaar wordt gehouden onder de lidstaten van de Europese Unie. Het CBS voert de enquête uit bij in Nederland gevestigde bedrijven met 10 of meer werkzame personen. De belangrijkste doelpopulatie bestaat uit de industrie- en dienstensector. De steekproef wordt getrokken uit het ABR en omvat ongeveer 15 duizend bedrijven. De grote ondernemingen zijn allemaal in de steekproef opgenomen voor zover het de R&D-uitgaven betreft. Er wordt niet bijgeschat voor bedrijven met minder dan 10 werkzame personen. Een consistente reeks is beschikbaar vanaf 2010.

Directe Raming Totalen (DRT)

De Directe Raming Totalen (DRT) combineren omzetgegevens uit drie verschillende bronnen. Voor de 350 grootste bedrijven komen de gegevens uit CONGO (consistentie grote ondernemingen). Voor de volgende 1.950 grootste bedrijven (Top-X) zijn deze gebaseerd op primaire dataverzameling, bijvoorbeeld uit de Productiestatistieken. Voor de rest van het bedrijfsleven dat niet voorkomt in

steekproeven van enquêtes worden btw-bestanden gebruikt. De data zijn bovendien per kwartaal beschikbaar.

De gegevens zijn pas vanaf 2014 beschikbaar. In de daaraan voorafgaande jaren is de DRT beschikbaar, maar aan minder verstreckende kwaliteitscontroles onderworpen. Wel zijn de omzetcijfers zeer actueel: er zijn momenteel al eerste voorlopige cijfers over het tweede kwartaal van 2020. De DRT is niet direct beschikbaar als microdatabestand voor externe gebruikers, maar analyses kunnen op verzoek door het CBS uitgevoerd worden.

Investerings in (im)materiële vaste activa bij niet-financiële bedrijven

Cijfers van investeringen gedaan door niet-financiële bedrijven. Onderscheid wordt gemaakt naar soorten vaste activa en naar nieuwe of bestaande activa. Onder materiële vaste activa worden onder andere verstaan gebouwen, grond, transportmiddelen en machines van een bedrijf. Immateriële vaste activa bestaan onder andere uit software, licenties, goodwill.

Patentaanvragen

Deze statistiek betreft twee microdatabestanden waarin patentaanvragen bij het European Patent Office (EPO) en Agentschap NL (nu RVO.nl), divisie NL Octrooiencentrum (NLOC) in de periode 2008-2010 zijn gerelateerd aan ondernemingengroepen uit het ABR van het CBS. Een patent geeft de houder (individueel of instituut) een aantal rechten op het gebied van exclusiviteit met betrekking tot de uitvinding. De wettelijke bescherming houdt in dat de eigenaar van het patent het recht heeft anderen uit te sluiten voor productie, gebruik, (aanbieden voor) verkoop of het importeren van de gepatenteerde uitvinding.

Een patent toegekend door het EPO is geldig in elk van de landen die lid zijn van het EPO (38 leden in 2014), mits de aanvrager zijn rechten in een land heeft gevalideerd. Een patent toegekend door het NLOC is alleen geldig in Nederland. Op basis van patentaanvragen kunnen aan innovatie gerelateerde indicatoren worden samengesteld. Sommige van de gegevens die wel uit patentregistraties te destilleren zijn, kunnen niet of moeilijk door andere indicatoren in kaart worden gebracht.

Polisadministratie

De Polisadministratie bevat gegevens over banen en is gebaseerd op data uit de loonaangiften van de Belastingdienst. De loonaangiften bevatten gegevens over inkomstenverhoudingen (uit de loonadministratie) van werkgevers en andere inhoudingsplichtigen. Het doel van de Polisadministratie is inzicht te krijgen in arbeidscontracten en loon van werknemers. De Polisadministratie van het UWV bevat naast kenmerken van de baan eveneens enkele gegevens over het bedrijf, waaronder het bedrijfseenheid identificatienummer (BEID). Op basis daarvan is het mogelijk de persoon en diens baan te koppelen aan een bedrijf en, uiteindelijk, de prestaties van dat bedrijf.

Privaat-Publieke Samenwerkingsverbanden (PPS) toeslag

Met de PPS-toeslag beoogt het ministerie van EZK private financiële inleg voor privaat-publieke samenwerking (PPS) tussen bedrijven en onderzoeksorganisaties te stimuleren. De toeslag bestaat uit een opslag van 30 procent bovenop het bedrag dat een bedrijf aan R&D in een onderzoeksorganisatie bijdraagt. Het instrument wordt uitgevoerd door RVO.nl. Bedrijven en kennisinstellingen die aan deze samenwerkingsverbanden deelnemen, zijn bij RVO.nl geregistreerd. Deze microdatabestanden zijn zogeheten restrictiebestanden. Toestemming van het ministerie van EZK en RVO is benodigd om analyses op deze bestanden uit te voeren.

Productiestatistieken (PS)

De Productiestatistieken geven een beeld van de werkgelegenheid en het financiële reilen en zeilen van bedrijven. De PS bevatten onder meer gegevens over omzet, kosten en toegevoegde waarde. Van de volgende bedrijfstakken worden PS samengesteld: landbouw, winning van delfstoffen, industrie, productie en distributie van energie en water, bouwnijverheid, reparatie van consumentenartikelen, groothandel en detailhandel, horeca, vervoer, opslag en communicatie, zakelijke en persoonlijke dienstverlening. De doelpopulatie van de PS bestaat uit de in de verslagperiode economisch actieve bedrijven met de hoofdactiviteit in een van de bovengenoemde bedrijfstakken.

De PS zijn enquêtes onder circa tien procent van de bedrijven. De steekproef is relatief groter onder de grotere bedrijven. Voor de kleinste bedrijven worden gegevens op het totaalniveau van grootteklasse-sector combinaties bijgeschat. Koppeling van gegevens uit de PS aan unieke bedrijven zal daarom zeker voor die kleinere ondernemingen niet altijd resultaten opleveren. De PS worden integraal waargenomen onder bedrijven vanaf 100 werkzame personen en in de meeste SBI-bedrijfstakken's ook vanaf 20 of 50 werkzame personen. De respons bij grote bedrijven is hoog (actief responsbeleid). Bij de kleinste bedrijven (meestal tot 10 personen) wordt er bij een aantal branches slechts eens in de drie jaar waargenomen. Daardoor kan het aantal waarnemingen van jaar tot jaar fluctueren.

Research Technological Development (RTD)

De R&D-enquête vraagt bedrijven gedetailleerd naar hun R&D-inspanningen. Het gaat daarbij niet alleen om uitgaven aan en inkomsten uit R&D, maar ook het R&D-personeel en het aantal bedrijven dat bij R&D betrokken is, vallen hieronder. De industrie- en dienstensector is de belangrijkste doelpopulatie. De steekproef omvat ongeveer vierduizend bedrijven met 10 werkzame personen of meer. De grote ondernemingen (100 of meer werkzame personen) zijn, voor zover het R&D-uitgaven betreft, allemaal in de steekproef opgenomen. De R&D activiteiten van bedrijven met minder dan 10 werkzame personen zijn bijgeschat op basis van het WBSO-register. Uitkomsten uit de R&D-enquête en de bijschatting worden samengevoegd om uitspraak te doen over de R&D-activiteiten van alle bedrijfsgrootten samen.

Statistiek Financiën niet-financiële Ondernemingen (SFO)

Het financiële reilen en zeilen wordt beschreven aan de hand van de balansen en resultatenrekeningen van in principe alle rechtspersoonlijkheid bezittende ondernemingengroepen (moedermaatschappij inclusief de bijbehorende groepsmaatschappijen) die in Nederland actief zijn in de niet-financiële sector. De SFO wordt samengesteld uit twee bronnen, enerzijds uit primaire waarneming (enquête) bij de grote ondernemingen en anderzijds uit secundaire bron (de vennootschapsbelastinggegevens van de belastingdienst) voor de kleine ondernemingen.

Wet Bevordering Speur- en Ontwikkelingswerk (WBSO) en Research & Development Aftrek (RDA)

Nederland kent sinds 1994 de WBSO als fiscale R&D-stimuleringsregeling voor bedrijven. Op basis van de WBSO kunnen bedrijven aanspraak maken op een aftrek (afdrachtvermindering) van de te betalen loonheffingen. Basis voor dit belastingvoordeel vormen de S&O-loonkosten van eigen personeel.

Tot 2012 was een belastingvoordeel op S&O-loonkosten (zelfstandigen: S&O-uren) via de WBSO de enige aftrekmogelijkheid voor een S&O-inspanning. In 2012 is een niet-looncomponent van S&O toegevoegd via een aparte regeling, de RDA. Dit betrof een aftrek binnen de vennootschapsbelasting voor bedrijven en voor zelfstandigen een aftrek binnen de inkomstenbelasting. De aftrek was gebaseerd op een bedrag aan kosten en uitgaven (met uitzondering van loonkosten) die direct toerekenbaar zijn aan eigen verricht S&O-werk. Dit omvat materiële kosten en uitgaven aan nieuwe

bedrijfsmiddelen. Ondernemers met gemiddeld minder dan 150 S&O-uren per maand konden binnen de RDA ook kiezen voor een forfait; een vastgesteld bedrag per S&O-uur.

In 2016 is de RDA samengevoegd met de WBSO. Sindsdien wordt binnen de WBSO ook een belastingvoordeel (afdrachtvermindering) verstrekt voor kosten en uitgaven ten behoeve van eigen S&O-werk. In de nieuwe WBSO-regeling geldt niet langer dat alleen ondernemingen met minder dan 150 S&O-uren gemiddeld per maand kunnen kiezen voor een forfaitair bedrag. In de nieuwe regeling hebben alle ondernemingen die optie, met daarbij een tweeschijvenregime.

Voor bedrijven die gebruik hebben gemaakt van het forfait, zijn geen gegevens beschikbaar over de niet-looncomponent van de S&O-uitgaven. Die bedrijven hebben een voordeel genoten afhankelijk van het aantal S&O-uren en hebben in verband daarmee geen gegevens over S&O-kosten of -uitgaven buiten de opgegeven loonkosten.

Deze microdatabestanden zijn zogeheten restrictiebestanden. Toestemming van het ministerie van EZK en RVO is benodigd om analyses op deze bestanden uit te voeren.

Bijlage 1. Kenmerken bronbestanden

Tabel B.1 Kenmerken van data beschikbaar binnen het CBS.

Bron	periode (zonder breuk)	populatie en bijzonderheden	type data	kennisinstellingen	toegankelijkheid
ABR	2006-2019	Alle in Nederland geregistreerde bedrijven en instellingen (incl. zelfstandigen).	Integraal	Ja	Beschikbaar in RA
BDK	2007-2019	Alle in Nederland geregistreerde bedrijven en instellingen (incl. zelfstandigen).	Integraal	Ja	Beschikbaar in RA
Btw	2007-2017	Statistische eenheden waarvan informatie uit de omzetbelasting beschikbaar is, exclusief eenheden met vrijstelling.	Integraal	Ja	Beschikbaar in RA
CIS	2010-2018*	Bedrijven in SBI 01 t/m 82 en vanaf 10 werkzame personen (geen bijschatting voor kleinere bedrijven).	Steekproef	Geen hoger onderwijs of UMC's, wel overige kennisinstellingen	Beschikbaar in RA
DRT	2014-2020	Actieve bedrijfseenheden met als hoofdactiviteit groothandel en handelsbemiddeling, vervoer en opslag, horeca, informatie en communicatie, specialistische zakelijke diensten, overige zakelijke dienstverlening en overige diensten.	Analysebestand op basis van verschillende bronnen van omzetgegevens	Geen hoger onderwijs of UMC's, wel overige kennisinstellingen	CBS-intern. Derden kunnen een verzoek voor maatwerk doen.
Investerings	2012-2018	Alle bedrijven in SBI-range B tot en met J en L tot en met N.	Steekproef tot 20 wp, daarna integraal. Bedrijven buiten de steekproef zijn geïmputeerd.	Nee	Beschikbaar in RA
Patenten	2008-2010	Alle patentaanvragen ingediend door Nederlandse aanvragers bij het EPO of het NLOC.	Integraal	Ja	Beschikbaar in RA
Polis	2006-2019	Banen van werknemers in Nederland.	Integraal	Ja	Beschikbaar in RA
PPS	2013-2019**	Nederlandse en buitenlandse bedrijven en onderzoeksorganisaties die gebruik maken van de PPS- toeslag.	Integraal	Ja	Beschikbaar in RA met toestemming EZK en RVO.nl

PS	1993-2018	Bedrijven in: landbouw, winning van delfstoffen, industrie, productie en distributie van energie en water, bouwnijverheid, reparatie van consumentenartikelen, groothandel en detailhandel, horeca, vervoer, opslag en communicatie, zakelijke en persoonlijke dienstverlening, milieudienstverlening en gezondheids- en welzijnszorg.	Steekproef	Nee	Beschikbaar in RA
RTD	2014-2018***	In Nederland gevestigde bedrijven met ten minste 10 werkzame personen worden geënkquêteerd, met de industrie en dienstensector als belangrijkste doelgroep. Bedrijven in grootteklassen 0 t/m 3 worden bijgeschat op basis van gegevens van de WBSO.	Steekproef	Geen hoger onderwijs of UMC's, wel overige kennisinstellingen	Beschikbaar in RA
WBSO en RDA	2004-2019****	Alle gebruikers met vastgestelde WBSO of RDA.	Integraal.	Nee	Beschikbaar in RA met toestemming EZK en RVO.nl

* Alleen even jaren. Er zijn ook data beschikbaar vanaf 1996, maar met ingang van verslagjaar 2010 zijn de cijfers ingrijpend veranderd ten opzichte van voorgaande jaren. Het gebruik van innovatiedata van vóór 2010 is dan ook af te raden voor tijdreeksanalyses. Het microdatabestand voor 2018 zal naar verwachting in oktober beschikbaar komen in de RA-faciliteit.

** Het samenstellen van deze koppelbestanden vindt op dit moment nog plaats. Naar verwachting komen zij in oktober beschikbaar in de RA-faciliteit.

*** Er heeft recent een herziening van de RTD plaatsgevonden vanwege aanpassingen aan de Frascati-richtlijnen voor het produceren van R&D-statistiek. Dat leidt onder meer tot een afname van R&D-bedrijven, omdat personen die minder dan 10 procent van hun tijd aan R&D besteden, niet meer worden meegeteld. Een andere aanpassing is dat ingeleend R&D-personeel ook meegenomen moet worden, waardoor de eigen uitgaven aan R&D zullen toenemen. Het microdatabestand voor 2018 zal naar verwachting in oktober 2020 beschikbaar komen in de RA-faciliteit.

**** Het koppelbestand voor 2019 komt naar verwachting in oktober beschikbaar in de RA-faciliteit.

Bijlage 2. Frequentietellingen

Dit overzicht dient om een *inschatting* te kunnen maken van het aantal bedrijven/instellingen dat in een beleidsanalyse opgenomen kan worden. Het betreft hier dan ook technische gegevens over de waargenomen **unieke** bedrijfseenheden in een bronbestand. Dit kan afwijken van wat er op StatLine staat over bijvoorbeeld aantallen bedrijven in een jaar, omdat daarbij gewerkt wordt met een specifiek peilmoment. Het daadwerkelijk aantal bedrijven (deelnemende bedrijven aan een instrument en eventuele controlegroep) dat in een statistisch onderzoek kan worden betrokken, wordt vanzelfsprekend pas duidelijk op het moment dat een koppeling tot stand wordt gebracht. Naar verwachting zal dit specifiek voor de RTD en CIS beperkt zijn.

Waar de vulling van de cellen voor een bepaalde variabele afwijkt van het totaal aantal bedrijven dat in de bron is opgenomen, zijn beide tellingen (totaal en respons) opgenomen. Het overzicht omvat de relevante variabelen zoals beschreven in de tabel in paragraaf 3.1. De vulling van andere variabelen in de bronbestanden kan anders zijn.

In het ABR zijn alle bedrijven en instellingen (inclusief zelfstandigen) geregistreerd. De vulling hiervan is dan ook volledig. Ter referentie is de telling van deze bron voor verslagjaar 2016 ook opgenomen. Voor dat jaar is gekozen in verband met de beschikbaarheidsperiode van andere statistieken, zoals de CIS. Deze referentie is dus de meest volledige telling van bedrijven volgens het principe van de bedrijfseenheden. In de frequentietellingen zijn de volgende bronnen niet meegenomen:

- Het Bedrijfsdemografisch Kader is een verbijzondering van het ABR. De dekking van deze bron is daarmee in principe gelijk aan die van het ABR;
- De uitvoeringsadministratie van de PPS-toeslag is nog niet definitief door het CBS aan het ABR gekoppeld. Frequentietellingen voor die bron zijn daarom niet in dit overzicht opgenomen;
- De Polisadministratie bevat alle banen van werknemers in Nederland over een bepaald verslagjaar. Wat betreft de vulling van het aantal banen is deze dus volledig dekkend.

Tabel B.2 Frequentietellingen databronnen, 2016 (vervolg op volgende pagina)

	ABR		RTD ⁷		CIS		Btw	
			R&D-uitgaven		Samenwerking		Omzet	
					innovatie			
	totaal		totaal	respons	totaal	respons	totaal	respons
<i>bedrijfseenheden</i>								
bedrijfstak								
totaal	1 727 245		6 760	2 370	7 340	4 140	1 727 240	1 523 840
A	75 130		80	30	40	20	75 130	57 890
B	465		70	20	70	35	465	420
C	65 185		1 815	1 055	1 995	1 415	65 185	60 970
D	1 225		65	15	70	30	1 225	1 120
E	1 805		110	40	85	45	1 805	1 660
F	166 690		285	55	255	120	166 690	159 520
G	249 005		885	170	1 785	840	249 005	234 090
H	41 820		340	40	535	205	41 820	39 235
I	58 100		125	10	100	25	58 100	55 310
J	92 375		380	195	640	460	92 375	86 825
K	101 025		165	45	180	105	101 020	87 785
L	27 625		55	5	50	20	27 625	25 195
M	344 735		745	325	1 145	680	344 735	324 475
N	72 535		505	65	390	145	72 535	68 155
O	820		170	40	0	0	820	730
P	76 205		145	20	0	0	76 205	65 270
Q	144 465		560	200	0	0	144 465	66 645
R	103 640		150	25	0	0	103 640	94 455
S	104 380		115	25	0	0	104 380	94 065
T	20		0	0	0	0	20	15
U	5		0	0	0	0	5	5
grootteklasse								
totaal	1 727 245		6 760	2 370	7 340	4 140	1 727 240	1 523 840
0-9 wp	1 663 285		160	30	0	0	1 663 280	1 466 885
10-249 wp	61 005		4 545	1 510	5 990	3 285	61 005	54 710
250+ wp	2 955		2 050	830	1 350	855	2 955	2 240

Bron: CBS en RVO.nl

Technische, ongewogen informatie over eenheden. Betekenis bedrijfstakcodes: (A) Landbouw, bosbouw en visserij, (B) Delfstoffenwinning, (C) Industrie, (D) Energievoorziening, (E) Waterbedrijven en afvalbeheer, (F) Bouwnijverheid, (G) Handel, (H) Vervoer en opslag, (I) Horeca, (J) Informatie en communicatie, (K) Financiële dienstverlening, (L) Verhuur en handel van onroerend goed, (M) Specialistische zakelijke diensten, (N) Verhuur en overige zakelijke diensten, (O) Openbaar bestuur en overheidsdiensten, (P) Onderwijs, (Q) Gezondheids- en welzijnszorg, (R) Cultuur, sport en recreatie, (S) Overige dienstverlening, (T) Huishoudens, (U) Extraterritoriale organisaties.

⁷ De cijfers van de RTD hebben betrekking op 2015.

Tabel 3. Frequentietellingen databronnen, 2016 (vervolg)

Investeringen		DRT		PS	WBSO		
		Omzet		Toegevoegde	S&O-(niet)-		
				waarde	loonkosten		
	totaal	totaal	respons	totaal	totaal	respons	
bedrijfseenheden							
bedrijfstak							
	totaal	1 213 675	640 050	604 740	56 505	17 085	16 045
	A	13 840	0	0	540	485	400
	B	465	0	0	290	20	20
	C	65 185	0	0	9 045	3 895	3 820
	D	1 225	0	0	560	45	45
	E	1 805	0	0	470	85	85
	F	166 690	0	0	3 855	470	445
	G	249 010	0	0	18 940	2 040	1 990
	H	41 820	38 750	36 685	3 830	95	95
	I	58 100	54 190	51 435	725	15	15
	J	92 375	85 265	80 680	5 940	3 645	3 340
	K	0	0	0	0	1 985	1 985
	L	27 625	0	0	1 025	55	55
	M	333 820	310 160	294 625	5 425	3 595	3 140
	N	72 535	66 990	63 420	5 350	340	330
	O	0	0	0	0	0	0
	P	0	0	0	0	75	65
	Q	0	0	0	0	110	100
	R	0	0	0	0	65	55
	S	89 190	84 695	77 900	510	60	55
	T	0	0	0	0	0	0
	U	0	0	0	0	0	0
grootteklasse							
	totaal	1 213 675	640 050	604 740	56 505	17 085	16 045
	0-9 wp	1 162 190	618 250	583 220	25 240	9 915	8 880
	10-249 wp	49 785	21 060	20 780	29 615	6 585	6 580
	250+ wp	1 690	740	740	1 645	585	585

Bron: CBS en RVO.nl

Het betreft technische, ongewogen informatie over de waargenomen eenheden.

Betekenis bedrijfstakcodes: (A) Landbouw, bosbouw en visserij, (B) Delfstoffenwinning, (C) Industrie, (D) Energievoorziening, (E) Waterbedrijven en afvalbeheer, (F) Bouwnijverheid, (G) Handel, (H) Vervoer en opslag, (I) Horeca, (J) Informatie en communicatie, (K) Financiële dienstverlening, (L) Verhuur en handel van onroerend goed, (M) Specialistische zakelijke diensten, (N) Verhuur en overige zakelijke diensten, (O) Openbaar bestuur en overheidsdiensten, (P) Onderwijs, (Q) Gezondheids- en welzijnszorg, (R) Cultuur, sport en recreatie, (S) Overige dienstverlening, (T) Huishoudens, (U) Extraterritoriale organisaties.