

Bijlage 11

Overzicht van microdata bij het CBS, ten behoeve van de evaluaties van de MIT- en de IPC-regeling

In deze bijlage van het aanbestedingsdocument wordt een overzicht gegeven van microdata bij het CBS die beschikbaar en potentieel bruikbaar zijn voor kwantitatieve analyse in de evaluaties van de regeling Mkb-Innovatiestimulering (Regio en) Topsectoren (MIT) en de regeling Innovatieprestatiecontracten. Het betreft: a) data voor effectindicatoren, b) data over waarneembare karakteristieken van bedrijven die al dan niet in combinatie met data genoemd onder effectindicatoren benut kunnen worden voor beschrijvende statistiek en c) 'matchingvariabelen', die al dan niet in combinatie met data genoemd onder effectindicatoren kunnen worden voor de vorming van controlegroepen met 'propensity score matching'. Er wordt steeds ook aangegeven voor welke jaren de gegevens naar verwachting beschikbaar zijn bij de start van het evaluatieonderzoek in februari 2023. Het overzicht is op basis van eerdere overzichtsinformatie van het CBS opgesteld door EZK en begin november 2022 afgestemd met het CBS, met name met het oog op een getrouwe weergave van de data die naar verwachting in februari 2023 beschikbaar zijn.

Per indicator zijn in onderstaande tabel 1 één of meerdere mogelijke bronnen aangegeven en de jaren waarvoor deze naar verwachting beschikbaar zijn in februari 2023. Deze bronnen worden vervolgens stuk voor stuk nader uitgewerkt, met onder andere een korte beschrijving van de populatie en de herkomst en karakteristieken van de data. Daar waar van toepassing wordt aangegeven of sprake is van restrictiebanden die alleen na toestemming van EZK/RVO (of andere partijen) beschikbaar kunnen worden gesteld. Dat speelt onder andere indien het gaat om bestanden gebaseerd op microgegevens afkomstig van RVO. In bijlage 1 van dit document worden de kenmerken van de CBS-bronnen overzichtelijk weergegeven, waarbij ook is gespecificeerd of het om steekproefgegevens op basis van enquêtes gaat dan wel om integrale waarneming op basis van registers. In bijlage 2 zijn frequentietellingen opgenomen voor enkele belangrijke variabelen in de beschreven bronnen.

De verbinding tussen de eenheden van de verschillende statistieken vindt plaats op basis van de zogeheten bedrijfseenheid (BEID). Dit is de unieke eenheid die bedrijven identificeert vanuit Algemeen Bedrijven Register (ABR), de ruggengraat van de bedrijvenstatistiek van het CBS. Met behulp van deze bedrijfseenheid (BEID) en in sommige gevallen ondernemingengroep (OG) als aggregaat van BEID's zijn de verschillende bronnen te koppelen. De in het evaluatieonderzoek te gebruiken microdata met betrekking tot toekenningen en afwijzingen bij de MIT en de IPC-regeling worden via KvK-nummers gekoppeld aan het ABR op BEID-niveau. Deze data maken geen deel uit van dit overzicht van microdata bij het CBS, maar zijn op uitgebreide wijze apart beschreven in de bijlagen 9 en 10 van het aanbestedingsdocument, naast besprekingen van die data in meer algemene zin (inclusief tellingen van waarnemingen) in de Terms of References voor de twee evaluaties (bijlagen 1 en 2 van het aanbestedingsdocument).

Tabel 1 Bronnenoverzicht indicatoren

Indicator	Bron	Periode (zonder breuken)
Effectindicatoren		
Uitgaven aan R&D	R&D-enquête (RTD)	2013-2020*
	Koppelbestanden WBSO/RDA	2004-2021
Financiering van R&D door bedrijven bij kennisinstellingen	R&D-enquête (RTD)	2013-2020*
	Koppelbestanden PPS-toeslag (TKI-grondslagdata)	2013-2021**
Omzet met nieuwe en verbeterde producten	Community Innovation Survey (CIS)	2010-2020***
Al dan niet product- of procesinnovatie gerealiseerd door bedrijven (binaire variabele)	Community Innovation Survey (CIS)	2010-2020***
Al dan niet samengewerkt bij innovatie door bedrijven met andere bedrijven (binaire variabele)	Community Innovation Survey (CIS)	2010-2020***
Al dan niet samengewerkt bij innovatie door bedrijven met kennisinstellingen (universiteiten en researchinstellingen onderscheiden; binaire variabelen)	Community Innovation Survey (CIS)	2010-2020***
Mate van samenwerking door bedrijven met kennisinstellingen in TKI-verband	Koppelbestanden PPS-toeslag (TKI-inzetdata)	2013-2021**
Patentaanvragen bij European Patent Office en OCNL	Patentaanvragen	2000-2010****
Toegevoegde waarde	Productiestatistieken (PS)	1993-2020
	Statistiek Financiën niet-financiële Ondernemingen (NFO)	2000-2020
Omzet	Aangifte omzetbelasting (btw)	2007-2021
	Directe Raming Totalen (DRT)	2014-2022
	PS	1993-2020
	NFO	2000-2020
Werkgelegenheid (fte)	Polisadministratie	2006-2022
Werkgelegenheidsgroei	Bedrijfsdemografisch Kader (BDK)	2007-2021
Arbeidsproductiviteit	Polisadministratie	2006-2022
	PS	1993-2020
	NFO	2000-2020

Kenmerken bedrijven, ten behoeve van beschrijvende statistiek		
Bedrijfstak, grootteklasse	ABR	2006-2021
Topsector, indien van toepassing	Topsectorenindeling van CBS in kader van Monitor topsectoren	2010-2014*****
Leeftijd, zelfstandig mkb	BDK	2007-2021
Matchingvariabelen		
Bedrijfstak, grootteklasse	ABR	2006-2021
Leeftijd, zelfstandig mkb	BDK	2007-2021
Investerings vaste activa	Investerings in (im)materiële vaste activa bij niet-financiële bedrijven	2006-2020
Kapitaalgoederenvoorraad	NFO	2000-2020

* In 2019 heeft een herziening van de RTD plaatsgevonden vanwege aanpassingen aan de richtlijnen in de Frascati Manual van de OECD voor de R&D-statistiek: <https://www.cbs.nl/nl-nl/achtergrond/2020/02/research-en-development-revisie-2019>. Dat heeft er onder meer toe geleid dat ingeleend R&D-personeel ook meegenomen wordt bij de R&D-activiteiten en -uitgaven van een bedrijf. Een andere aanpassing is dat personen die minder dan 0,1 vte op jaarbasis aan R&D-activiteiten besteden, niet meer worden meegeteld in de R&D-statistiek. Cijfers volgens de methodiek na revisie zijn beschikbaar vanaf het verslagjaar 2013.

** De PPS-toeslagdata bestaan uit TKI-grondslagdata en TKI-inzetdata. Bij de grondslagdata betreft het grondslagprojecten lopend t/m 2020, waarmee grondslag is gegenereerd in de jaren t/m 2021.

*** 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. In 2020 heeft een revisie van de innovatiestatistiek plaatsgevonden, betrekking hebbend op het verslagjaar 2018, naar aanleiding van aanpassingen van richtlijnen in de Oslo Manual van de OECD voor de innovatiestatistiek. Dit is gepaard gegaan met een herontwerp van de innovatie-enquête (CIS). Inhoudelijk is een belangrijke verandering geweest dat geen onderscheid meer wordt gemaakt tussen technologische en niet-technologische innovatie. In de nieuwe methodiek wordt systematisch van een breed innovatiebegrip uitgegaan, waarin technologische en niet-technologische innovatie zijn geïntegreerd. Voor nadere info is te raadplegen: <https://www.cbs.nl/nl-nl/publicatie/2020/42/ict-kennis-en-economie-2020>, paragraaf 6.1; <https://www.bedrijvenbeleidinbeeld.nl/bouwstenen-bedrijvenbeleid/innovatie/innovatie-activiteiten>.

**** Het CBS heeft een nieuwe octrooistatistiek ontwikkeld, met data lopend van 2013 t/m 2019. De microdata komen naar verwachting in april 2023 beschikbaar voor externe onderzoekers in de RA-omgeving van het CBS.

***** Voor de jaren 2010 t/m 2014 zijn de maatwerkdata beschikbaar in de 'Remote Access'-omgeving van het CBS. Voor de jaren 2015 en 2016 zijn beperkte updates gedaan door het CBS. In overleg met het CBS kunnen mogelijkheden worden gezien om die desgewenst ook te benutten in de 'Remote Access'-omgeving.

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) - Innovatiestatistiek

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. Bedrijven met 200 of meer werkzame personen worden integraal waargenomen. 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 1 900 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 verstrekende kwaliteitscontroles onderworpen. Wel zijn de omzetcijfers zeer actueel.

De data zijn standaard beschikbaar via RA voor bedrijven met als hoofdactiviteit groothandel en handelsbemiddeling, vervoer en opslag, horeca, informatie en communicatie, specialistische zakelijke diensten, overige zakelijke dienstverlening en overige diensten. Voor overige sectoren (niet omvattende landbouw, bosbouw en visserij en financiële dienstverlening) is info over maatwerk te verkrijgen bij het CBS:

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 Octrooiencentrum Nederland (OCNL) in de periode 2000-2010 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 OCNL 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.

Het CBS heeft een recentelijk een nieuwe octrooistatistiek ontwikkeld, betrekking hebbend op de verslagjaren 2013 t/m 2019. Het betreft hier weer patentaanvragen vanuit Nederland bij het EPO en OCNL. Zie voor nadere info: <https://www.cbs.nl/nl-nl/onze-diensten/methoden/onderzoeksomschrijvingen/korte-onderzoeksbeschrijvingen/patentaanvragers-en-aanvragen-uit-nederland>; <https://opendata.cbs.nl/statline/#/CBS/nl/navigatieScherm/thema?themaNr=85288>). Deze gegevens komen naar verwachting in de RA-omgeving beschikbaar in april 2023.

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: TKI-grondslagdata en TKI-inzetdata

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.

De TKI-grondslagdata bevatten gegevens over private bijdragen aan R&D uitgevoerd door kennisinstellingen. Voor zover het om private cash bijdragen gaat, betreft het financiering van R&D die wordt uitgevoerd door kennisinstellingen. Daarnaast is in beperkte mate sprake van private bijdragen in natura, waarbij bedrijven met eigen R&D-inspanningen zelf aan projecten kunnen deelnemen.

De TKI-inzetdata bevatten gegevens over de inzet van PPS-toeslagmiddelen door TKI's in publiek-private samenwerkingsverbanden. Die gegevens geven integraal inzicht in projecten waarin bedrijven en kennisinstellingen gezamenlijk R&D uitvoeren in TKI-verband.

Deze microdatabestanden zijn registraties afkomstig van RVO, die aan het Algemeen Bedrijvenregister van het CBS zijn gekoppeld. Het 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 gehouden onder bedrijven vanaf 100 werkzame personen en in de meeste SBI-bedrijfstakken 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 en uitgaven aan R&D. Dit omvat ook financieringsbronnen voor de R&D-uitgaven en financiering door bedrijven van extern uitgevoerde R&D, waaronder R&D bij kennisinstellingen. Ook de omvang van het R&D-personeel en het aantal bedrijven dat bij R&D betrokken is, vallen onder deze dataverzameling. 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 (NFO)

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 NFO wordt samengesteld uit twee bronnen, enerzijds uit primaire waarneming (enquête) bij de grote ondernemingen (onderdeel SFGO van NFO) en anderzijds uit secundaire bron (de vennootschapsbelastinggegevens van de Belastingdienst) voor de kleine ondernemingen (onderdeel SFKO van NFO). Ondernemingen met een balanstotaal vanaf 40 miljoen euro (tot 2010 23 miljoen euro worden hierbij als grote ondernemingen aangemerkt en ondernemingen met een balanstotaal onder die grens als kleine ondernemingen).

Topsectorenindeling in het kader van de Monitor topsectoren van het CBS

Ten behoeve van de Monitor topsectoren en daaraan gerelateerd analyses van het gebruik van beleidsinstrumenten door topsectoren die het CBS in de periode 2012-2018 heeft uitgevoerd in opdracht van het ministerie van EZ(K), heeft het CBS een indeling van bedrijven ontwikkeld met een onderscheid naar topsectoren en niet-topsectoren en binnen de topsectoren een onderscheid naar afzonderlijke topsectoren.

Vier topsectoren zijn volledig op basis van SBI-codes afgebakend, Dat betreft de topsectoren Agri&food, Chemie, High tech systemen en materialen en Life sciences & health. Voor de overige vijf topsectoren is aanvullend maatwerk op bedrijfsniveau verricht, te weten voor de topsectoren

Creatieve industrie, Energie, Logistiek; Tuinbouw en uitgangsmaterialen en Water. Hierbij zijn individuele bedrijven geselecteerd met behulp van extra informatie over de topsector. Deze informatie is afkomstig uit secundaire bronnen, zoals ledenlijsten van brancheverenigingen. Dit heeft geresulteerd in een lijst maatwerkbedrijven voor de vijf 'maatwerktopsectoren'. Voor de jaren 2010 tot en met 2016 zijn aparte maatwerklijsten bijgehouden, met als peildatum 1 december van die jaren. Ze zijn, gekoppeld aan het ABR, ingezet voor berekeningen met indelingen naar topsectoren voor de daaropvolgende statistiekjaren 2011 tot en met 2017. -Zie voor meer informatie: <https://www.cbs.nl/-/media/pdf/2018/40/181008-monitor-topsectoren-2018.pdf>. De topsectorindeling van het CBS is te benutten in de evaluatie van de IPC-regeling, om in aansluiting op wat het CBS in het verleden al gedaan heeft voor de jaren 2011-2013, het aandeel van bedrijven binnen topsectoren te analyseren binnen gebruikers van de IPC-regeling (zie paragraaf 8 van de Terms of Reference voor de evaluatie van de IPC-regeling). Voor de jaren 2010 t/m 2014 zijn de maatwerkdata beschikbaar in de 'Remote Access'-omgeving van het CBS. Voor de jaren 2015 en 2016 zijn beperkte updates gedaan door het CBS. In overleg met het CBS kunnen mogelijkheden worden gezien om die desgewenst ook te benutten in de 'Remote Access'-omgeving.

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 WBSO/RDA-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	Kennisinstel- lingen	Toegankelijk- heid (RA = 'Remote Access'- omgeving van CBS)
ABR	2006- 2021	Alle in Nederland geregistreerde bedrijven en instellingen (incl. zelfstandigen)	Integraal (register)	Ja	Beschikbaar in RA
BDK	2007- 2021	Alle in Nederland geregistreerde bedrijven en instellingen (incl. zelfstandigen)	Integraal (register)	Ja	Beschikbaar in RA
Btw	2007- 2021	Statistische eenheden waarvan informatie uit de omzetbelasting beschikbaar is, exclusief eenheden met vrijstelling	Integraal (register)	Ja	Beschikbaar in RA
CIS	2010- 2020*	Bedrijven in SBI 01 t/m 82 en vanaf 10 werkzame personen (geen bijschatting voor kleinere bedrijven)	Steekproef	Nee	Beschikbaar in RA
DRT	2014- 2022	Standaard beschikbaar in RA: actieve bedrijfseenheden met als hoofdactiviteit groothandel en handelsbemiddeling, vervoer en opslag, horeca, informatie en communicatie, specialistische zakelijke diensten, overige zakelijke dienstverlening en overige diensten. Voor overige sectoren (niet omvattende landbouw, bosbouw en visserij en financiële dienstverlening) is info over maatwerk te verkrijgen bij het CBS.	Analysebesta nd op basis van verschillende bronnen van omzetgegeve ns	Geen hoger onderwijs of UMC's, wel overige kennisinstellinge n	Deels beschikbaar in RA
Investerings	2006- 2020	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

NFO	2000-2020	Rechtspersoonlijkheid bezittende ondernemingengroepen (b.v.'s, n.v.'s en coöperaties) die in Nederland actief zijn in de niet-financiële sector	Integrale waarneming; kleinere bedrijven op basis van vpb-aangiften (register)	Nee	Beschikbaar in RA
Patentaanvragen	2000-2010**	Alle patentaanvragen ingediend door Nederlandse aanvragers bij het EPO of OCNL	Integraal	Ja	Beschikbaar in RA
Polisadministratie	2006-2022	Banen van werknemers in Nederland	Integraal (register)	Ja	Beschikbaar in RA
PPS	2013-2021***	Bedrijven en onderzoeksorganisaties in Nederland die actief zijn in TKI-verband voor wat betreft grondslagvorming en inzet van PPS-toeslag in PPS-projecten	Integraal (register)	Ja	Beschikbaar in RA met toestemming EZK en RVO
PS	1993-2020	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 welzijnzorg	Steekproef	Nee	Beschikbaar in RA
RTD	2013-2020****	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
Topsectoren-indeling	2010-2014*****	Bedrijven en onderzoekorganisaties in Nederland die deel uitmaken van de topsectoren	Integraal (register)	Geen hoger onderwijs of UMC's, wel overige kennisinstellingen	Beschikbaar in RA

WBSO en RDA	2004-2021	Alle gebruikers met vastgestelde WBSO of RDA	Integraal (register).	Nee	Beschikbaar in RA met toestemming EZK en RVO
-------------	-----------	--	-----------------------	-----	--

* 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. In 2020 heeft een revisie van de innovatiestatistiek plaatsgevonden, betrekking hebbend op het verslagjaar 2018, naar aanleiding van aanpassingen van richtlijnen in de Oslo Manual van de OECD voor de innovatiestatistiek. Dit is gepaard gegaan met een herontwerp van de innovatie-enquête (CIS). Inhoudelijk is een belangrijke verandering geweest dat geen onderscheid meer wordt gemaakt tussen technologische en niet-technologische innovatie. In de nieuwe methodiek wordt systematisch van een breed innovatiebegrip uitgegaan, waarin technologische en niet-technologische innovatie zijn geïntegreerd. Voor nadere info is te raadplegen: <https://www.cbs.nl/nl-nl/publicatie/2020/42/ict-kennis-en-economie-2020>, paragraaf 6.1; <https://www.bedrijvenbeleidinbeeld.nl/bouwstenen-bedrijvenbeleid/innovatie/innovatie-activiteiten>.

** Het CBS heeft een nieuwe octrooistatistiek ontwikkeld, met data lopend van 2013 t/m 2019. De microdata komen naar verwachting in april 2023 beschikbaar voor externe onderzoekers in de RA-omgeving van het CBS.*** De PPS-toeslagdata bestaan uit TKI-grondslagdata en TKI-inzetdata. Bij de grondslagdata betreft het grondslagprojecten lopend t/m 2020, waarmee grondslag is gegenereerd in de jaren t/m 2021.

**** In 2019 heeft een herziening van de RTD plaatsgevonden vanwege aanpassingen aan de richtlijnen in de Frascati Manual van de OECD voor de R&D-statistiek: <https://www.cbs.nl/nl-nl/achtergrond/2020/02/research-en-development-revisie-2019>. Dat heeft er onder meer toe geleid dat ingeleend R&D-personeel ook meegenomen wordt bij de R&D-activiteiten en -uitgaven van een bedrijf. Een andere aanpassing is dat personen die minder dan 0,1 vte op jaarbasis aan R&D-activiteiten besteden, niet meer worden meegeteld in de R&D-statistiek. Cijfers volgens de methodiek na revisie zijn beschikbaar vanaf het verslagjaar 2013.

***** Voor de jaren 2010 t/m 2014 zijn de maatwerkdata beschikbaar in de 'Remote Access'-omgeving van het CBS. Voor de jaren 2015 en 2016 zijn beperkte updates gedaan door het CBS. In overleg met het CBS kunnen mogelijkheden worden gezien om die desgewenst ook te benutten in de 'Remote Access'-omgeving.

Bijlage 2. Frequentietellingen

In deze bijlage zijn frequentietellingen opgenomen voor een aantal belangrijke variabelen in de beschreven bronnen in het steekjaar 2016. 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 een aantal relevante variabelen uit de bronbestanden. 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 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 ¹		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

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.

¹ De cijfers van de RTD hebben betrekking op 2015 en zijn gebaseerd op de R&D-statistiek voorafgaand aan de in 2019 doorgevoerde revisie.

Tabel B.2 Frequentietellingen databronnen, 2016 (vervolg)

Investerings		DRT		PS	WBSO	
		Omzet		Toegevoegde	S&O-(niet)-	
				waarde	loonkosten	
	totaal	totaal	respons	totaal	totaal	respons
<i>bedrijfseenheden</i>						
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	618 250	583 220	25 240	9 915	8 880
	190					
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

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.