



Paper

Vergelijking internationale tijdreeksen Monitor Brede Welvaart

April 2026

Vergelijking internationale tijdreeksen Monitor Brede Welvaart

Over deze publicatie

Dit rapport beschrijft een analyse van de kwaliteit en bruikbaarheid van de internationale data. Dit wordt gedaan aan de hand van vier opgestelde criteria: nauwkeurigheid en plausibiliteit, tijdigheid, consistentie en vergelijkbaarheid. Wanneer de internationale data niet voldoet aan het criterium voor consistentie is het mogelijk om de data te corrigeren voor breuken (o.a. door gewijzigde definities of methoden) en uitbijters (extreme waarden). In dit onderzoek wordt hiervoor een methode ontwikkeld. Deze gecorrigeerde reeks wordt enkel gebruikt voor het bepalen van de trend, de originele tijdreeks wordt niet veranderd. Het toetsen van de data aan de kwaliteitscriteria en het bepalen van een betrouwbaardere trend draagt bij aan de kwaliteit van een dergelijke vergelijking tussen landen en maakt de analyse geschikter voor eventuele toepassing in de monitor. Daarmee kunnen de resultaten meer zeggingskracht en relevantie krijgen voor beleid.

Publicatiedatum 13-4-2026 06:30

Vergelijking internationale tijdreeksen Monitor Brede Welvaart

1. Inleiding

Het Centraal Bureau voor de Statistiek (CBS) heeft in 2017 de opdracht gekregen van het ministerie van Economische Zaken om een Monitor Brede Welvaart (MBW & SDG's) te ontwikkelen. Het CBS heeft sindsdien jaarlijks in mei een update van de verder ontwikkelde monitor uitgebracht ten behoeve van het verantwoordingsdebat in de Tweede Kamer. De monitor geeft een beeld van de huidige situatie van de brede welvaart in Nederland.

In de monitor worden ook de ontwikkelingen in de data geanalyseerd en geduid. De voornaamste analyses zijn het bepalen van de trendmatige ontwikkeling, de meest recente jaarmutatie en de positie van Nederland binnen de Europese Unie (EU). Het CBS heeft een methode ontwikkeld om de trendmatige veranderingen en de positie te vergelijken met de trend en positie van andere Europese landen. Op die manier kan worden onderzocht of Nederland ten opzichte van Europese landen voorloopt en verder uitloopt (een hoge positie en een relatief sterke trend), of achterloopt en verder achterop raakt (een lage positie en een relatief zwakke trend), of haar voorsprong verliest of haar achterstand inloopt. Voordat deze methode onderdeel van de monitor kan worden, zijn er nog resterende methodologische vraagstukken die opgelost moeten worden.

Tot nu toe is er in de MBW & SDG's alleen gebruik gemaakt van internationale data om voor één specifiek jaar een vergelijking te maken met Nederland. Om hiernaast ook een vergelijking te maken van de trend, in combinatie met de positie, moet een hele tijdreeks met internationale data worden gebruikt. De kwaliteit van de uitkomsten van de analyse hangt voor een groot deel af van de kwaliteit van de internationale data. Omdat het CBS deze data niet zelf maakt is onvoldoende informatie over de kwaliteit hiervan beschikbaar. Pas wanneer de data kwalitatief voldoende is bevonden, kan de Nederlandse trend en positie van een indicator zinvol vergeleken worden met de trend en positie van andere landen.

Dit rapport beschrijft een analyse van de kwaliteit en bruikbaarheid van de internationale data. Dit wordt gedaan aan de hand van vier opgestelde criteria: nauwkeurigheid en plausibiliteit, tijdigheid, consistentie en vergelijkbaarheid. Wanneer de internationale data niet voldoet aan het criterium voor consistentie is het mogelijk om de data te corrigeren voor breuken (o.a. door gewijzigde definities of methoden) en uitbijters (extreme waarden). In dit onderzoek wordt hiervoor een methode ontwikkeld. Deze gecorrigeerde reeks wordt enkel gebruikt voor het bepalen van de trend, de originele tijdreeks wordt niet veranderd. Het toetsen van de data aan de kwaliteitscriteria en het bepalen van een betrouwbaardere trend draagt bij aan de kwaliteit van een dergelijke vergelijking tussen landen en maakt de analyse geschikter voor eventuele toepassing in de monitor. Daarmee kunnen de resultaten meer zeggingskracht en relevantie krijgen voor beleid.

Publicatiedatum 13-4-2026 06:30

Vergelijking internationale tijdreeksen Monitor Brede Welvaart

2. Onderzoeksmethode

2.1 Inleiding

In de MBW worden ontwikkelingen bekeken aan de hand van drie analyses: de positie van Nederland binnen de EU, de trendmatige ontwikkeling en de recente jaarommutatie. Voor het vergelijken van de positie van Nederland met andere EU-landen is het belangrijk dat het niveau (de waarde) van de indicator onderling kan worden vergeleken. De trendmatige ontwikkeling wordt bepaald op de middellange termijn. Dit betekent dat er een trendperiode van acht jaar wordt gebruikt. Op basis van de data van deze jaren wordt een lineaire trend geschat ($y=ax+b$) waarbij een significante waarde van a bepaalt of de trend stijgt of daalt.

Door trend en positie voor alle Europese landen te vergelijken met de waarden voor Nederland kan worden onderzocht of Nederland bijvoorbeeld voorloopt en uitloopt of juist achterloopt en verder achterop raakt. Om een dergelijke vergelijking zinvol te maken moeten de gebruikte data (paragraaf 2.2) aan bepaalde kwaliteitscriteria voldoen. De opgestelde criteria worden in de paragraaf 2.3 toegelicht. Aan de hand van deze criteria kan de bruikbaarheid van een indicator in een trend-positie vergelijking worden ingeschat (hoofdstuk 3). Voor een trend-positie vergelijking zijn minimaal tien EU-landen nodig met geschikte tijdreeksen.

Aangezien in de MBW & SDG's Eurostat de grootste bron is voor internationale vergelijking wordt in dit onderzoek uitsluitend gebruik gemaakt van data afkomstig van Eurostat.¹⁾ Het onderzoek betreft een eerste verkenning van de bruikbaarheid van deze data waardoor slecht een beperkte selectie van deze indicatoren kan worden onderzocht. De geselecteerde indicatoren representeren een verscheidenheid aan onderwerpen.

In de originele datasets kunnen breuken en uitbijters zitten waardoor er geen betrouwbare trend kan worden berekend. Wanneer dat het geval is wordt het berekenen van de trend gedaan op basis van data gecorrigeerd voor dergelijke breuken en uitbijters. De originele data, die ook gebruikt worden voor het bepalen van de positie, worden hierbij niet vervangen. In hoofdstuk 4 wordt deze methode toegelicht.

2.2 Indicatoren en beschikbare informatie

Om een representatief beeld te kunnen geven is bij het selecteren van de indicatoren gezorgd voor een goede spreiding van onderwerpen, datatypes (index, percentages en absolute waarden) en vormen van dataverzameling. Zo komen de indicatoren terug in de Monitor Brede Welvaart binnen de verschillende thema's, dimensies en de Sustainable Development Goals.

Dit heeft geleid tot een selectie van de 13 indicatoren in Tabel 2.2.1. Van deze indicatoren worden de onderliggende tijdreeksen geanalyseerd.

2.2.1 Overzicht geanalyseerde indicatoren

Nr.	Indicator	Dimensie	Thema	SDG	Raamwerk
1	Relatieve armoede	-	-	1	EU-SILC
2	Toegevoegde waarde milieusector	-	-	9	EGSS
3	Individuele consumptie	Hier en nu	Materiële welvaart	8	ESA 2010
4	Vogels boerenland	-	-	15	OECD
5	Gewerkte uren in de zorg	-	-	3	ESA 2010
6	Neonatale sterfte	-	-	3	
7	Gezonde levensverwachting mannen	Hier en nu / later	Gezondheid	3	EU-SILC
8	Langdurige werkloosheid	Hier en nu	Arbeid en vrije tijd	8	EU-LFS
9	Totale woonquote (huur + koop)	Hier en nu	Wonen	11	EU-SILC
10	Milieuproblemen	Hier en nu	Milieu	12	EU-SILC
11	Gemiddelde schuld per huishouden	Later	Economisch kapitaal	10	ESA2010
12	Stikstofoverschot	Later	Natuurlijk kapitaal	15	GNB
13	Invoer fossiele energiedragers	Elders	Milieu en grondstoffen	8	EW-MFA

Van deze indicatoren zijn op de [website van Eurostat](#) de data voor alle landen te downloaden. Hiernaast kunnen ook bijbehorende 'flags' of markeringen, weergegeven per observatie, worden gedownload. Deze markeringen geven extra informatie over een datapunt. Eurostat definieert onder andere flags voor breuken (breaks: b) in de data, of er een afwijkende definitie (definition differs: df) wordt gehanteerd, of het cijfers met een lage betrouwbaarheid betreft (low reliability: u), of het een schatting (estimation: e) betreft en of het cijfer nog voorlopig (provisional: p) is. Aan een observatie kunnen meerdere van deze flags worden meegegeven.

Verder zijn op de Eurostat website uitgebreide metadata beschikbaar.²⁾ Deze metadata zijn voor elke indicator op dezelfde manier gestructureerd en bevat veel relevante informatie over de opgestelde kwaliteitscriteria (zie paragraaf 2.3). Wel verschilt de mate van detail van deze metadata per indicator. Zo zijn er indicatoren waarvoor de metadata enkel ingaan op het achterliggende raamwerk voor het opstellen van de data (zie kolom raamwerk). Voor andere indicatoren zijn meer onderwerp- of indicatorspecifieke metadata aanwezig.

2.3 Kwaliteitscriteria

In dit onderzoek zijn verschillende kwaliteitscriteria opgesteld die belangrijk zijn om aan te voldoen om tot een informatieve vergelijking te komen tussen internationale reeksen. Door reeksen aan deze criteria te toetsen kan met meer zekerheid worden gezegd of de onderlinge relatie tussen landen werkelijke ontwikkelingen weerspiegelen en niet voortkomen uit verschillende methoden of inconsistenties in de data. De tijdreeksen kunnen worden beoordeeld op in hoeverre ze voldoen aan deze criteria. Aan de hand hiervan kan worden besloten of een indicator in een trend-positievergelijking meegenomen kan worden.

In dit onderzoek zijn vier kwaliteitscriteria gedefinieerd: tijdigheid, nauwkeurigheid en plausibiliteit, consistentie en vergelijkbaarheid (tussen landen). Deze criteria zijn geïnspireerd op de Europese Code of Practice voor statistiek, waar ook een kwaliteitsraamwerk bij hoort. Deze richtlijnen zijn leidend voor het CBS als het gaat om datakwaliteit en zetten de standaard voor de statistische bureaus in de Europese Unie. De praktijkcode bestaat uit 16 beginselen met betrekking tot het institutionele kader, statistische procedures en statistische output. Voorbeelden van deze beginselen zijn: sterk kwaliteitsbewustzijn, deugdelijke methoden, nauwkeurigheid en betrouwbaarheid, tijdigheid en punctualiteit en samenhang, vergelijkbaarheid en consistentie. In het definiëren van de vier kwaliteitscriteria ligt de focus op (zo veel mogelijk meetbare) eisen die kunnen worden getoetst aan de hand van de tijdreeksen en de metadata. De tijdreeksen worden in het volgende hoofdstuk getoetst aan deze criteria die hieronder in meer detail worden uitgediept.

Nauwkeurigheid en plausibiliteit

Nauwkeurigheid en plausibiliteit gaan over de mate van waarschijnlijkheid dat de gerapporteerde statistiek de werkelijkheid benaderd. Het beoordelen van de nauwkeurigheid en plausibiliteit van de tijdreeks begint bij het vaststellen van de bron van de tijdreeks. Indicatoren kunnen worden opgesteld vanuit bestaande brondata waarvoor specifieke kwaliteitsraamwerken gelden waarvan de implementatie wordt getoetst (in dit geval) door Eurostat. Informatie hierover is te vinden in de uitgebreide metadata op de website van Eurostat. Ook wordt via contactpersonen van het CBS bij Eurostat en personen binnen het CBS die betrokken zijn bij het opstellen van deze brondata hierover meer informatie opgehaald. Er zijn echter ook statistieken die buiten een dergelijk raamwerk worden opgesteld, wat een inschatting van de betrouwbaarheid minder eenvoudig maakt.

Voor het bepalen van een trend is het verder belangrijk om na te gaan of deze datapunten geen uitbijters bevatten. Tijdreeksen kunnen extreme waarden bevatten, die uitbijters worden genoemd. Het aantal uitbijters kan bijdragen aan een inschatting van de nauwkeurigheid en plausibiliteit van een tijdreeks.

Tijdigheid

Het vergelijken van tijdreeksen is alleen informatief wanneer van deze tijdreeks recente data beschikbaar zijn. We noemen een reeks tijdig genoeg wanneer in een trendperiode (minimaal acht jaar) het meest recente datapunt niet ouder is dan vijf jaar voor het einde van deze periode. De tijdigheid is dan t-5. Wanneer er bijvoorbeeld een trendperiode van 2015-2024 wordt beschouwd, mag het meest recente datapunt niet ouder zijn dan 2019. Tijdreeksen die niet aan deze basisvoorwaarde voldoen zijn niet geschikt voor analyse. Voor de tijdreeksen die hier wel aan voldoen kan de tijdigheid verschillen, waarbij t-1 het hoogst haalbare is.

Daarnaast wordt geïnventariseerd wat het eerste en het laatste jaar is met een observatie voor Nederland, het aantal landen met data vanaf 2005 en hoeveel resultaten nog voorlopig zijn. Deze gegevens vormen de basis voor de mogelijkheid tot het opstellen van een tijdreeks met voldoende data om een uitspraak te doen over de tijdigheid van de indicator. Verder is voor het aantal landen vanaf 2005 gekozen om een beeld te krijgen van de doorlooptijd van de statistiek, en daarmee een beeld kan geven voor de andere kwaliteitscriteria vanwege eventuele langlopende kwaliteitsraamwerken.

Consistentie

Om van een tijdreeks een trend te kunnen bepalen moeten er voldoende datapunten beschikbaar zijn. Omdat $y = ax + b$ twee onbekenden bevat, zijn minimaal drie datapunten nodig om een trend te kunnen vaststellen. Bovendien wordt de trend bepaald voor de middellange termijn: de afgelopen acht jaar. Hiermee is de basis voor een consistente tijdreeks minimaal drie datapunten in de afgelopen acht jaar. Tijdreeksen die hier niet aan voldoen zijn niet geschikt voor analyse. Voor de tijdreeksen die hier wel aan voldoen kunnen klassen worden toebedeeld op basis van het aantal beschikbare datapunten in de trendperiode.

Deze datapunten moeten voor het bepalen van een trend verder onderdeel zijn van een consistente reeks zonder wijzigingen in methoden (breuken). Tijdreeksen bevatten vaak breuken. We onderscheiden methodebreuken en trendbreuken. COVID zorgde voor veel trendbreuken in bijvoorbeeld data van ziekenhuizen. Een wijziging in een definitie van bijvoorbeeld gezonde levensverwachting levert een methodebreuk op. De aanwezigheid van breuken hoeft niet te betekenen dat de reeks ongeschikt is voor een trend-positie vergelijking. Breuken kunnen namelijk worden gerepareerd waardoor consistente reeksen ontstaan voor het bepalen van een trend. De methode om breuken te repareren wordt verder beschreven in hoofdstuk 4.

Een bepaalde mate van garantie dat de tijdreeks wordt gecontinueerd is ook onderdeel van het criterium consistentie. De beoordeling hiervan vindt echter al plaats wanneer een indicator wordt opgenomen in de MBW en was daarom geen onderdeel van dit onderzoek.

Vergelijkbaarheid

Tot slot is een vergelijking van de trend en positie van landen onderling alleen informatief wanneer de data vergelijkbaar zijn met die van andere landen. Over de (geografische) vergelijkbaarheid wordt informatie gegeven in de metadata van Eurostat. Wanneer landen dezelfde methoden en definities hanteren is de tijdreeks vergelijkbaar. Methoden en definities worden vastgelegd in onderliggende raamwerken vanuit Eurostat. De mate van detail in deze raamwerken geeft een indicatie voor de vergelijkbaarheid van de opgeleverde data.

Hiernaast is de flag 'df', voor een afwijkende gehanteerde methode ten opzichte van de methode die door Eurostat in het raamwerk wordt voorgesteld, ook bepalend voor de vergelijkbaarheid. Wanneer de definitie afwijkt kan de positie tussen landen niet direct worden vergeleken. Een verschil in definitie kan ook voor een vergelijking van de trendontwikkeling een probleem vormen. Wanneer een land een andere definitie hanteert voor een indicator zal ook de ontwikkeling van de indicator anders zijn dan voor andere landen. Een trend kan dan niet meer goed vergeleken worden.

De kwaliteitscontrole door Eurostat wordt in dit onderzoek als leidend gezien vanwege de geldende kwaliteitseisen die Eurostat hanteert. Er is niet gekeken naar landelijke kwaliteitsrapportages. Op basis van enkel de tijdreeksen en flags van Eurostat kan worden besloten een land uit de trend- en positievergelijking te halen.

¹⁾ De indicator 'Vogels boerenland' worden wel op de Eurostat website gepubliceerd maar hierbij wordt aangegeven dat Eurostat niet genoemd moet worden als bron. Eurostat ontvangt de data niet direct van de deelnemende landen en voert op de data zelf geen kwaliteitscontrole uit.

²⁾ De metadata is geraadpleegd in januari 2026.

Publicatiedatum 13-4-2026 06:30

Vergelijking internationale tijdreeksen Monitor Brede Welvaart

3. Beoordeling van kwaliteitscriteria

3.1 Inleiding

De dertien geselecteerde indicatoren worden in dit hoofdstuk beoordeeld op de vier geïntroduceerde kwaliteitscriteria. Per criterium zijn een aantal (meetbare) sub-criteria onderscheiden op basis waarvan de indicator wordt getoetst.

Om een indicator op te nemen in de trend-positievergelijking moeten minimaal 10 landen (inclusief Nederland) kunnen worden vergeleken en daarmee aan de kwaliteitseisen voldoen. De criteria worden beoordeeld aan de hand van de data met bijbehorende flags en beschikbare metadata. Deze informatie is geraadpleegd in januari 2026. Verder zijn de bevindingen getoetst bij, en eventueel aangevuld door, experts op het gebied van deze (bron)data binnen het CBS (met uitzondering van 'vogels boerenland').

Per criterium wordt voor de indicatoren een klasse toebedeeld: Laag (L), Midden (M) of Hoog (H). Om meer nuance te kunnen aanbrengen is een klasse hiertussen, bijvoorbeeld midden tot hoog (M-H), ook mogelijk. De vier kwaliteitscriteria worden eerst individueel beschreven en beoordeeld, waarna een algemene klasse wordt gegeven voor de indicator. De klassen moeten worden gezien als een indicatie van de bruikbaarheid van de tijdreeksen voor een trend-positievergelijking en dus niet van de algemene kwaliteit van de statistiek.

3.2 Nauwkeurigheid en plausibiliteit

In Tabel 3.2.1 zijn de resultaten van de analyse voor de criteria nauwkeurigheid en plausibiliteit opgenomen. De kwaliteit van de brondata is voor een groot deel bepalend voor de kwaliteit van de resulterende indicator. Afhankelijk van de brondata geldt er bijvoorbeeld een handleiding met bijbehorende kwaliteitseisen voor het opstellen van de statistieken. Eurostat en de lidstaten hechten veel belang aan de kwaliteit van de statistieken en door Eurostat worden hier voor sommige indicatoren uitvoerige kwaliteitschecks op gedaan. In de metadata wordt vaak uitspraak gedaan door Eurostat over de kwaliteit van de statistiek, zie de kolom 'kwaliteit en plausibiliteit uit metadata'.

Voor bijna alle (bron)statistieken bestaan er raamwerken vanuit Eurostat (met uitzondering van 'vogels boerenland'). Het detailniveau van deze raamwerken en de kwaliteitsgarantie hangt onder andere af van de leeftijd van de statistiek (zie ook de tabel voor tijdigheid in de volgende paragraaf). Zo zijn de raamwerken van bijvoorbeeld de Nationale Rekeningen (ESA) en de EU-SILC al lang in ontwikkeling en wordt kwaliteit op deze manier bovengemiddeld geborgd (M-H of H). Andere statistieken als de 'toegevoegde waarde milieusector' zijn relatief nieuw (verplicht sinds 2017) en daarom nog volop in ontwikkeling, wat invloed heeft op de nauwkeurigheid en plausibiliteit van de data (M).

De manier van dataverzameling heeft hiernaast ook invloed op de kwaliteit. Data kunnen worden opgesteld vanuit registerdata of door middel van (telefonische) enquêtes. Hoe data verzameld worden kan overigens ook voor één indicator per land verschillen. Landen kunnen hiernaast ook een combinatie van deze methoden voor het opstellen van een indicator gebruiken. Het is lastig om een waardeoordeel aan de manier van dataverzameling toe te kennen. Sommige informatie kan alleen worden verkregen vanuit enquêtes zoals 'milieuproblemen' (M-H), of op basis van een zo goed mogelijke inschatting vanuit beschikbare informatie zoals 'stikstofoverschot' (M). De regulatie van de steekproef van een enquête, zoals de omvang of spreiding, heeft invloed op de nauwkeurigheid van de resultaten. Hoe groter de steekproef met bijbehorende eisen, hoe nauwkeuriger de resulterende statistiek zal zijn (M tot M-H). Een aantal statistieken wordt opgesteld vanuit registerdata, wat de nauwkeurigheid over het algemeen bevordert (M-H tot H). Tot slot zijn er ook statistieken waarvoor, afhankelijk van de beschikbare bronnen, een combinatie van registerdata en enquêtedata wordt gebruikt of waarvoor de gebruikte bronnen per land verschillen. Dit heeft, naast mogelijke invloed op de kwaliteit, ook gevolgen voor de vergelijkbaarheid (M tot M-H).

Van de geselecteerde indicatoren is er geen indicator waar het op basis van plausibiliteit en nauwkeurigheid direct wordt afgeraden om een trend-positievergelijking te maken. De insteek van Eurostat is ook om data te publiceren met voldoende kwaliteit om een informatieve vergelijking tussen landen te kunnen maken. Voor drie indicatoren is de klasse M toebedeeld, vier krijgen de klasse M-H mee en voor zes wordt de kwaliteit en plausibiliteit van de indicator als hoog beoordeeld.

3.2.1 Classificering op basis van nauwkeurigheid en plausibiliteit

Nr.	Indicator	Klasse	Nauwkeurigheid en plausibiliteit uit metadata
1	Relatieve armoede	M-H	Op basis van EU-SILC raamwerk, combinatie van (telefonische) enquêtes en registerdata, verschillende methoden per land waardoor de kwaliteit ook kan verschillen. Voor de enquêtes worden modelvragen opgesteld die letterlijk vertaald moeten worden. Er gelden gedetailleerde regels voor het bepalen van de steekproef. Eurostat voert uitvoerige kwaliteitschecks uit.
2	Toegevoegde waarde milieusector	M	Relatief nieuwe statistiek die pas verplicht is sinds 2017. De methoden zijn nog in ontwikkeling. Eurostat voert wel kwaliteitscontroles uit.
3	Individuele consumptie	H	Wordt bepaald a.d.h.v. het ESR en Systeem van Nationale Rekeningen, langlopend raamwerk met uitvoerige kwaliteitschecks door Eurostat.
4	Vogels boerenland	M-H	Kwaliteit is volgens de metadata goed, aangezien Europese totalen convergeren. Eurostat ontvangt de data niet van de lidstaten en voert geen kwaliteitschecks uit.
5	Gewerkte uren in de zorg	H	Wordt bepaald a.d.h.v. het ESR en Systeem van Nationale Rekeningen, langlopend raamwerk met uitvoerige kwaliteitschecks door Eurostat.
6	Neonatale sterfte	M	Vanuit basisregistratie personen en medische geboorte registers: levend geboren kinderen en sterfgevallen onder 28 dagen. Onderdeel van de Demographic Balance. Verschillen in geboorteregistraties, b.v. omtrent de duur van de zwangerschap, wel belangrijk voor nauwkeurigheid. In metadata geen informatie over nauwkeurigheid.
7	Gezonde levensverwachting mannen	H	Combinatie van EU-SILC (zie '1. Relatieve armoede') voor bepalen van 'gezonde' jaren en data over sterfgevallen. De algehele accuraatheid wordt door Eurostat als hoog bestempeld.
8	Langdurige werkloosheid	H	Accuraatheid van EU-LFS wordt als hoog bestempeld door Eurostat, voldoet aan door Eurostat opgestelde criteria voor betrouwbaarheid. Verzameld d.m.v. interviews aan een grote steekproef, onderhevig aan gebruikelijke fouten gerelateerd aan deze manier van dataverzameling.
9	Totale woonquote (huur + koop)	M-H	Op basis van EU-SILC (zie 1. 'Relatieve armoede').
10	Milieuproblemen	M-H	Op basis van EU-SILC (zie 1. 'Relatieve armoede').
11	Gemiddelde schuld per huishouden	H	Wordt bepaald a.d.h.v. het ESR en Systeem van Nationale Rekeningen, dit zijn langlopende raamwerken met uitvoerige kwaliteitschecks door Eurostat. Brondata afkomstig van banken.
12	Stikstofoverschot	M	Nauwkeurigheid en betrouwbaarheid hangt af van de onderliggende data voor het bepalen van het overschot. Opgebouwd uit: mestproductie (ingeschat op basis van hoeveelheid vee), import en export van mest, gebruik van (in)organische meststof en stikstof onttrekking. Niet al deze data is voor elk land beschikbaar en van dezelfde kwaliteit, daardoor beperktere nauwkeurigheid.
13	Invoer fossiele energiedragers	H	De nauwkeurigheid van EW-MFA wordt als goed beschouwd. Er zijn methodologische handleidingen, validatie procedures en kwaliteitsrapportages. De statistiek wordt al lang opgesteld (op Eurostat vanaf 1995). Onderliggende data (Internationale Handel) gaat achteruit door verlaagde lastendruk. Dit geldt echter nog niet voor fossiele energiedragers.

3.3 Tijdigheid

In onderstaande tabel is het resultaat te zien voor het criterium tijdigheid. Voor alle indicatoren voldoet een groot genoeg aantal landen (minimaal tien, inclusief Nederland) aan de basiseis voor tijdigheid namelijk t-5. De data voor de indicator 'vogels boerenland' worden door steeds minder landen data aangeleverd. Waar tussen 2010 en 2016 nog 22 landen data aanleverden zijn dit er in 2021 maar 16 en in 2022 slechts 12. Wanneer indicatoren in een trend-positievergelijking worden opgenomen moet periodiek gecontroleerd worden of nog steeds voldoende landen aan het criteria voor tijdigheid voldoen. De indicator 'stikstofoverschot' wordt vanaf verslagjaar 2026 verplicht, wat het aantal landen dat kan voldoen aan het criterium tijdigheid zal verhogen.

Het hoge aantal voorlopige cijfers bij 'individuele consumptie' en 'gewerkte uren in de zorg' komt waarschijnlijk voort uit de recente vijfjaarlijkse revisie van de Nationale Rekeningen.³⁾ Over het algemeen geldt voor cijfers in de Nationale Rekening wel een hoge tijdigheid.

Voor een aantal van de indicatoren komt het laatste datapunt uit 2023. Aan deze indicatoren is de klassen M-H toebedeeld. Opvallend is dat dit ook data vanuit de EU-SILC betreft terwijl deze altijd snel beschikbaar zijn op basis van de Verordening, daarom hebben deze indicatoren alsnog de klasse H toebedeeld gekregen.

Het criterium tijdigheid lijkt op basis van deze analyse in de meeste gevallen geen obstakel te zijn voor de toepassing van een indicator in een trend-positievergelijking. Wel loopt voor een aantal indicatoren de tijdigheid achteruit waardoor het aan te raden is om het criterium tijdigheid van t-5 periodiek te controleren.

3.3.1 Classificering op basis van tijdigheid

Nr.	Indicator	Klasse	Jaar eerste datapunt NL	Laatste datapunt	Landen met data vanaf 2005	Aantal landen met tijdigheid t-5	Voorlopig (landen vanaf 2015)
1	Relatieve armoede	H	1995	2024	25	27	1
2	Toegevoegde waarde milieusector	M-H	2001	2023	1	27	3
3	Individuele consumptie	H	1995	2024	27	27	10
4	Vogels boerenland	M-H	1995	2023	18	18	0
5	Gewerkte uren in de zorg	H	1995	2024	28	28	11
6	Neonatale sterfte	M-H	1995	2023	26	25	0
7	Gezonde levens-verwachting mannen	H	2005	2023	28	27	4
8	Langdurige werkloosheid	H	2009	2024	2	27	0
9	Totale woonquote (huur + koop)	H	2005	2024	26	27	1
10	Milieuproblemen	H	2005	2023	26	28	0
11	Gemiddelde schuld per huishouden	H	1995	2024	26	25	2
12	Stikstofoverschot	M-H	1995	2023	28	19	0
13	Invoer fossiele energiedragers	H	1996	2024	27	27	0

3.4 Consistentie over tijd

Klassen voor consistentie worden gebaseerd op een combinatie van het aantal landen dat voldoet aan het criterium van aantal landen met minimaal drie datapunten in de afgelopen acht jaar en aanvullende achtergrond informatie uit de metadata. Bij elke reeks wordt door voldoende landen (minimaal tien) voldaan aan het basiscriterium van minimaal drie observaties in de afgelopen acht jaar (minimaal 19, maximaal 28). De indicator 'toegevoegde waarde milieusector' is een relatief nieuwe statistiek waardoor methoden die landen hanteren momenteel nog in ontwikkeling zijn. Hierdoor veranderen definities en methoden nog in de loop der tijd of worden ze aangescherpt. Door deze wijzigingen is een vergelijking (en het berekenen van een trend) voor een land momenteel nog beperkt informatief (L-M).

Voor sommige reeksen worden breuken gerepareerd door Eurostat, bijvoorbeeld voor 'langdurige werkloosheid'. In dit geval zijn er dan geen breuken aangegeven in de data maar wordt in de metadata benoemd of breuken al gecorrigeerd zijn. Er zijn ook reeksen waarvoor, per land per jaar, wel wordt aangegeven dat er breuken in de reeks aanwezig zijn. Deze breuken kunnen met behulp van het schatten van een dummy-variabele op basis van een sterk correlerende reeks van een ander land, worden gecorrigeerd. Hierdoor kan er als nog een brouwbare trend worden bepaald waarbij de originele reeksen niet worden gewijzigd. Met behulp van deze methode hoeft de aanwezigheid van breuken geen grote gevolgen te hebben voor de bruikbaarheid van de data. Deze methode wordt uitgewerkt in hoofdstuk 4 en als voorbeeld toegepast op de reeksen 'relatieve armoede' en 'gezonde levensverwachting mannen'. Aan deze indicatoren kan daarom voor het criterium consistentie alsnog een Midden tot Hoge klasse worden toebedeeld. Wel zijn er een aantal indicatoren (zie tijdigheid) waarvoor steeds minder landen data lijken aan te leveren, wat consistentie in de toekomst zou kunnen belemmeren. Dit leidt nu niet tot de toedeling van een lagere klasse, maar geeft wederom aan dat de toebedeling van klassen periodiek zou moeten worden herzien.

Slechts één indicator zou op basis van het criterium consistentie afvallen voor een trend-positievergelijking, namelijk 'toegevoegde waarde milieusector'. Het overgrote deel is goed te vergelijken, acht indicatoren worden geclassificeerd als M-H en voor vier indicatoren wordt de vergelijkbaarheid over de jaren als hoog gezien.

3.4.1 Classificering op basis van consistentie

Nr.	Indicator	Klasse	Consistentie uit metadata	Aantal landen > 3 obs. over t-8	Breuken
1	Relatieve armoede	M-H	Framework borgt de vergelijkbaarheid over tijd. Revisie in 2021, aangegeven als breuk.	27	11
2	Toegevoegde waarde milieusector	L-M	Aangezien definities en methoden nog in ontwikkeling zijn is een vergelijking over verschillende jaren momenteel nog beperkt informatief is.	27	5
3	Individuele consumptie	H	Langlopend raamwerk borgt vergelijkbaarheid, elke 5 jaar wordt de tijdreeks vanaf 1995 gereviseerd.	27	0
4	Vogels boerenland	M-H	Gehele tijdreeks wordt door landen opnieuw berekend bij nieuwe levering.	19	0
5	Gewerkte uren in de zorg	H	Langlopend raamwerk borgt vergelijkbaarheid, elke 5 jaar wordt de tijdreeks vanaf 1995 gereviseerd.	28	1
6	Neonatale sterfte	M-H	Niet verplichte statistiek: niet alle landen leveren consistent genoeg data. Er zijn op dit moment nog voldoende landen die aan het consistentie criterium voldoen. Verder kunnen geografische veranderingen van landen de populatie beïnvloeden, maar betreft een index dus het effect zal klein zijn.	24	0
7	Gezonde levensverwachting mannen	M-H	Op basis van EU-SILC (zie 1. 'relatieve armoede') en aantal sterfgevallen.	28	25
8	Langdurige werkloosheid	M-H	Herontwerp van EU-LFS in 2021, landen hoeven geen gecorrigeerde tijdreeks aan te leveren maar een breuk-schatting op basis hiervan heeft Eurostat een correctie van de tijdreeks gemaakt.	27	12
9	Totale woonquote (huur + koop)	M-H	Op basis van EU-SILC (zie 1. 'relatieve armoede').	27	11
10	Milieuproblemen	M-H	Op basis van EU-SILC (zie 1. 'relatieve armoede').	28	7
11	Gemiddelde schuld per huishouden	H	Langlopend raamwerk borgt vergelijkbaarheid over tijd, elke 5 jaar wordt de tijdreeks vanaf 1995 gereviseerd.	26	0
12	Stikstofoverschot	M-H	Over het algemeen is er een goede vergelijking mogelijk. Landen wordt gevraagd data terug te reken bij het implementeren z van een nieuwe methode.	19	0
13	Invoer fossiele energiedragers	H	Vergelijkbaarheid is goed door heldere statistische concepten en definities. In de data zijn trendbreuken voor de hele tijdreeks al hersteld.	27	0

3.5 Vergelijkbaarheid tussen landen

Voor twee indicatoren wordt in de metadata door Eurostat nadrukkelijk afgeraden om een directe vergelijking van de data te maken tussen landen, resulterend in klasse L. Zo wordt een vergelijking van de positie voor 'vogels boerenland' afgeraden. Deze indicator betreft een indexcijfer waardoor de 'positie' erg afhankelijk is van de stand in het indexjaar, die niet altijd even nauwkeurig is bepaald. Voor 'stikstofoverschot' wordt een vergelijking over het algemeen afgeraden door verschillende methoden en bronnen.

De indicator 'toegevoegde waarde milieusector' valt ook nog niet goed te vergelijken tussen landen, omdat de statistiek nog in ontwikkeling is en landen elk in een verschillend stadium van implementatie verkeren. Tot slot wordt de indicator 'totale woonquote' op vergelijkbaarheid als laag geclassificeerd, woonlasten worden verschillend gemeten en niet alle landen houden zich aan de door Eurostat voorgeschreven definitie.

Wederom zorgt een uitgebreid raamwerk met kwaliteitschecks vanuit Eurostat ervoor dat voor een groot aantal indicatoren de vergelijkbaarheid tussen landen als midden tot hoog kan worden omschreven. Vaak zijn definities goed afgebakend maar kan er nog verschil optreden in de door landen gehanteerde methoden en gebruikte bronnen. Dit hangt onder anderen af van de beschikbaarheid van bijvoorbeeld registerdata, waarvan in Nederland relatief veel beschikbaar is. Ondanks de verschillende manieren van dataverzameling wordt de vergelijkbaarheid van de methoden tussen landen wel door Eurostat gecontroleerd en deze wordt in de metadata beschreven.

Voor drie indicatoren wordt momenteel een trend-positievergelijking afgeraden door een te beperkte vergelijkbaarheid van de data tussen landen, één indicator krijgt de klasse M toebedeeld. Voor vier indicatoren is de vergelijkbaarheid goed met een klasse van M-L en voor vijf indicatoren wordt de vergelijkbaarheid tussen landen als hoog gezien.

3.5.1 Classificering op basis van vergelijkbaarheid

Nr.	Indicator	Klasse	Aantal landen met flag 'definition differs' vanaf 2005	Vergelijkbaarheid uit metadata
1	Relatieve armoede	M-H	0	De vergelijkbaarheid tussen landen wordt vanuit het EU-SILC raamwerk en checks vanuit Eurostat goed geborgd. Door de verschillende manieren van data verzameling kunnen er echter wel verschillen optreden tussen landen.
2	Toegevoegde waarde milieusector	L-M	0	Nieuwe statistiek. Momenteel verschillen methoden, gebruikte bronnen en compleetheid nog tussen landen.
3	Individuele consumptie	H	0	Langlopend raamwerk van Nationale Rekeningen borgt vergelijkbaarheid tussen landen.
4	Vogels boerenland	L	12	Trend is te vergelijken maar de positie moet niet worden vergeleken, index data is erg volatiel. Veel data met de flag definition differs.
5	Gewerkte uren in de zorg	H	0	Langlopend raamwerk van Nationale Rekeningen borgt vergelijkbaarheid tussen landen.
6	Neonatale sterfte	M	0	Versillen tussen landen in de registratie van neonatale sterfte, afhankelijk van het aantal weken zwangerschap en geboortegewicht. Door de zeldzaamheid kan de statistiek voor kleine landen sterk fluctueren.
7	Gezonde levensverwachting mannen	M-H	0	Op basis van EU-SILC (zie 1. 'relatieve armoede') en aantal sterfgevallen.
8	Langdurige werkloosheid	H	2	Vergelijkbaarheid voor de EU-LFS wordt als erg hoog omschreven.
9	Totale woonquote (huur + koop)	L	0	Op basis van EU-SILC, maar woonlasten worden wel verschillend gemeten door de lidstaten (geïmputeerd, a.d.h.v. register-data of doormiddel van enquêtes waarin de vraagstelling niet altijd gelijk is). De definitie zoals voorgeschreven door Eurostat wordt niet altijd gehanteerd.
10	Milieuproblemen	M-H	2	Op basis van EU-SILC (zie 1. 'relatieve armoede').
11	Gemiddelde schuld per huishouden	H	0	Raamwerk van Nationale Rekeningen borgt vergelijkbaarheid tussen landen. Geen invloed op klasse maar door Nederlands hypotheek-stelsel heeft Nederland een hoge gemiddelde schuld per huishouden t.o.v. andere landen.
12	Stikstofoverschot	L	0	Volgens Eurostat moet er geen directe vergelijking worden gemaakt tussen landen vanwege verschillende methoden en data bronnen.
13	Invoer fossiele energiedragers	H	0	Vergelijkbaarheid tussen landen is hoog door heldere concepten en definities. Aangezien landen verschillende bronnen kunnen gebruiken bij het opstellen van de EW-MFA kan de scope en kwaliteit verschillen. Voor fossiele energiedragers wordt de vergelijkbaarheid echter als hoog gezien.

3.6 Geschiktheid per indicator

In onderstaande tabel worden de gecombineerde klassen per indicator gegeven ('Totaal'). De laagst toegekende klasse per kwaliteitscriterium bepaalt hierbij de bruikbaarheid van een indicator. Hoe lager de toebedeelde klasse, des te voorzichtiger moet worden omgegaan met het trekken van conclusies uit een trend-positievergelijking.

Voor vier van de 13 indicatoren valt een trend-positie vergelijking momenteel in meer of mindere mate af te raden, namelijk: 'toegevoegde waarde milieusector', 'vogels boerenland', 'totale woonquote' en 'stikstofoverschot' (klasse L tot L-M). Er is één indicator waarvoor in zijn totaliteit de klasse M wordt toebedeeld. Voor vier indicatoren wordt een klasse M-H gegeven en voor vier indicatoren wordt de toepasbaarheid als hoog ingeschat (na correctie van breuken en uitbijters). De vergelijkbaarheid tussen landen lijkt het vaakst een belemmering te vormen. Op het criterium tijdigheid werd een vergelijking voor geen enkele indicator afgeraden.

De hoogste klassen worden toebedeeld aan economische indicatoren. Deze zijn in vergelijking met indicatoren gerelateerd aan milieu langer in ontwikkeling en vaak nauwkeuriger te bepalen. Verder kan de geschiktheid van een indicator wijzigen over tijd. Zo wordt 'stikstofoverschot' vanaf verslagjaar 2026 een verplichte statistiek waardoor de klassen voor alle kwaliteitscriteria naar verwachting zullen toenemen. Een periodieke toetsing, bijvoorbeeld om de vijf jaar, is om deze reden gewenst om de betrouwbaarheid te garanderen.

3.6.1 Geschiktheid per indicator

Nr.	Indicator	Totaal	Nauwkeur- igheid en plausibi- liteit	Tijdigheid	Consistentie	Vergelijk- baarheid
1	Relatieve armoede	M-H	M-H	H	M-H	M-H
2	Toegevoegde waarde milieusector	L-M	M	M-H	L-M	L-M
3	Individuele consumptie	H	H	H	H	H
4	Vogels boerenland	L	M-H	M-H	M-H	L
5	Gewerkte uren in de zorg	H	H	H	H	H
6	Neonatale sterfte	M	M	M-H	M-H	M
7	Gezonde levens-verwachting mannen	M-H	H	H	M-H	M-H
8	Langdurige werkloosheid	M-H	H	H	M-H	H
9	Totale woonquote (huur en koop)	L	M-H	H	M-H	L
10	Milieuproblemen	M-H	M-H	H	M-H	M-H
11	Gemiddelde schuld per huishouden	H	H	H	H	H
12	Stikstofoverschot	L	M	M-H	M-H	L
13	Invoer fossiele energiedragers	H	H	H	H	H

³⁾ *Periodiek worden de nationale rekeningen gereviseerd. Daarbij worden nieuwe bronnen, methoden en concepten doorgevoerd in de nationale rekeningen, zodat het beeld van de Nederlandse economie weer optimaal aansluit bij alle onderliggende statistieken, bronnen en internationale richtlijnen voor het samenstellen van de nationale rekeningen.*

4. Correcties voor breuken en uitbijters in de data

4.1 Inleiding

In de dataset van Eurostat wordt systematisch aangegeven of er sprake is van een methodebreuk in de data. In dat geval zijn onderlinge datapunten niet goed met elkaar vergelijkbaar. Een gevolg is dat er geen trend kan worden bepaald. In dit hoofdstuk wordt uitgewerkt hoe daar mee om te gaan. Bij het CBS is veel ervaring opgedaan in het repareren van methodebreuken (Van den Brakel, Kraan, Roels, 2010, Methodenreeks; Thema: methode breuken en Lodder, 2007, Het repareren van trendbreuken). In dit hoofdstuk wordt de methodiek toegelicht en toegepast op een aantal reeksen in dit onderzoek.

Allereerst wordt gekeken naar de hoeveelheid breuken in de reeksen die in dit onderzoek centraal staan. In Bijlage 1 wordt weergegeven hoeveel landen er zijn met breuken per statistiek, of er in de tijdreeks van Nederland breuken zijn, en hoeveel reeksen van landen er zijn met meer dan vijf breuken. Opvallend is dat er een paar statistieken zijn met heel veel breuken. Dat zijn:

- Relatieve armoede , nr. 1
- Gezonde levensverwachting, nr. 7

Het is mogelijk om een redelijke schatting van het methode-effect te maken per tijdreeks met breuken van elk land voor elke statistiek. Niet in elke reeks zit een breuk en bovendien zitten de breuken niet altijd in hetzelfde jaar. Dat biedt mogelijkheden om de breuken te repareren.

In grote lijnen komt het repareren van een specifieke reeks van een land (doelreeks) op het volgende neer:

- Er wordt geconstateerd dat er methodebreuken zijn in de doelreeks, aan de hand van de informatie van Eurostat. Er wordt bekeken in welk jaren er breuken zijn. Vervolgens wordt bekeken welke reeksen van andere landen sterk correleren met de doelreeks, die geen methodebreuken bevatten in dezelfde jaren waarin de doelreeks die heeft. Het selecteren van deze reeksen wordt gedaan met stapsgewijze regressie. Die reeksen worden als regressoren gebruikt in een regressiemodel.
- Door aan het regressiemodel een dummy-variabele toe te voegen, kan het methode-effect worden geschat. Een dummy-variabele heeft de waarde 1 voorafgaand aan de breuk, en 0 in de periode daarna. De geschatte parameter van de dummy-variabele is het methode-effect.
- Als het methode-effect is geschat, kan een gecorrigeerde reeks worden berekend door van de oorspronkelijke reeks het methode-effect op te tellen of af te trekken (afhankelijk van de richting van het methode-effect).

Aan de hand van voorbeelden kan dit geïllustreerd worden. We beginnen met een analyse van uitbijters en breuken in de reeksen van relatieve armoede en gezonde levensverwachting.

Uitbijters:

In de statistiek voor relatieve armoede en gezonde levensverwachting zijn voor 28 landen reeksen te vinden in de database van Eurostat. Voor het vinden van uitbijters is een functie `tsclean` in het softwarepakket R beschikbaar. Naast het vinden van uitbijters geeft deze functie meteen een alternatieve schatting van het datapunt met de uitbijter. Er zijn in de voorbeeldreeksen voor een aantal landen uitbijters gevonden (tabel 4.1.1).

4.1.1 Uitbijters in de statistiek voor relatieve armoede en gezonde levensverwachting

Nr.	Reeks	Landen met minimaal 1 uitbijter
1	Relatieve armoede	België (2006), Hongarije (2006), Duitsland (2006), Zweden(2007)
7	Gezonde levensverwachting	België (2020), Bulgarije (2016, 2022), Kroatië (2013), Denemarken (2012, 2019), Hongarije (2023)

Breuken:

De breuken voor de twee voorbeeldreeksen worden nu meer diepgaand geanalyseerd. Breuken helemaal aan het begin van de reeks zijn niet relevant omdat ze geen invloed op de trend hebben die op basis van de gerepareerde reeksen gemaakt zal worden. Voor de reeksen van relatieve armoede is een overzicht te vinden van het aantal breuken per jaar en land in tabel 4.1.2.

4.1.2 Breuken in de reeksen per land voor de statistiek relatieve armoede

Landen	Breuken bij reeksen Relatieve armoede (Nr. 1)
Nederland, Bulgarije	2016
België, Ierland, Hongarije	2019
Denemarken, Duitsland, Frankrijk	2020
Luxemburg, Kroatië	Respectievelijk 3 en 2 breuken in 2020-2024
Bulgarije, Roemenië, Zweden	Niet relevante breuken in 2006, 2007, 2008
De overige 14 landen	Geen breuken
Verenigd Koninkrijk	Geen recente cijfers

Voor de reeksen van gezonde levensverwachting is een overzicht gegeven in tabel 4.1.3.

4.1.3 Breuken in de reeksen per land voor de statistiek gezonde levensverwachting

Landen	Breuken bij reeksen Gezonde levensverwachting (Nr. 7)
Roemenië, Slovenië, Zweden	2010
Zweden	2014
Kroatië	2012
Letland	2013
Italië	2017
Nederland, Italië, Bulgarije	2016
Denemarken	2020
Hongarije, Bulgarije, Polen	2022
België	2019 (uitbijter)
Frankrijk, Duitsland	2015, 2020, 2022, 2023
De overige 10 landen	Geen breuken
Verenigd Koninkrijk	Geen recente cijfers
Luxemburg	3 breuken in 2020-2023 en veel fluctuatie
Portugal	Alleen uitbijters in 2012 en 2013

Voor het repareren geldt het algemene principe dat reeksen van landen die in hetzelfde jaar een breuk hebben als het land waarvan de reeks gerepareerd dient te worden, niet als regressoren kunnen worden gebruikt. Dus Bulgarije kan bij relatieve armoede niet als regressor worden gebruikt voor de reparatie van de reeks van Nederland, omdat het eveneens in 2016 een breuk heeft. Hetzelfde geldt voor de reeksen van gezonde levensverwachting. Daarbij kan ook de reeks van Italië niet gebruikt worden voor breukreparatie van de Nederlandse reeks.

Alle reeksen van landen waarbij geen breuken optreden, kunnen bij uitstap wel gebruikt worden als regressor. Reeksen waarbij de breuk in een ander jaar plaatsvindt, kunnen opgenomen worden als regressor, mits er voor die breuk een extra dummy wordt meegenomen in de schatting. Die dummy wordt niet gebruikt in de reparatieslag, zoals eerder beschreven, maar wordt alleen als regressor gebruikt om de breuk van de regressor "op te vangen".

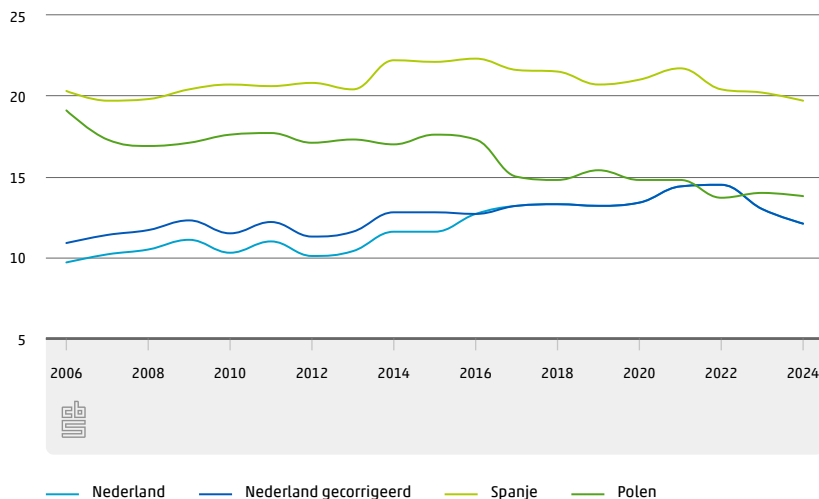
4.2 Voorbeelden van breuken en uitbijters

De reeks voor de relatieve armoede in Nederland wordt weergegeven in grafiek 4.2.1. Eurostat geeft door middel van "flags" aan dat er sprake is van een breuk. Een groot aantal reeksen voor andere landen bevatten echter geen breuk voor 2016. Veel reeksen bevatten voor geen enkel jaar een breuk. Uit een stapsgewijze regressie, waarbij alleen de sterkst gecorreleerde reeksen overblijven, blijkt dat met name de reeks voor Spanje en Polen sterk gecorreleerd zijn met de reeks van Nederland. Deze reeksen worden ook weergegeven in grafiek 4.2.1.

Met de stapsgewijze regressie kan ook het effect worden bepaald voor de breuk in 2016. Deze wordt vastgesteld op -1,21. Dat betekent dat door de breuk de data van de jaren voorafgaand aan 2016 systematisch een waarde van -1,21 te laag zijn ten opzichte van de data van 2016 en later. We nemen aan dat we de meest recente data als correct beschouwen, en verhogen daarom alle waarden voor de reeks voorafgaand aan het jaar 2016 met het getal 1,21. Dit resulteert in de gecorrigeerde reeks 'Nederland gecorrigeerd'. Deze reeks wordt weergegeven in grafiek 4.2.1.

4.2.1 Relatieve armoede met correctie voor een breuk

% van de bevolking beneden nationale armoedegrens

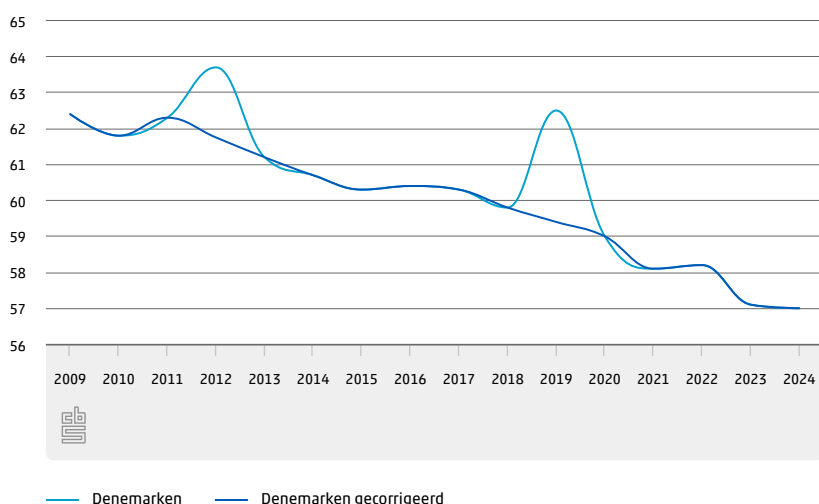


Opgemerkt moet worden dat de samenhang met de andere reeksen, alleen in relatieve zin wordt meegenomen. Het is dus niet relevant dat de reeksen van Spanje en Polen veel hoger liggen. Het draait allemaal om de ontwikkeling in de tijd, de groeicijfers, niet om de absolute waarde.

In sommige gevallen zijn er geen methodebreuken maar wel uitbijters. Dat zijn extreme waarden die om verschillende redenen kunnen optreden. Het is gebruikelijk om voorafgaand aan de trendbepaling deze uitbijters te vervangen door een schatting. Een voorbeeld is hieronder weergegeven voor de reeks van Gezonde Levensverwachting (GLV) in Denemarken. Deze reeks bevat volgens onze analyse twee uitbijters: in 2012 en 2019. Deze waarden worden vervangen door een schatting. In de praktijk komt dit neer op een interpolatie. Dit resulteert in de gecorrigeerde reeks 'Denemarken gecorrigeerd'.

4.2.2 Gezonde levensverwachting met correctie voor uitbijters

jaren (bij geboorte)



4.3 Technische toelichting

In deze toelichting wordt een methodologische verantwoording gegeven van de gevolgde werkwijze die is toegepast bij het corrigeren van breuken. Allereerst wordt de methode geschetst, en vervolgens wordt deze gebruikt voor een concreet voorbeeld.

4.3.1 Methode

Alvorens de breuken te repareren, worden eventuele uitbijters in de reeks geïdentificeerd en vervangen door een schatting. Nadat de uitbijters zijn verwijderd, worden de breuken aangepakt door modelmatig dummy-variabelen toe te voegen aan een model dat een schatting levert van de reeks. Als we normaliter een reeks kunnen weergeven met het model:

$$Y = X \times \beta + \epsilon$$

Dan kunnen we door het toevoegen van een dummy-variabele (D) de breuk weergeven:

$$Y = X \times \beta + D \times \delta + \epsilon$$

We kunnen dit model schatten en dat levert geschatte waarden op, zeg b en d, voor de parameters van het model (β en δ).

Het effect van de breuk wordt vaak weergegeven door de variabele D (dummy) met de parameter δ . Voorafgaand aan het jaar van de breuk krijgt de variabele D de waarde 1, en vanaf het breukjaar de waarde 0. De reden hiervoor is dat het uitgangspunt is dat de meest recente cijfers als het beste worden gezien, en de oudere cijfers dus gecorrigeerd moeten worden.

Vervolgens wordt de breuk in de reeks gerepareerd door de gecorrigeerde waarde voor Y (Y_c) te berekenen:

$$Y_c = Y - D \times d$$

(d is de geschatte waarde van δ)

Opgemerkt moet worden dat in reeksen soms met name aan het begin veel ontbrekende waarden en breuken zijn. Het heeft weinig zin om die te repareren. Enerzijds is dat technisch lastig, omdat er dan te weinig datapunten van zowel voor als na de breuk zijn. Anderzijds is het ook niet zinvol omdat de reeksen worden gebruikt om een trend te bepalen over de laatste 8 jaar, 2017-2024. In de praktijk betekent dit dat we voor de reeks van relatieve armoede vanaf 2006 gaan werken met de data, en voor gezonde levensverwachting vanaf 2009. Bij de laatste reeks zijn er bij veel reeksen breuken in 2008, die doorgaans niet goed geschat kunnen worden.

4.3.2 Uitwerking van de methode voor relatieve armoede in Hongarije

Bij wijze van voorbeeld van de methode om voor breuken te corrigeren, wordt hier het resultaat voor relatieve armoede in Hongarije weergegeven. Allereerst is er een behoorlijke uitbijter aan het begin van de reeks. Deze is vervangen door een schatting.

Vervolgens is geconstateerd dat er volgens Eurostat een breuk is in 2019. Hoewel Eurostat geen duidelijke definitie geeft van wat de breuk betekent, is in de data duidelijk waarneembaar dat 2018 een stuk lager ligt dan 2019. Vervolgens wordt op basis van correlaties en stapsgewijze regressie bepaald welke reeksen voldoende samenhangen met die van Hongarije, opdat een adequate schatting van de breuk kan plaatsvinden. Van alle geteste variabelen blijken de reeksen van relatieve armoede in Polen en Slovenië de beste schatting op te leveren. Hieronder zijn de schattingsresultaten weergegeven. De t-waarde geeft aan of een parameter significant van de waarde 0 verschilt en is daarmee een maat voor de sterkte van de correlatie. Dat is bij alle regressoren het geval, zelfs zeer sterk bij Slovenië (t-waarde is 10,1). De mate van significantie wordt met sterretjes aangegeven (*). De VIF-score geeft aan of er sprake is van multicollineariteit. Multicollineariteit betekent dat regressoren onderling teveel correleren, waardoor de parameters van het model niet meer goed geschat kunnen worden. De gevonden VIF-scores liggen onder 5 en er kan dus geconcludeerd worden dat er geen sprake is van multicollineariteit. De Adjusted R² geeft aan in welke mate het model de data door het model correct geschat wordt. Een waarde van 0,85 betekent dat 85% van de variantie door het model wordt verklaard. Dat is bijzonder hoog en goed.

4.3.2.1 Schattingsresultaten voor het model van relatieve armoede in Hongarije

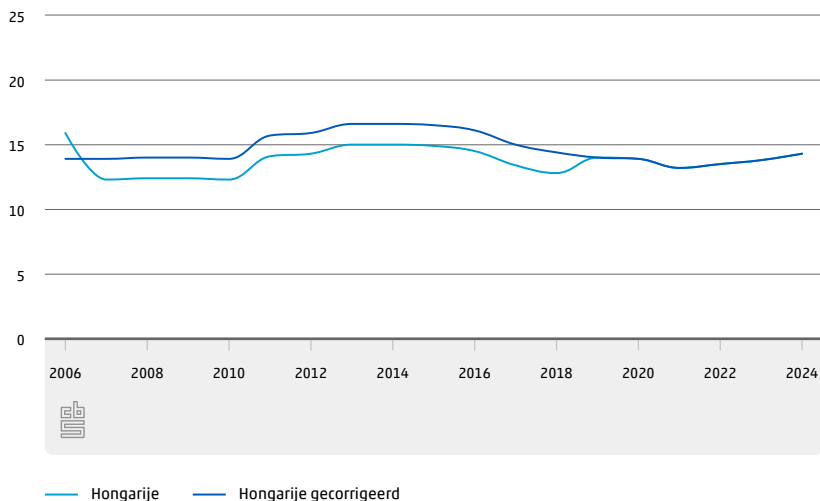
Regressor	Parameter	t-waarde	Significantie	VIF
Intercept	-1,3	-0,68	-	-
Polen	0,23	2,46	*	2,85
Slovenië	0,956	10,1	***	1,2
D_2019	-1,6	-4,83	***	3,2

Adjusted R² = 0,85

De geschatte coëfficiënt D_2019 is negatief, voor een dummy die 1 is voor jaren 2005 t/m 2018 en 0 voor de jaren er na. Gevolg is dat de gecorrigeerde reeks hoger ligt dan de originele reeks voor de periode, 2006-2018, waarin de dummy gelijk aan 1 is. Voor de periode dat de dummy 0 is, is de waarde gelijk aan de originele reeks. De gecorrigeerde reeks is in grafiek 4.3.2.2 weergegeven als 'Hongarije gecorrigeerd'.

4.3.2.2 Relatieve armoede, originele en gecorrigeerde data

% van de bevolking beneden nationale armoedegrens

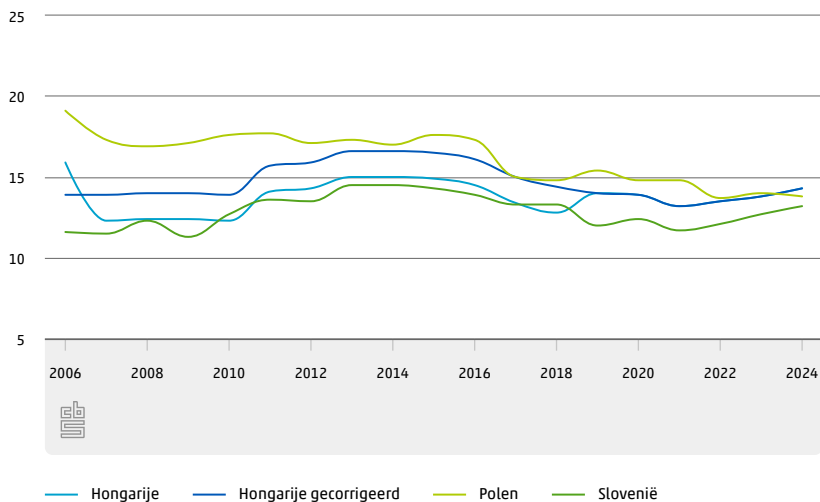


Als we naar grafiek 4.3.2.2 kijken dan zien we direct dat het verwijderen van de uitbijter een groot effect heeft en dat de grafiek vloeiender doorloopt tussen 2018 en 2019. Wat ook opvalt is de sterke stijging tussen 2010 en 2011. Het lijkt zeker mogelijk dat ook daar een methodebreuk heeft plaatsgevonden, hoewel dat niet uit de informatie van Eurostat blijkt. Als er een model geschat zou worden met een extra dummy voor dat tijdvak, zou dat zeker tot een correctie kunnen leiden.

Als we echter de reeksen van Polen en Slovenië ook in beeld brengen, ontstaat een genuanceerder beeld. Het is duidelijk zichtbaar dat met name Slovenië en Hongarije behoorlijk gelijk oplopen in de reeks van relatieve armoede. We zien ook bij Slovenië een behoorlijke stijging in het tijdvak 2009-2011. Het suggereert dat de sterke stijging in 2010 bij Hongarije mogelijk toch een realistisch beeld geeft. Mogelijk is de sterke stijging een gevolg van de kredietcrisis in 2008.

4.3.2.3 Relatieve armoede, originele en gecorrigeerde data

% van de bevolking beneden nationale armoedegrens



Hiermee wordt ook inzichtelijk dat het proces geen automatisch rekenproces is, maar dat per reeks gekeken moet worden naar wat een realistisch verloop zou kunnen zijn. Een keihard bewijs dat er geen methodebreuk was in 2010 in Hongarije is er niet, maar vanwege de gerede twijfel hierover wordt de informatie van Eurostat zonder wijziging overgenomen.

Publicatiedatum 13-4-2026 06:30

Vergelijking internationale tijdreeksen Monitor Brede Welvaart

5. Conclusies en aanbevelingen

In dit rapport is een methode voor het bepalen van de bruikbaarheid van internationale reeksen in een trend-positie vergelijking voor indicatoren uit de MBW onderzocht. Zo zijn er kwaliteitscriteria opgesteld waaraan indicatoren kunnen worden getoetst: nauwkeurigheid en plausibiliteit, tijdigheid, consistentie over tijd en vergelijkbaarheid tussen landen. Ook is een methode ontwikkeld om van reeksen met breuken en uitbijters een betrouwbare trend te kunnen bepalen.

In de beoordeling van de geschiktheid van een indicator is de laagste classificering van de kwaliteitscriteria doorslaggevend. Van de onderzochte indicatoren waren er vier indicatoren die afvielen: 'toegevoegde waarde milieusector', 'vogels boerenland', 'totale woonquote' en 'stikstofoverschot'. Voor twee van deze indicatoren werd ook in de metadata van Eurostat een vergelijking tussen landen afgeraden. Aan de andere kant waren er ook veel indicatoren met een bovengemiddelde ofwel hoge vergelijkbaarheid. Wanneer er aan een indicator een gemiddelde bruikbaarheid wordt toegekend moet voorzichtig worden omgegaan met het trekken van conclusies op basis van een trend-positie vergelijking. In deze analyse was er één indicator welke in zijn totaliteit als gemiddeld geclassificeerd werd: 'neonatale sterfte'.

Aan de hand van de in dit onderzoek geformuleerde criteria kunnen in de toekomst meer indicatoren op eenzelfde wijze op bruikbaarheid worden getoetst. De metadata van Eurostat is in de analyse erg behulpzaam maar het doornemen hiervan kost relatief veel tijd. Om de informatie op waarde te kunnen schatten is het hiernaast voor indicatoren nodig geweest om de conclusies te toetsen bij experts binnen het CBS.

Voordat de metadata worden geanalyseerd kan op basis van meetbare minimale kwaliteitscriteria een eerste selectie worden gemaakt voor de toepasbaarheid van indicatoren in een trend-positievergelijking. Verder zou ervoor gekozen kunnen worden om indicatoren op basis van het achterliggende raamwerk te classificeren. Het achterliggende raamwerk en de looptijd hiervan blijkt veel informatie te geven over de indicatoren, hiernaast zijn voor een aantal indicatoren ook alleen metadata beschikbaar van het overkoepelende raamwerk en niet van de indicator zelf. Of de aanname van een gelijke classificering van de indicatoren binnen een raamwerk plausibel is, kan per raamwerk worden afgewogen.

Verder kan de bruikbaarheid van een indicator veranderen in de loop van de tijd. Zo kan een statistiek zich verder ontwikkelen (bijvoorbeeld 'toegevoegde waarde milieusector') of juist in nauwkeurigheid achteruit gaan in de tijd doordat minder landen data kunnen aanleveren ('vogels boerenland') of de kwaliteit van de brondata kan wijzigen ('invoer fossiele energiedragers'). De bruikbaarheid van een indicator zal om deze reden bij voorkeur periodiek moeten worden herzien.

Voor toepasbaarheid in een trend-positievergelijking moet zowel de positie als de trend van een indicator te vergelijken zijn. Wanneer in een reeks breuken en uitbijters zitten kan er geen vergelijkbare trend worden bepaald. Met de in dit onderzoek ontwikkelde methode kunnen reeksen met behulp van een dummy-variabele worden geschat om te komen tot een gecorrigeerde reeks. Met deze consistente reeks kunnen betrouwbare trends worden geschat. Voor het bepalen van de trend is de data gecorrigeerd voor zowel breuken als uitbijters. Voor het bepalen van de onderlinge positie van landen kan ook een correctie worden gedaan van uitbijters van andere landen dan Nederland, dit gebeurt momenteel niet. Hierbij moeten de data van landen echter wel geschat worden en aangepast in de gepresenteerde reeksen, wat bij het bepalen van de trend niet het geval was.

De ontwikkelde methode voor het bepalen van betrouwbare trends in dit onderzoek kan direct worden toegepast op andere indicatoren. Wel is het toepassen van de methode mensenwerk en zal er geen sprake zijn van een volledig geautomatiseerd proces. Voorwaarde voor deze methode is ook dat Eurostat de methodebreuken blijft aangeven.

Indicatoren in de MBW kunnen worden getoetst aan de in dit onderzoek opgestelde criteria en reeksen kunnen voor een nauwkeurige trendbepaling worden ontdaan van breuken en uitbijters. Na een toetsing aan de kwaliteitscriteria en na eventuele correctie van de reeksen voor het berekenen van de trend, kan een indicator worden opgenomen in een trend-positie vergelijking in publicaties. De MBW bevat veel indicatoren en bovendien moet de bruikbaarheid periodiek worden herzien. Er zijn meerdere manieren voorgesteld voor het maken van een efficiëntieslag in de beoordeling van een indicator. Zo kan de controle van kwantitatieve kwaliteitscriteria worden geautomatiseerd, en kan ervoor worden gekozen om de focus te verleggen naar het overkoepelende raamwerk. Of de aanname, van een gelijke indeling van klassen voor de verschillende indicatoren binnen een raamwerk, opgaat voor het specifieke raamwerk kan worden geverifieerd bij een expert binnen het CBS.

Publicatiedatum 13-4-2026 06:30

Vergelijking internationale tijdreeksen Monitor Brede Welvaart

Bijlage A

A.1 Uitbijters en breuken in de reeksen van Eurostat

Nr	Indicator	# Landen met uitbijters	Breuken NL 2005-2024	# breuken	# landen met breuken	# landen met meer dan 5 breuken	Opmerking
1	Relatieve armoede	4	2016	38	18	1	vanaf 2006: 27 breuken
2	Toegevoegde waarde milieusector	2	0	6	5	0	
3	Individuele consumptie	0	0	1	1	0	
4	Vogels boerenland	5	0	0	0	0	
5	Gewerkte uren in de zorg	7	0	3	2	0	
6	Neo-natale sterfte	5	0	0	0	0	
7	Gezonde levensverwachting mannen	5	2008, 2016	49	25	1	vanaf 2009: 31 breuken, in 17 reeksen
8	Langdurige werkloosheid	2	0	20	12	1	
9	Totale woonquote (huur en koop)	2	2016	23	16	1	
10	Milieuproblemen	3	2024	23	19	1	
11A	Gemiddelde schuld per huishouden	0	0	0	0	0	
11B	Gemiddelde schuld per huishouden	1		0	0	0	
12	Stikstofoverschot	2	0	0	0	0	
13	Invoer fossiele energiedragers	3	0	1	1	0	