

Methodebeschrijving tijdreeks productie-index van de industrie

10

R.H.M. van der Stegen

Publicatiedatum CBS-website: 11 mei 2010



Verklaring van tekens

.	= gegevens ontbreken
*	= voorlopig cijfer
**	= nader voorlopig cijfer
x	= geheim
–	= nihil
–	= (indien voorkomend tussen twee getallen) tot en met
0 (0,0)	= het getal is kleiner dan de helft van de gekozen eenheid
niets (blank)	= een cijfer kan op logische gronden niet voorkomen
2008–2009	= 2008 tot en met 2009
2008/2009	= het gemiddelde over de jaren 2008 tot en met 2009
2008/'09	= oogstjaar, boekjaar, schooljaar enz., beginnend in 2008 en eindigend in 2009
2006/'07–2008/'09	= oogstjaar, boekjaar enz., 2006/'07 tot en met 2008/'09

In geval van afronding kan het voorkomen dat het weergegeven totaal niet overeenstemt met de som van de getallen.

Colofon

Uitgever

Centraal Bureau voor de Statistiek
Henri Faasdreef 312
2492 JP Den Haag

Prepress

Centraal Bureau voor de Statistiek - Grafimedia

Omslag

TelDesign, Rotterdam

Inlichtingen

Tel. (088) 570 70 70
Fax (070) 337 59 94
Via contactformulier: www.cbs.nl/infoservice

Bestellingen

E-mail: verkoop@cbs.nl
Fax (045) 570 62 68

Internet

www.cbs.nl

Methodebeschrijving tijdreeks productie-index van de industrie

Samenvatting: Deze nota beschrijft de methode voor het maken van een lange tijdreeks van de productie-index van de industrie. Als eerste wordt een korte beschrijving gegeven van de productie-index. Dan volgt een inventarisatie van de beschikbare data. Het derde punt is een beschrijving van de methode om de verschillende reeksen te koppelen. Hierbij komen onder andere de uitgangspunten, de SBI-herzieningen in de loop van de tijd en de methode voor het uiteindelijke koppelen aan bod.

Afkortingen

ABR: algemeen bedrijvenregister
NR: nationale rekeningen
PI: productie-index van de industrie

1. Inleiding

Het CBS stelt al sinds de jaren vijftig maandelijks productie-indices samen van de industriële productie. Vanaf 1969 meet de productie-index (PI) de volume ontwikkeling van de toegevoegde waarde van de industrie. Als er in dit rapport wordt gesproken over productie wordt eigenlijk het volume van de toegevoegde waarde bedoeld. De productie-index industrie is een veel gebruikte indicator voor de kortetermijn golfbeweging in de economie en daarmee een van de belangrijkste conjuncturele maandindicatoren. Omdat de structuur van de economie en meetmethoden veranderen zijn er basisverleggingen nodig. Er is gekozen om deze om de ongeveer 5 jaar door te voeren. Bij deze vijfjaarlijkse basisverleggingen van de PI is steeds een nieuwe reeks gestart en niet gekoppeld aan de oudere reeksen. Er is wel een overlappende periode waarin beide reeksen zijn berekend. Belangrijke gebruikers van de PI zijn DNB, CPB, economische onderzoeksbureaus en afdelingen van bijvoorbeeld grote banken en economisch onderzoekers uit de universitaire wereld. Deze gebruiken voor veel conjunctuuronderzoek de PI onder andere 1) als referentiereeks voor het dateren en beschrijven van de conjunctuur, 2) als basis voor het schatten van tijdreeksmodellen en 3) als basis voor het onderzoeken van de samenhang van bepaalde variabelen met de hoofdconjunctuur.

Sommige van deze instellingen hebben bij gebrek aan officiële tijdreeksen zelf de oude PI-reeksen doorgesplitst, waardoor er inmiddels meerdere varianten van in omloop zijn. In andere gevallen gebruiken m.n. universitaire onderzoekers bijvoorbeeld reeksen van de VS voor het schatten van modellen.

In het kader van het speerpunt Lange Tijdreeksen is een officiële, consistente CBS-indexreeks gemaakt van de industriële productie vanaf de jaren vijftig, o.b.v. het door het CBS gepubliceerde materiaal.

1.1 Opbouw van deze nota

Na deze inleiding wordt begonnen met een kort overzicht van definities (paragraaf 2) en met het inventariseren van de beschikbare data (paragraaf 3). Daarna wordt de methode beschreven waarmee de reeksen worden gekoppeld waarbij rekening gehouden wordt met de uitgangspunten en de SBI-herziening.

2. Theoretische achtergrond en definities

De productie-index is een maandstatistiek die de volumeontwikkeling van de productie van een bedrijfstak in de loop van de tijd beschrijft. Er wordt daarbij gewerkt met een

basisjaar waarin de jaarproductie op 100 wordt gesteld ¹⁾. Gedurende een aantal jaren wordt de volumeontwikkeling uitgedrukt ten opzichte van dit basisjaar. Na verloop van een aantal jaren wordt een nieuw basisjaar gedefinieerd en daarmee een nieuwe reeks gestart. Het concept van basisjaar en –verlegging loopt sterk parallel aan dat bij de consumenten- of producentenprijsindex.

De ontwikkeling van de productie wordt geraamd via waarneming van de ontwikkeling in de volgende grootheden:

- Omzet en prijzen
- Gewerkte uren
- Geproduceerde hoeveelheden
- Orders

Hieronder wordt de raming uitgewerkt op basis van omzet (of eigenlijk wordt dan gedefleerde omzet ²⁾ bedoeld). De raming zou op analoge wijze kunnen worden uitgevoerd door i.p.v. omzet één van de andere bovengenoemde grootheden te nemen.

Elke maand i van jaar j wordt omzet o_{ij} waargenomen bij bedrijven in een bedrijfstak. Na afloop van het basisjaar wordt de gemiddelde omzet O_j opgemaakt uit de 12 o_{ij} 's in het basisjaar. De productie-index voor maand i in het basisjaar is nu gedefinieerd als:

$$(1): p_{i,j} = \frac{o_{i,j}}{O_j} * 100$$

Naast de productie-index als zodanig kennen we ook het groeicijfer van de productie in maand i ten opzichte van dezelfde maand een jaar eerder:

$$(2): q_{i,j} = \frac{p_{i,j}}{p_{i,j-1}} = \frac{o_{i,j}}{o_{i,j-1}}$$

Daarmee kan p_i in de latere jaren ook worden berekend door de jaarlijkse groeivoeten q_i te koppelen aan de PI in het basisjaar. In de praktijk wordt die laatste aanpak gevolgd, waarbij dan in vergelijking (2) de omzetten worden meegenomen van alle bedrijven die in beide waarnemingen voorkomen. Ook bij het koppelen van reeksen met verschillende basisjaren (zie Hoofdstuk 4) zal een dergelijke methode gevolgd worden.

Merk op dat, p_{ij} hetzelfde seizoenspatroon vertoont als o_{ij} , daarentegen zijn in q_{ij} alleen de veranderingen van het seizoenpatroon van jaar tot jaar zichtbaar.

Nadat voor k verschillende bedrijfstakken $p_{i,j,k}$ is opgemaakt kunnen we ook een geaggregeerde productie-index $P_{i,j}$ berekenen:

$$(3): P_{i,j} = \sum_k a_k p_{i,j,k}$$

De coëfficiënten a_k staan bekend als het wegingschema dat bepaald wordt in het basisjaar. De coëfficiënten worden afgeleid uit de verhoudingen van de toegevoegde waarde tussen de verschillende bedrijfstakken in de nationale rekeningen.

3. Inventarisatie beschikbare data

De beschikbare data over de periode vanaf 1953 zijn in verschillende opzichten niet homogeen. Zo zijn er voor verschillende bedrijfstakken verschillen in detailniveau over de tijd, bijvoorbeeld 4 digit SBI versus 2 digit SBI ³⁾. Voorts zijn er veranderingen in de SBI-indeling zelf. Het aantal versies/updates dat van een reeks beschikbaar is, verschilt als gevolg van de bewaarstrategie gevolgd bij dat basisjaar. Indien er verschillende versies beschikbaar zijn, speelt vooral de laatste update voor aanpassing aan NR een belangrijke rol. In

¹⁾ De productie gemiddeld over het basisjaar is 100, echter omdat de productie per maand verschilt, zijn maandwaarden i.h.a. ongelijk aan 100.

²⁾ Merk op dat met omzet steeds bedoeld wordt de omzet na deflering met de producentenprijsindex; hoe die deflering plaatsvindt vergt een beschrijving op zichzelf, maar is in deze context niet van belang en wordt daarom buiten beschouwing gelaten.

³⁾ Het aantal digit SBI geeft de mate van detail van de bedrijfstak indeling weer. 4 digit SBI is gedetailleerder dan 2 digit. Een 2 digit code kan worden afgeleid uit een langere, bijvoorbeeld 4 digit code door de laatste, in dit geval 2, cijfers weg te laten.

deze sectie wordt eerst opgesomd welke reeksen volgens de huidige bewaarstrategie aanwezig zouden kunnen zijn en vervolgens welke reeksen ook daadwerkelijk aanwezig zijn.

3.1 *Theoretische versies van de data*

3.1.1 *Bewerkingen op de data*

Elke maand wordt een zogenaamde 'oorspronkelijke' reeks PI gemaakt, daaruit een 'gemiddelde dag' reeks van de PI en uit de 'gemiddelde dag' reeks wordt een 'seizoensgecorrigeerde' reeks van de PI gemaakt. In een 'gemiddelde dag' reeks wordt de 'oorspronkelijke' reeks gecorrigeerd voor voornamelijk het verschil in werkdagen tussen de maanden. Hierdoor worden maanden onderling vergelijkbaar en dezelfde maand in verschillende jaren vergelijkbaar. Bij seizoenscorrectie wordt nog een extra correctie uitgevoerd waardoor de maanden onderling nog beter vergelijkbaar worden. Er wordt gecorrigeerd voor het feit dat in veel bedrijfstakken de productie nu eenmaal een inherent seizoenspatroon kent.

3.1.2 *Versies van de data*

Van de reeks maandcijfers productie-index bestaan verschillende versies:

1. De eerste raming; o.b.v. de data die binnen zijn 5 weken na het einde van de maand T
2. De tweede raming, o.b.v. de grotere hoeveelheid data die binnen zijn 5 weken na afloop van maand T+1
3. De derde raming, o.b.v. de grotere hoeveelheid data die binnen zijn 5 weken na afloop van maand T+2
4. Soms zijn er nog meer ramingen om te corrigeren voor nagekomen veranderingen in de data van de enquête. Dit is de meest nauwkeurige versie gebaseerd op microdata.
5. In deze versie wordt alleen nog maar op meso- en macroniveau aanpassingen aangebracht, microdata wordt niet meer gebruikt. De bewerkte data wordt aangepast aan de volumeontwikkeling van de toegevoegde waarde volgens NR: Wanneer het jaargroei cijfer voor een bepaald jaar in de NR blijkt af te wijken van de productie-index wordt de productie-index aangepast per bedrijfstak aan NR met een correctie die over het hele jaar hetzelfde is. Na correctie komt de gemiddelde waarde over het jaar van de PI overeen met de toegevoegde waarde van NR. Dit gebeurt bovendien meerdere keren, eerst na NR voorlopig, dan na NR nader voorlopig en na NR definitief jaar en vervolgens bij elke revisie.

3.1.3 *Aggregatie niveaus*

Het meest gedetailleerde niveau waarop de PI wordt samengesteld is een bedrijfsklasse-indeling op 4 digit niveau. Door middel van wegingsfactoren worden reeksen geaggregeerd.

3.1.4 *Basisjaren*

Ongeveer elke 5 jaar wordt er begonnen met het opnieuw vaststellen van een basisjaar en daarmee met een nieuwe reeks. Meestal worden de wijzingen in de methodiek van berekenen van de PI met een nieuwe reeks ingevoerd, zoals wijziging van de SBI.

3.2 *Aanwezige reeksen*

3.2.1 *Indeling naar basisjaren*

Voor de verschillende basisjaren zijn verschillende reeksen, op zowel 2 digit SBI als geaggregeerd ⁴⁾, beschikbaar zoals hieronder opgesomd. Bij elk basisjaar wordt er aangegeven vanaf wanneer er nieuwe data aan een reeks is toegevoegd (dwz data gebaseerd op nieuwe waarnemingen die nog niet eerder gepubliceerd is, dit in tegenstelling tot data die al gebruikt is bij een eerder basisjaar). Het eerste deel van de reeks wijkt af van het voorgaande basisjaar als gevolg van onder andere nieuwe classificaties en nieuwe

⁴⁾ Er wordt niet van 1 digit SBI maar van geaggregeerd SBI niveau gesproken omdat vanaf 2 digit volgens een andere indeling wordt geaggregeerd die niet overeenkomt met het weglaten van het laatste cijfer van de 2 digit code.

wegingen over dezelfde periode. Later in dit rapport zal alleen de nieuwe data gebruikt worden voor het maken van de tijdreeks.

- 1958=100, alleen gepubliceerde definitieve data vanaf januari 1953 t/m december 1965. Er is slechts 1 set definitieve reeksen gevonden waardoor het niet mogelijk is om de verschillende versies van de data te onderscheiden. Alleen gemiddelde dag data, geen oorspronkelijke data op 2 digit SBI en hoger geaggregeerd niveau. Deze reeksen zijn in 1966 beschreven in nr 5 van Statistische en econometrische onderzoeken.
- 1963=100, loopt vanaf januari 1963 t/m augustus 1974, nieuwe data zijn toegevoegd vanaf januari 1969, dit betekent dat de eerste publicatie van de data voor januari 1969 niet beschikbaar is, alleen de definitieve. Er is alleen data op 2 digit SBI en hoger geaggregeerd niveau beschikbaar. Deze reeksen zijn in 1968 beschreven in nr 1 van Maandstatistiek van de industrie.
- 1970=100, loopt vanaf januari 1970 t/m december 1979, nieuwe data zijn toegevoegd vanaf september 1974. Er is alleen data op 2 digit SBI en hoger geaggregeerd niveau beschikbaar. Deze reeksen zijn in 1974 beschreven in nr 10 van Maandstatistiek van de industrie.
- 1975=100 loopt vanaf januari 1973 t/m januari 1984, nieuwe data zijn toegevoegd vanaf september 1979. Er is alleen data op 2 digit SBI en hoger geaggregeerd niveau beschikbaar. Deze reeksen zijn in 1980 beschreven in Maandstatistiek van de industrie.
- 1980=100 loopt vanaf januari 1979 t/m augustus 1989, nieuwe data zijn toegevoegd vanaf januari 1984. Er is alleen data op 2 digit SBI en hoger geaggregeerd niveau beschikbaar. Deze reeksen zijn in 1984 beschreven in nr 8 van Maandstatistiek van de industrie.
- 1985=100 loopt vanaf januari 1985 t/m september 1994, nieuwe data zijn toegevoegd vanaf augustus 1989. Er is alleen data op 2 digit SBI en hoger geaggregeerd niveau beschikbaar. Deze reeksen zijn in 1990 beschreven in nr 1 van Maandstatistiek van de industrie.
- 1990=100 loopt vanaf januari 1990 t/m december 1998, nieuwe data zijn toegevoegd vanaf oktober 1994, de publicatiereeksen (2 digit en hoger geaggregeerd) worden niet gebruikt voor de tijdreeks omdat er vanaf 1990 op 4 digit niveau gedetailleerdere reeksen beschikbaar zijn.
- 1995=100 loopt vanaf januari 1990 t/m juli 2003, nieuwe data zijn toegevoegd vanaf januari 2000, de publicatiereeksen (2 digit en hoger geaggregeerd) worden niet gebruikt voor de tijdreeks omdat er op 4 digit niveau gedetailleerdere reeksen beschikbaar zijn.
- 2000=100 loopt vanaf januari 1995 t/m nu, nieuwe data zijn toegevoegd vanaf augustus 2003, de publicatiereeksen (2 digit en hoger geaggregeerd) worden niet gebruikt voor de tijdreeks omdat er op 4 digit niveau gedetailleerdere reeksen beschikbaar zijn.

De reden om de reeksen van de basisjaren 1990, 1995 en 2000 op 4 digit niveau te gebruiken is dat deze daarmee nauwkeuriger van SBI 93 naar SBI 2008 omgerekend kunnen worden. Voor 1990 is deze informatie niet beschikbaar en kan alleen een koppeling gemaakt worden tussen aggregaten waarbij de veranderingen ten opzichte van SBI 2008 relatief klein zijn.

3.2.2 Indeling naar bedrijfstakken

Vanaf 1990 is er voldoende informatie beschikbaar voor het samenstellen van reeksen (zowel oorspronkelijk, als gemiddelde dag als seizoensgecorrigeerd) voor alle bedrijfstakken. Voor 1990 zijn er alleen gemiddelde dag en oorspronkelijke reeksen voor enkele bedrijfstakken:

- Nijverheid en energie (exclusief bouw) gemiddelde dag vanaf basisjaar 1958 en oorspronkelijk vanaf 1963
 - Delfstoffenwinning gemiddelde dag vanaf basisjaar 1958 (SBI 06-09)
 - Energiebedrijven en afvalbeheer gemiddelde dag vanaf basisjaar 1958 (SBI 35–36)
 - Industrie gemiddelde dag vanaf basisjaar 1958 en oorspronkelijk vanaf 1963 (SBI 10–33)
 - Voeding- en genotmiddelen gemiddelde dag vanaf basisjaar 1958 (SBI 10–12)
 - Vlees- en zuivelindustrie gemiddelde dag vanaf basisjaar 1958 (SBI 101+105)
 - Drankenindustrie gemiddelde dag vanaf basisjaar 1958 (SBI 11)

- Tabaksindustrie gemiddelde dag vanaf basisjaar 1958 (SBI 12)
- Overige voedingsmiddelenindustrie gemiddelde dag vanaf basisjaar 1958 (SBI 102–104, 106–109)
- Textielindustrie gemiddelde dag vanaf basisjaar 1958 (SBI 13)
- Kledingindustrie gemiddelde dag vanaf basisjaar 1970 (SBI 14)
- Leder- en schoenenindustrie gemiddelde dag vanaf basisjaar 1970 (SBI 15)
- Papierindustrie gemiddelde dag vanaf basisjaar 1958 (SBI 17) ⁵⁾
- Grafische industrie gemiddelde dag vanaf basisjaar 1990 (SBI 18)
- Aardolie-industrie gemiddelde dag vanaf basisjaar 1970 (SBI 19)
- Chemie en farmaceutische industrie gemiddelde dag vanaf basisjaar 1958 (SBI 20+21)
- Rubber en kunststofproductenindustrie gemiddelde dag vanaf basisjaar 1958 (SBI 22)
- Bouwmaterialenindustrie gemiddelde dag vanaf basisjaar 1958 (SBI 23)
- Metaal gemiddelde dag vanaf basisjaar 1958 (SBI 24–30+33)
 - Basismetaal gemiddelde dag vanaf basisjaar 1970 (SBI 24)
 - Overige metaal gemiddelde dag vanaf basisjaar 1970 (SBI 25–30+33)

3.2.3 SBI herzieningen

In de periode 1953 t/m 2008 zijn er verschillende bedrijfstakindelingen gebruikt. De tijdreeks wordt gepubliceerd onder de meest recente indeling, namelijk SBI 2008. De verschillen tussen de voorgaande indelingen worden hieronder weergegeven:

- Vanaf 2009 wordt met SBI 2008 gewerkt. In de bijlage wordt een overzicht gegeven in werkzame personen van de verschuivingen tussen de verschillende bedrijfstakken als gevolg van de overgang van SBI 93 op SBI 2008. De belangrijkste verschillen tussen SBI 2008 en SBI 93 zijn:
 - De splitsing van uitgevers en drukkerijen. De uitgevers (221 in SBI 93) gaan uit de industrie. De uitgevers en drukkerijen zijn als aparte reeksen vanaf januari 1990 beschikbaar.
 - De splitsing van de reparatie in de metaal (34+35 onder SBI93). Dit wordt een aparte SBI 33 onder SBI 2008 maar blijft binnen de metaalindustrie. Door het samen nemen van reeksen levert dit geen probleem op.
 - De sloop, 37 uit SBI 93 gaat uit de industrie volgens SBI 2008. Echter volgens de transitie matrix is dit slechts een kleine bedrijfstak ten opzichte van het totaal van de publicatiereeks SBI 36+37. Daarom wordt dit verwaarloosd.
 - De nijverheid onder SBI 2008 bevat officieel geen energiebedrijven en afvalbeheer (35–36 in SBI 2008), terwijl deze bedrijfstak bij basisjaren t/m 2000 dat wel het geval was. Er wordt daarom besloten ook voor de periode ná 2000 de nijverheid inclusief energiebedrijven en afvalbeheer te publiceren.
- Reeksen met basisjaar 1990, 1995 en 2000 zijn op basis van SBI 93
- Er zijn geen grote verschillen tussen SBI 74 en SBI 93. Met SBI 74 zijn de volgende reeksen met basisjaren gemaakt: 1970 t/m 1985
- Reeksen met basisjaren 1958 en 1963 zijn op basis van bedrijfstaktellingen gemaakt. Het is onbekend hoe groot de verschillen zijn met SBI 74.
- Nationale rekeningen wegingen zijn beschikbaar op consistente basis vanaf 1969. Echter in de periode tot 1987 alleen voor een geaggregeerde bedrijfstakindeling van 40 bedrijfstakken.

4. Methode

Voordat de methode voor het maken van de tijdreeks beschreven wordt, worden eerst de uitgangspunten waaraan de methode moet voldoen op een rij gezet. De methodebeschrijving valt uiteen in het bepalen van de groei van de productie en vervolgens het koppelen van de groei tot een reeks. Daarna wordt de ijking op volumeontwikkelingen van de toegevoegde waarde uit de nationale rekeningen beschreven.

⁵⁾ Onduidelijk is of papier in 1958 bestaat uit alleen papier of de combinatie papier en grafische industrie. Voor de zekerheid is de reeks met basisjaar 1958 weggelaten en begint de tijdreeks de informatie uit de reeks met basisjaar 1963.

4.1 Mogelijke methoden voor samenstellen van de tijdreeks

Er zijn verschillende mogelijkheden voor het samenstellen van de tijdreeks productie-index. In deze paragraaf worden een aantal mogelijke oplossingsstrategieën besproken waarna in de volgende paragrafen de definitieve oplosmethode wordt uitgewerkt:

1. De meest voor de hand liggende oplossing is om de reeksen van een bepaald basisjaar als uitgangspunt te nemen, bijvoorbeeld het laatste basisjaar. Vervolgens worden de mutaties uit de reeksen van de voorgaande basisjaren gehaald en deze gekoppeld aan uitgangsreeksen. Als dit gedaan wordt, blijkt het seizoenspatroon in de reeksen langzaam op te blazen waardoor aan het eind van de reeks het seizoenspatroon een veel grotere amplitude heeft dan in de oorspronkelijke reeks.
2. De overlap tussen de reeksen van de opeenvolgende basisjaren kan gebruikt worden in een tijdsafhankelijk (structureel) tijdreeksmodel. Dit model berekent een tijdsafhankelijke correctie voor de te koppelen reeks. Omdat vervolgens hier aan nog weer een andere reeks wordt gekoppeld met een tijdreeksmodel krijg je correctiefactor op correctiefactor, en dit wordt nog een paar keer herhaald tot alle reeksen gekoppeld zijn. Aangezien er nog geen ervaring is met deze werkwijze is de kans van slagen niet groot. De kans is groot dat naar mate de correcties toenemen de oude reeksen steeds minder herkenbaar zijn.
3. Om de oude reeksen terug te kunnen herkennen in de tijdreeks is het nodig dat de veranderingen niet te groot worden in de tijd. Een methode om dit bereiken is het ongewijzigd overnemen van de basisjaren in de tijdreeks. Het opvullen van de periode tussen de basisjaren kan op verschillende manieren gedaan worden. Een paar suggesties zijn:
 - a. Door in de overlappende periode, het gemiddelde te nemen van de twee reeksen
 - b. Door een tijdreeksmodel een geleidelijke overgang te laten berekenen van de ene reeks naar de andere.
 - c. Door een correctiefactor te berekenen die per maand wordt toegepast zodanig dat de overgang van het ene basisjaar op het andere zonder breuken verloopt. Vanwege het gegarandeerde resultaat en de eenvoud wordt deze methode in de komende paragrafen verder uitgewerkt.

4.2 Uitgangspunten

De uitgangspunten voor het maken van een tijdreeks zijn:

- Microdata (in de zin van data van individuele bedrijven in de verschillende perioden) zijn niet meer beschikbaar, daardoor bestaat het basismateriaal waarmee de tijdreeksen gemaakt worden uit de gepubliceerde reeksen 2 digit en hoger geaggregeerd of, indien aanwezig, de 4 digit reeksen.
- De tijdreeks wordt opgebouwd uit de volgende data:
 - In de periode voor 1990 wordt de tijdreeks opgebouwd uit gepubliceerde reeksen, dus op een hoog aggregatieniveau (2 digit en hoger). Detailniveau is niet beschikbaar.
 - In de periode 1990 t/m nu wordt de tijdreeks opgebouwd uit deelreeksen op 3 of 4 SBI-digit niveau. Deze reeksen zijn in het algemeen niet gepubliceerd (de hogere aggregaten van deze reeksen natuurlijk wel). Met deze reeksen is een nauwkeurige implementatie van de overgang van SBI93 naar SBI2008 mogelijk die niet mogelijk is voor 1990.
- Omdat er voor 1990 geen onderliggende detailreeksen (SBI 3-digit en fijner) beschikbaar zijn, is het voor 1990 niet mogelijk te corrigeren voor:
 - Wijzigingen in classificaties door de invoering van nieuwe versies van de SBI. Hierdoor kunnen verschuivingen tussen de 2-digitgroepen binnen de industrie plaatsvinden. Een bepaald gedeelte van een bedrijfstak hoorde bijvoorbeeld bij de oude versie van de SBI wel tot een bepaalde bedrijfstak en bij de nieuwe versie niet meer. Omdat de onderliggende detailreeksen ontbreken en in een aantal gevallen voor de oude jaren ook de informatie over de SBI wijzigingen ontbreekt, is het niet mogelijk om zulke fouten te repareren of te detecteren. De enige oplossing kan zijn om een reeks niet te publiceren als het verschil groot is. Deze analyse is zeer subjectief omdat deze alleen gebaseerd is op de naam van de deel-bedrijfstak in het schakelschema die verschuift van SBI-groep.

- Wijzigingen in de scope van de industrie. Er worden activiteiten aan de industrie toegevoegd of verdwijnen eruit als gevolg van classificatiewijzigingen. Dit is vooral van invloed op de aggregaten nijverheid en industrie. Als het verschil te groot is, kan de enige oplossing zijn om een aggregaat niet te publiceren. Evenals als bij het vorige punt is deze analyse alleen gebaseerd op de naam van de deel-bedrijfstak in het schakelschema, en daardoor erg subjectief.
- Wijzigingen in definities, waarderingen en dergelijke. Deze hangen samen met NR-revisies, waarin bijvoorbeeld de productiegrens (minimale productie om in de enquête te zijn opgenomen) kan zijn veranderd (zoals de registratie van investeringen in eigen beheer) of de waardering (basisprijzen versus factorkosten e.d.).
- De ontwikkeling van het jaargemiddelde van de groei van de productie-index dient overeen te komen met de volumeontwikkeling van de toegevoegde waarde in de nationale rekeningen. Hierdoor wordt de consistentie met de toegevoegde waarde uit de nationale rekeningen bereikt. De toegevoegde waarde is beschikbaar vanaf 1969. In de periode vóór 1969 geeft de tijdreeks PI echter de volumeontwikkeling van de productie weer.
- Seizoenspatronen van de basisjaren moeten herkenbaar blijven in de tijdreeks. Dit betekent dat er een vloeiende overgang, gebaseerd op gemeten ontwikkeling van de productie, ontwikkeld moeten worden om van het ene basisjaar naar het andere te komen.
- Indien mogelijk wordt de ontwikkeling van de PI van een maand ten opzichte van dezelfde maand vorig jaar gehaald uit de cijfers vóór aanpassing aan nationale rekeningen, dus uit de cijfers die gebaseerd zijn op maandenquêtes. Ná aanpassing aan NR zit in de reeksen van de PI een stap tussen december en januari als gevolg van het uniform aanpassen van de PI in de verschillende jaren. Omdat deze stap ontstaat door de jaarlijkse aanpassing en geen gemeten effect is, wordt deze gezien als ongewenst.
- Voor 1990 is het uitgangspunt voor het maken van de tijdreeksen de gemiddelde dagreeksen omdat deze reeksen in de meeste bedrijfstakken beschikbaar zijn. Deze worden per bedrijfstak als kettingindex gekoppeld voor de verschillende basisjaren. Het maken van seizoensgecorrigeerde reeksen wordt aan de gebruiker overgelaten.
- Vanaf 1990 zijn voor alle variabelen oorspronkelijke reeksen beschikbaar (d.w.z. de reeksen zonder seizoens- en gemiddelde dagcorrectie, zie par 3.2.1). Alleen voor de aggregaten nijverheid en industrie zijn langere oorspronkelijke reeksen beschikbaar. Voor deze aggregaten worden ook langere tijdreeksen gemaakt. Het verwijderen van de gemiddelde dag correctie uit de gemiddelde dag reeksen om terug te rekenen naar oorspronkelijke reeksen is niet mogelijk omdat er verschillende methoden in de loop van de jaren gebruikt zijn en de onderliggende detailreeksen elk een ander schema voor dagcorrectie hadden. In de loop van de tijdreeks zijn er verschillende methoden gebruikt voor gemiddelde dag correctie, de verschillen hiertussen worden genegeerd.
- Voor 1990 worden er geen reeksen geaggregeerd om reeksen opnieuw op te bouwen zoals de industrie en de nijverheid uit de onderliggende bedrijfstakken of zoals een aggregaat van de kleding-, schoenen- en leerindustrie (de indeling was voor 1970 kleding+schoenen en leer en na 1970 kleding en schoenen+leer). De reden hiervoor is dat het effect van veranderende gewichten in de reeksen niet goed meegenomen kan worden door gebrek aan gegevens. Dit heeft tot gevolg dat deze data niet gepubliceerd wordt.
- De tijdreeks wordt gepubliceerd onder SBI 2008.

4.3 Gebruikte methoden

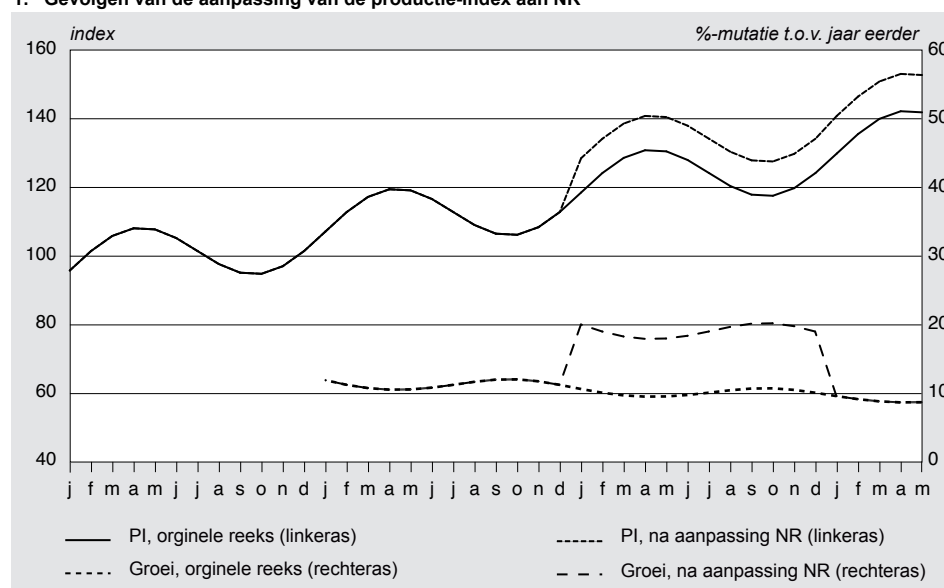
4.3.1 Berekening van de groei

Zoals beschreven in paragraaf 3.1.2 zijn er verschillende versies van de data beschikbaar. Het doel is om de meest nauwkeurige data te gebruiken, gebaseerd op microdata vóór de aanpassing aan de NR. De reden hiervoor is dat er een betere methode is voor de aanpassing van de PI aan de toegevoegde waarde van NR die het verloop van de PI minder verstoort.

Het verschil in een bepaald jaar tussen de productieontwikkeling volgens de PI en NR wordt in de huidige praktijk gelijkmatig verdeeld over de 12 maanden en bij de PI opgeteld. Hierdoor ontstaat een sprong tussen december van het vorige jaar en januari van het jaar daarna. Het is erg onwaarschijnlijk dat deze sprong ook daadwerkelijk in 1 keer van

december op januari is opgetreden. In het fictieve vereenvoudigd voorbeeld hieronder wordt dit uitgelegd. In de praktijk wordt a) de eerste drie jaar en b) met NR-revisies de PI aan volgende versies van de nationale rekeningen aangepast, maar in dit voorbeeld wordt slechts 1 aanpassing weergegeven. In de figuur zijn de originele PI-reeks en de reeks na aanpassing aan NR in het bovenste gedeelte weergegeven. De eerste 2 jaar liggen de PI reeksen over elkaar. In juni van het vierde jaar wordt door NR voor het derde jaar bepaald dat de productie hoger was waardoor de PI over het derde jaar wordt opgetild. De maanden januari t/m mei van het vierde jaar die al gemeten zijn door de PI worden meegenomen met deze productiestijging. De groei van de productie wordt in het onderste gedeelte getekend. Het eerste jaar is het niet mogelijk om de groei te berekenen, in het tweede jaar liggen de lijnen op elkaar. Als gevolg van de aanpassing aan NR springt de groei in het derde jaar in 1 maand naar een hoger niveau. Voor de maanden januari t/m mei van het vierde jaar is er nog nieuwe informatie dus die houden dezelfde, niet aangepaste, groei.

1. Gevolgen van de aanpassing van de productie-index aan NR



Omdat deze sprongen in de groei van de PI als gevolg van de aanpassingen aan NR onwaarschijnlijk zijn, is besloten om gebruik te maken van de laatste waarneming vóór de eerste aanpassing aan NR en wordt het productieverval tussen de PI en NR op een andere, meer geleidelijke, manier in de reeks verwerkt, zie paragraaf 4.3.4.

Hieruit volgt de verhouding $q_{i,j} = \frac{P_{i,j}}{P_{i,j-1}}$ van de productie $p_{i,j}$ die in de verdere berekeningen gebruikt gaat worden.

4.3.2 Productie tussen 2 basisjaren

Nu de verhouding $q_{i,j} = \frac{P_{i,j}}{P_{i,j-1}}$ op basis van waarneming (dus vóór aanpassing aan NR) in de vorige paragraaf bepaald is, wordt deze gebruikt om 2 basisjaren te koppelen. De eerste stap in de gevolgde methode is om 2 opeenvolgende basisjaren te nemen en de gemiddelde groei tussen de basisjaren te bepalen uit de maandontwikkelingen. In de uitwerking van de methode wordt aangenomen dat om de 3 jaar een nieuw basisjaar wordt samengesteld. Met deze verhouding wordt de niveauverhouding bepaald van de gemiddelde productie-index P_j tussen de basisjaren j en $j+3$. Deze stap is nodig om in latere fase de aanpassingen per maand te minimaliseren.

De gemiddelde groei tussen basisjaren is dan:

$$(4): \frac{P_{j+3}}{P_j} = \frac{\sum_{i=1}^{12} P_{i,j+3}}{\sum_{i=1}^{12} P_{i,j}} = \frac{1}{12} * \sum_{i=1}^{12} \left(\prod_{k=1}^3 q_{i,j+k} \right)$$

met $p_{i,j}$ de productie-index in maand i van jaar j , k is de index voor de jaren tussen de basisjaren. In het algemeen zal de verhouding voor de verschillende maanden afwijken van de verhouding tussen de basisjaren.

Als tweede stap worden alle verhoudingen van januari tussen de basisjaren aan elkaar geschakeld en daarna van februari enz. De zo berekende verhouding per maand zal niet gelijk zijn aan de verhouding tussen de 2 basisjaren voor die maand.

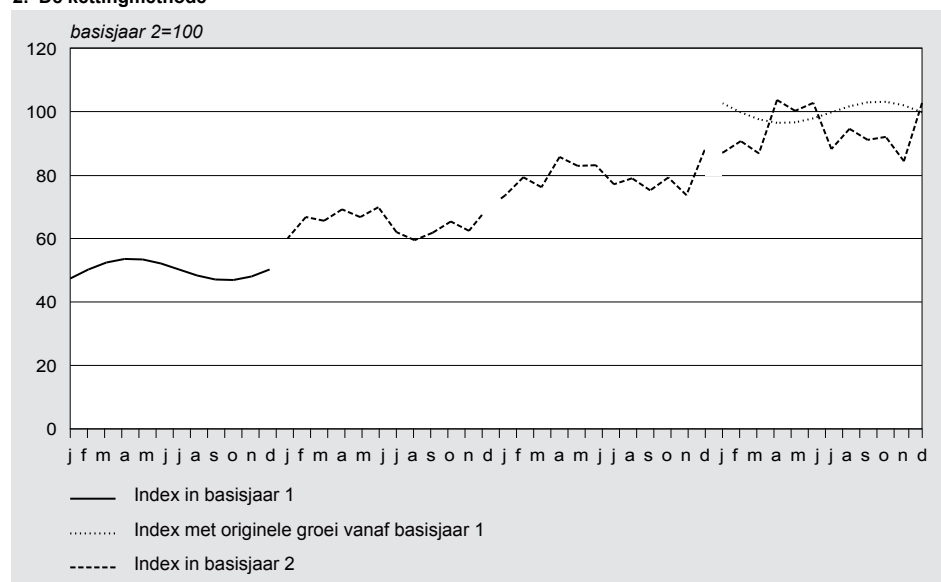
Als derde stap wordt het verschil ∂q_i gelijkmatig verdeeld over de jaren zodat elke maand met de geschakelde groei exact van het ene basisjaar in het volgende basisjaar uitkomt.

$$(5) : \frac{P_{i,j+3}}{P_{i,j}} = \prod_{k=1}^3 (q_{i,j+k} + \partial q_i)$$

Deze procedure wordt tussen alle basisjaren uitgevoerd. Het resultaat is een reeks waarbij het ene basisjaar vloeiend overloopt in het volgende met minimale aanpassing in de jaren daartussen.

Ter verduidelijking van deze procedure is de figuur hieronder gegeven. In deze figuur zijn 2 basisjaren gegeven en 2 tussenliggende jaren⁶⁾. Het eerste basisjaar is het uitgangspunt. Per maand wordt dit vermenigvuldigd met de gemiddelde groei uit de waarneming van de PI (stap 1) om zo het eerste tussenliggende jaar te krijgen. Dit eerste tussenjaar wordt weer vermenigvuldigd met de groei om het tweede tussenjaar te krijgen. Tot slot wordt het tweede tussenjaar vermenigvuldigd met de groei om het laatste jaar te krijgen (stap 2).

2. De kettingsmethode



Echter het patroon van het laatste jaar wijkt af van het gewenste rechterbasisjaar waar het op moet aansluiten. Voor elk van de 12 maanden worden de verschillen lineair verdeeld over de jaren zodat bij elke groei 1/3 van het verschil wordt opgeteld. Als de procedure nu herhaald wordt met de gecorrigeerde groei wordt wel exact aangesloten op het rechterbasisjaar (stap 3).

4.3.3 Productie vóór het eerste basisjaar en na het laatste basisjaar

Vóór het eerste basisjaar en ná het laatste basisjaar worden de reeksen gewoon geschakeld, er worden geen correcties uitgevoerd bij voorbeeld op basis van het eerst of laatst waargenomen jaar.

4.3.4 Aansluiting aan de nationale rekeningen met behulp van de Dentonmethode

Als de reeksen zijn samengesteld moeten deze worden aangesloten op de ontwikkeling van de toegevoegde waarde uit de nationale rekeningen. Dit wordt gedaan met de Dentonmethode. Deze methode past de jaargemiddelden van een maandreeks van de PI aan bij vooraf opgegeven waarden per jaar bijvoorbeeld een indexreeks van de toege-

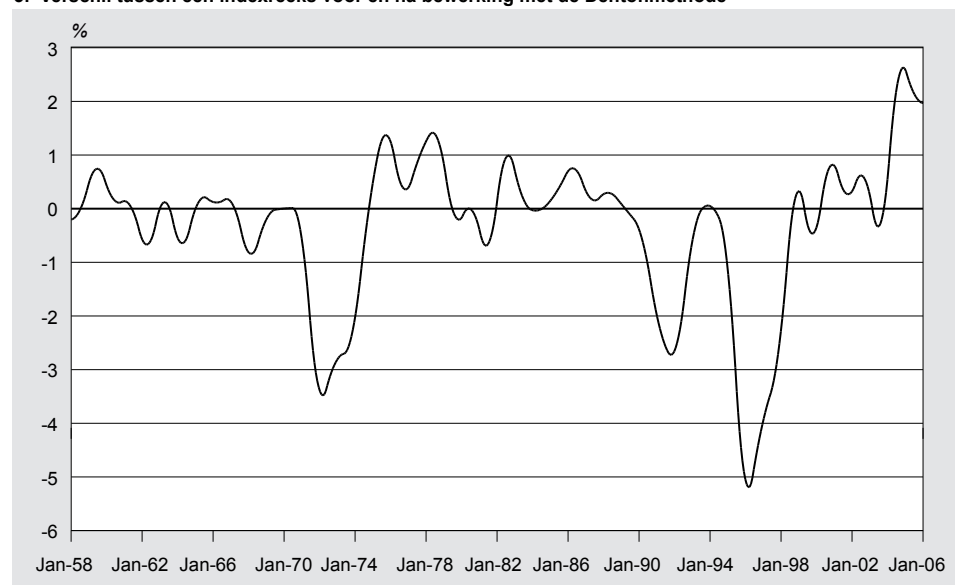
⁶⁾ In het figuur zijn als voorbeeld slechts 2 tussenjaren weergegeven, bij de echte PI zijn er in de meeste gevallen 4 tussenjaren. Elke 5 jaar is er een basisjaar.

voegde waarde volgens NR ⁷⁾. Dit gebeurt onder de aanname dat een aanpassing de ene maand niet veel mag verschillen van een naastgelegen maand ⁸⁾. Dit is met name van belang in december-januari waarbij in de ‘gebruikelijke’ methode met de aanpassingen voor 2 verschillende jaren gewerkt wordt, hetgeen tot vreemde ontwikkelingen kon leiden. Denton heeft in 1971 zijn methode gepubliceerd om maand of kwartaaldata aan te passen aan jaartotalen (Denton (1971)). De gebruikte implementatie van de Dentonmethode is uitgebreid beschreven in de publicatie Bikker e.a..(2007). Deze methode heeft de volgende voordelen:

- De december–januari stap is weg die wel ontstond bij de ‘gebruikelijke’ manier van aanpassen aan de NR.
- De originele seizoenspatronen van de basisjaren blijven zoveel mogelijk behouden. En daarnaast is er een geleidelijke overgang van het seizoenspatroon tussen de verschillende basisjaren.

Nadeel van de Dentonmethode is dat het een complexe wiskundige programmatuur vereist.

3. Verschil tussen een indexreeks voor en na bewerking met de Dentonmethode



Figuur 3 geeft het verschil voor een voorbeeldreeks voor en na bewerking met de Dentonmethode. Het is duidelijk dat de aanpassing geleidelijk verloopt in plaats van stapsgewijs. De gemiddelde grootte van de aanpassingen over een jaar moeten gelijk zijn aan de gebruikelijke methode met een uniforme stap per jaar. De Dentonmethode wordt op 2 manieren geïmplementeerd, namelijk voor een jaar dat bestaat uit 12 maanden en voor een jaar dat bestaat uit 9 perioden (8 maanden en de periode juni, juli, augustus en september). Deze laatste variant is nodig om te voorkomen dat de verhouding tussen maanden verandert, waardoor het niet mogelijk is om vakantiespreiding te detecteren.

4.4 Implementatie voor 1990

Zoals in paragraaf 3.2.1 vermeld zijn er voor 1990 zijn er alleen 2 digit reeksen beschikbaar. Met de methode uit paragraaf 4.3.1 worden de groei percentages bepaald die volgens de methode uit 4.3.2 gekoppeld worden tot een indexreeks. In het begin wordt de methode uit sectie 4.3.3 gebruikt. Deze reeksen worden niet verder geaggregeerd en

⁷⁾ In twee gevallen voor de aardolie-industrie geeft de NR een negatieve waarde in lopende prijzen voor de toegevoegde waarde in combinatie met een positieve waarde in constante prijzen. Hierdoor is het niet mogelijk om een goede indexreeks te maken. Dit wordt opgelost door een lopend gemiddelde te gebruiken bestaande uit $(1/12 \cdot \text{jaar } j-1) + (5/6 \cdot \text{jaar } j) + (1/12 \cdot \text{jaar } j+1)$.

⁸⁾ Dit betekent dat er ook in het basisjaar kleine aanpassingen gemaakt kunnen worden.

hoeven niet geconverteerd te worden voor SBI-wijzigingen omdat deze op het 2 digit niveau van slechts zeer beperkte invloed zijn.

4.5 Implementatie na 1990

In tegenstelling tot voor 1990, is het basismateriaal na 1990 4 digit oorspronkelijke reeksen. Met dezelfde methode als in sectie 4.4 worden er reeksen gemaakt. Nu zijn er extra bewerkingen nodig omdat deze gedetailleerde reeksen 1) te aggregeren en 2) om te rekenen van SBI 93 naar SBI 2008 worden. Stapsgewijs ziet de berekening er als volgt uit:

- In de eerste stap worden reeksen van 1990 t/m 2007 gemaakt op 4 digit, SBI 93 niveau⁹⁾. Om de reeksen onderling vergelijkbaar te maken wordt met behulp van Denton methode reeksen op niveau van de toegevoegde waarde uit de nationale rekeningen gebracht (Bikker (2007)). Hierdoor is de ontwikkeling van de reeksen onderling vergelijkbaar en consistent met de nationale rekeningen. Het resultaat is $a_k p_{i,j,k}$ uit vergelijking 3, dit zijn 4 digit index reeksen p onder SBI 93 met basisjaar 2005=100, gewicht a op basis van de toegevoegde waarde en k de index voor de reeksen onder SBI 93.
- De volgende stap rekent de SBI 93 gegevens met index k van de maand-reeksen uit de vorige stap en de toegevoegde waarde (per jaar) uit de nationale rekeningen om naar SBI 2008 met index l op 4 digit niveau. Hiervoor wordt het schakelschema van het algemeen bedrijven register (ABR) van december 2008 gebruikt dat aangeeft het percentage werkzame personen dat van een SBI 93 naar een SBI 2008 gaat op 4 digit niveau. (dus $\sum_l A_{k,l} = 100\%$). Dit levert

$$(6): p_{i,j,l} = \frac{a_l p_{i,j,l}}{a_l} = \frac{\sum_k A_{k,l} a_k p_{i,j,k}}{\sum_k A_{k,l} a_k}$$

- Aan de reeksen van de productie-index wordt het jaar 2008 toegevoegd. Dit jaar is al vanuit microdata op basis van SBI 2008 bepaald.
- De 4 digit reeksen worden gewogen en vervolgens geaggregeerd naar 2 digit niveau en hoger met vergelijking 3.

4.6 Koppelen van de reeksen van voor 1990 en na 1990

De geaggregeerde reeksen vanaf 1990 worden gekoppeld aan de reeksen van voor 1990. Er wordt geen rekening gehouden met dat de meeste reeksen voor 1990 gemiddelde dag zijn en na 1990 oorspronkelijk. Tijdens de werkdagcorrectie wordt dit gecorrigeerd. Er zijn nu 2 bewerkingen nodig, namelijk aanpassen aan de toegevoegde waarde uit de NR en de werkdagcorrectie uit de seizoenscorrectie.

- De tijdreeks van de PI heeft nog niet de gewenste jaar op jaar ontwikkeling. In het kader van de 1 cijfer gedachte wordt de PI, indien mogelijk, aangepast aan de toegevoegde waarde uit de NR. Deze is beschikbaar vanaf 1969, in de periode daarvoor wordt de meest recente jaarcijfers van de productie index gebruikt. Om de gewenste jaargroei te krijgen is een extra Denton-slag nodig (Bikker 2007)¹⁰⁾.
- De laatste bewerking is het verwijderen van residueel werkdagenpatroon met X-12. Er wordt hierbij gebruik gemaakt van de mogelijkheid van X-12 om 2 verschillen perioden van werkdagcorrectie te hebben.

5. Referenties

- Problemen bij het samenstellen productie-indexcijfers; In: Statistische en econometrische onderzoeken, nr 5, p 15–39, CBS, 1966

⁹⁾ 2008 is al samengesteld onder SBI 2008 en wordt daarom later in stap 2 toegevoegd.

¹⁰⁾ Omdat er een verschil is tussen het aggregeren van gewogen indexreeksen van de toegevoegde waarde (gevolgde methode) en het eerst aggregeren van de toegevoegde waarde en dan er een indexreeks van maken (gewenst resultaat), voldoen de aggregaten van de indexreeksen van de productie-index niet meer aan de jaar-op-jaar ontwikkeling van de toegevoegde waarde. Hierdoor is de Dentonmethode ook na 1990 nog een keer nodig.

- Hoeveelheidsindexcijfers van de industriële productie (1963=100); In: Maandstatistiek van de industrie, jrg 10(1968), nr 1, p 73–85
- Hoeveelheidsindexcijfers van de industriële productie 1970=100; In: Maandstatistiek van de industrie, jrg 16(1974), nr 10, p 729–742
- Hoeveelheidsindexcijfers van de productie in de nijverheid 1975=100; In: Maandstatistiek van de industrie, jrg 22(1980), nr 10, p 729–742
- Hoeveelheidsindexcijfers van de productie in de nijverheid 1980=100; In: Maandstatistiek van de industrie, jrg 26(1984), nr 8, p 12–32
- Hoeveelheidsindexcijfers van de productie in de nijverheid 1985=100; In: Maandstatistiek van de industrie, jrg 32(1990), nr 1, p 7–20
- Bikker, Reinier, Jacco Daalmans en Nino Mushkudiani; Methodenreeks: Thema: macro-integratie, deelthema: inpassen; CBS(2007) 06-03-2008. (<http://www.cbs.nl/>)
- Denton, Frank. T.; Adjustment of monthly or quarterly series to annual totals: An approach based on quadratic minimization; Journal of the American statistical association, vol 66, no 333 (mar. 1971), p 99–102

6. Bijlage 1: transitie matrix van SBI 93 naar SBI 2008

Hiernaast is de transitie matrix van SBI 93 naar SBI 2008 weergegeven voor de industrie. De verdeling van de verschillende bedrijfstakken is gebaseerd op werkzame personen. Vertikaal staan mogelijke detailreeksen onder SBI 2008. Horizontaal staan eerst de aggregaatreeksen, vervolgens gepubliceerde detailreeksen en tot slot indien nodig een uitsplitsing om te kijken of het mogelijk is om op een 1 op 1 overgang te komen.

Som van WP	sbi93 pub		sbi93 pub2		sbi93 pub3														Eind totaal												
	11-14	15-16	11-14	15-16	11-14	15-16	17	18	19	20+36-37	21	22	23	24	25	26	27-35	28	29-33	34+35	40+41 #N/B	40+41 #N/B	40+41 #N/B								
sbi08 pub	schakel						16	151+155		20	21	22	23	23	23	24	25	26	27	28	29-33	34+35									
06-09	1-1	7342				16	151	155	37	20	21	22	23	23	24	25	26	27	28	29-33	3511	3520	3530	34+35+355							
101+105	1-1						13317	11248																8	7350						
102-104+106-109	1-1	75					4279																		24565						
11	1-1																								73317						
12	1-1						6616																		6616						
13	1-1						3971																		3971						
14	1-1																								12525						
15	1-1																								3891						
16+31-32	1-1																								1777						
17	1-1									16889															124934						
18	1-1									99323															19830						
19	1-1										19830														35073						
20-21	1-1																								5713						
22	1-1																								62182						
23	1-1																								31656						
24	1-1																								28652						
25	1-1																								21812						
26-28+33	1-1																								91968						
29-30	1-1																								127						
35-36	1-1																								37387						
#N/B	1-1																								22838						
Eind totaal		7417	68963	6616	3971	17596	11248	12697	3735	1761	16889	19830	33614	35073	5713	596	61711	31656	28652	21812	92211	148612	8139	5363	3549	5138	24316	22838	6237892	6276809	7042358