



Centraal Bureau voor de Statistiek

Divisie Macro-economische statistieken en publicaties
Sector Ontwikkeling en ondersteuning

*Postbus 4000
2270 JM Voorburg*

Tijdelijke inpasmethode werkzame beroepsbevolking

Ria Okkerse-Ruitenbergh, George van Leeuwen

Kennisgeving:

De in dit rapport weergegeven opvattingen zijn die van de auteurs en komen niet noodzakelijk overeen met het beleid van het Centraal Bureau voor de Statistiek.

Projectnummer:

102667-Euro

BPA-nummer:

2005-22-MOO

Datum:

23 februari 2005

TIJDELIJKE INPASMETHODE WERKZAME BEROEPSBEVOLKING

Samenvatting: Dit document beschrijft de verbeterde inpasmethode om het cijfer voor de werkzame beroepsbevolking met als belangrijkste bron de enquête beroepsbevolking, in te passen in de arbeidsrekeningen. Deze methode zal toegepast worden vanaf het eerste kwartaal 2004 tot medio 2005.

Trefwoorden: werkzame beroepsbevolking, enquête beroepsbevolking, arbeidsrekeningen.

1. Inleiding

Arbeidsrekeningen (AR) is een integratiekader waarbinnen verschillende bronnen met elkaar geconfronteerd worden. Uit de AR wordt op kwartaalbasis het aantal banen van werknemers gepubliceerd. Sinds februari 2003 wordt maandelijks het geïntegreerde cijfer over de werkzame beroepsbevolking gepubliceerd. De mutatie van het aantal banen van werknemers heeft als belangrijkste bron de enquête werkgelegenheid en lonen (EWL). De werkzame beroepsbevolking op maandbasis heeft als belangrijkste bron de enquête beroepsbevolking (EBB). De EBB levert ook de onderliggende gegevens om de verschillen tussen het begrip ‘werkzame beroepsbevolking’ en ‘banen van werknemers’ te kwantificeren. Op basis van dit bronmateriaal uit de EBB kan dus ook een cijfer berekend worden voor banen van werknemers. Dit cijfer wordt niet gepubliceerd.

Het aantal banen van werknemers berekend op basis van de EBB en uit de AR zijn in de afgelopen jaren uit elkaar gaan lopen. Dit terwijl de begrippen van ‘banen van werknemers’ op elkaar aansluiten, met uitzondering van in- en uitgaande pendel¹. In- en uitgaande pendel vormt maar een klein aandeel van de banen van werknemers en vormt niet de oorzaak van de waargenomen verschillen. De herkomst van de gegevens van EBB en EWL zijn echter niet dezelfde. De EBB is een huishoudensenquête, en de EWL een bedrijvenenquête. Besloten is de inpasmethode van de EBB op de AR te wijzigen.

2. Inpasmethode

De inpasmethode begrenst het verschil tussen de jaarontwikkeling banen werknemers met als bron de EBB en volgens AR. Deze methode wordt toegepast vanaf het eerste kwartaal 2004. Tot dat moment wordt uitgegaan van de reeds

¹ Werknemers die in Nederland werken en in het buitenland wonen en omgekeerd.

gepubliceerde cijfers. De methode zal toegepast worden tot de revisie van de Nationale Rekeningen die medio 2005 voltooid is.

2.1 Van werkzame beroepsbevolking naar banen van werknemers

De inpassmethode gaat uit van het inpassen op basis van het begrip: banen van werknemers. Na het inpassen van 'banen van werknemers' worden op basis van EBB gegevens de groepen die het begrip 'banen van werknemers' scheiden van het begrip 'werkzame beroepsbevolking' weer toegevoegd.

Allereerst wordt daarom uit de EBB brongegevens het cijfer voor banen van werknemers bepaald. Het verschil tussen het begrip 'AR banen van werknemers', 'EBB banen van werknemers' en 'werkzame beroepsbevolking' wordt in onderstaand schema gegeven.

$$\begin{aligned} &AR \text{ banen, werknemers} = \\ &EBB \text{ banen, werknemers} + \text{binnenkomende pendel} - \text{uitgaande pendel} = \\ &EBB \text{ werkzame beroepsbevolking} \\ &\quad - \text{werkzame beroepsbevolking, zelfstandigen} \\ &\quad + \text{personen werkzaam} < 12 \text{ uur, 15-64 jaar, werknemer} \\ &\quad + \text{bijbanen 15-64 jaar als werknemer} \\ &\quad + \text{hoofd- en bijbanen 65 jaar en ouder, werknemers} \\ &\quad + \text{binnenkomende pendel} - \text{uitgaande pendel..} \end{aligned}$$

De samenhang tussen deze cijfers is uitvoerig beschreven door de Vriesⁱ.

2.2 Schatting AR-cijfer

Het geïntegreerde AR-kwartaalcijfer voor de banen van werknemers komt één maand later beschikbaar dan de EBB over dezelfde periode. Om de EBB toch te kunnen inpassen moeten we een voorspelling maken voor het AR-cijfer. Dit doen we met behulp van een zogenaamd State-Space Model (SSM). In deze aanpak wordt het niveau van de werkgelegenheid gezien als een toestand die is opgebouwd uit verschillende componenten ('States') en waarover in de loop der tijd via metingen inzicht wordt verkregen (zie bijvoorbeeld Durbin en Koopmanⁱⁱ). Met behulp van de methodologie van Kalman filtering wordt informatie over het (relatieve) belang van de onderliggende componenten van de tijdreeks in een iteratief proces opgebouwd. Door steeds meer informatie te gebruiken (de ruimte in de tijdsdimensie te vergroten), krijgen we een steeds beter beeld van de meest recente 'States'.

Het SSM voor de reeks banen werknemers in de AR bestaat uit twee (stelsels van) vergelijkingen: a) één vergelijking die de 'States' relateert aan de meting en b) een set van 1 of meer vergelijkingen die het gedrag van de 'States' in de tijd beschrijft. Dit leidt tot de volgende (stelsels van) vergelijkingen:

$$y_t = Z\alpha_t + G\varepsilon_t$$

$$\alpha_t = J\alpha_{t-1} + H\varepsilon_t,$$

met y_t de vector van kwartaalwaarnemingen voor de banen van werknemers volgens de AR voor kwartaal $t \{t = 1, \dots, T\}$ en α_t de zogenaamde ‘State’ vector voor hetzelfde kwartaal. De matrices Z, J, G en H zijn zogenaamde systeemmatrices. Hiervan dienen Z en J respectievelijk voor a) het beschrijven van de ‘link’ tussen de ‘State’ variabelen en de metingen en b) voor het beschrijven van de transitie van de ‘States’

Voor de ‘State’ variabelen volgen we een veel gebruikte variant van het zogenaamde ‘unobserved-components’ model (UCM). Deze variant beziet een tijdreeks als zijnde opgebouwd uit drie componenten: een trend-cycle component met een vast lokaal niveau en een stochastische helling, een seizoencomponent en een onregelmatige component (statistische ruis).

De specificatie van de meetvergelijking voor de AR-cijfers in termen van ‘State’ variabelen luidt dus als volgt:

$$y_t = Trend_t + S_t + \varepsilon_t,$$

met $Trend_t$ het niveau van de trend-cycle, S_t de seizoencomponent in het AR-cijfer en ε_t een onafhankelijk en normaal verdeelde storing van de meetvergelijking $\varepsilon_t \sim NID(0, \sigma_\varepsilon^2)$.

De volgende stap is het specificeren van de modellen voor de componenten. Omdat de AR-reeks een tamelijk glad trendmatig verloop kent, lijkt voor de trendcomponent een model met een ‘smoothed’ trend het meest geëigend. Volgens dit model volgt de trend een lokaal niveau met stochastische drift. De specificatie voor de trendcomponent is dus:

$$Trend_t = Trend_{t-1} + \beta_{t-1}$$

$$\beta_t = \beta_{t-1} + \varepsilon_\beta.$$

Het niveau van de trend wordt dus gelijk gesteld aan het niveau van de trend in het voorafgaande kwartaal vermeerderd met een stochastische helling (de verandering van de trend) β . De stochastiek in de trendspecificatie wordt vertegenwoordigd door ε_β , een onafhankelijk en normaal verdeelde storingsterm: $\varepsilon_\beta \sim NID(0, \sigma_\beta^2)$.

Omdat we ervan uit mogen gaan dat er ook seizoenpatronen in het AR-cijfer voorkomen, nemen we tevens een seizoencomponent in het model op. We veronderstellen dat het ‘seizoen’ niet deterministisch is, maar stochastisch. Met andere woorden: de som van de seizoencomponenten binnen een jaar heeft verwachting nul. Dit leidt tot de volgende vergelijking voor de seizoencomponent:

$$S_t + S_{t-1} + S_{t-2} + S_{t-3} = \varepsilon_S,$$

met ε_S een onafhankelijk en normaal verdeelde storing: $\varepsilon_S \sim NID(0, \sigma_S^2)$.

2.3 Bandbreedte tussen de jaarontwikkeling EBB en AR

Vervolgens wordt gekeken of het verschil in jaarontwikkeling tussen het berekende EBB cijfer en het AR cijfer voor banen van werknemers binnen een bepaalde bandbreedte valt. Als de jaarmutatie buiten de bandbreedte valt, wordt het EBB cijfer aangepast tot op de grens van de bandbreedte. De bandbreedte komt voort uit een analyse van de betrouwbaarheidsintervallen van de individuele reeksen.

Nadat het eerste voorlopige EBB-cijfer voor de werkzame beroepsbevolking gepubliceerd wordt, komt er vanuit arbeidsrekeningen eerst een flash en later het reguliere cijfer voor banen van werknemers. Op basis van de steeds betrouwbaardere informatie kan het EBB-cijfer voor de werkzame beroepsbevolking aangepast worden. Er wordt een bandbreedte van 0,8% gebruikt bij een voorspelling, 0,8% bij een flash en 0,4% bij een regulier AR cijfer.

2.4 Interpolatie van de mutatie

Omdat het AR cijfer per kwartaal beschikbaar is, en het EBB cijfer maandelijks, moet de mutatie van het EBB cijfer vervolgens geïnterpoleerd worden. De methode hiervoor is dezelfde als die bij de introductie van het inpassen van de maandcijfers EBB is gebruiktⁱⁱⁱ. Deze methode komt erop neer dat als de mutatie voor het kwartaal jan-mrt 30 duizend was en er een kwartaal eerder geen aanpassingen waren, het driemaandsgemiddelde (dec-feb) met 20 duizend wordt aangepast en het driemaandsgemiddelde (nov-jan) met 10 duizend.

2.5 Extrapolatie van de mutatie

Voor de eerste twee driemaandsgemiddelde (EBB) die volgen na een kwartaal dat is ingepast, is een iets afwijkende inpasmethode nodig. Zolang een kwartaal nog niet compleet is, geeft de hierboven beschreven methode nog onvoldoende handvatten om de driemaandsgemiddelden in te passen. Als het cijfer zonder meer gepubliceerd zou worden, geeft dit een sprong in het cijfer die pas als het kwartaal compleet is, gecorrigeerd zal worden.

Tabel 1 geeft een fictief voorbeeld van de werkwijze bij het inpassen van eerste twee driemaandsgemiddelden die volgen na een ingepast kwartaal. Allereerst wordt er een voorspelling gemaakt van de jaarontwikkeling van het AR cijfer voor het eerstvolgende complete kwartaal. Dit gebeurt op basis van het SSM zoals boven beschreven. Hiermee wordt een jaarontwikkeling van banen van werknemers AR berekend. Deze jaarontwikkeling wordt geïnterpoleerd voor de tussenliggende maanden. De EBB driemaandsgemiddelden worden hiermee vergeleken en hierop eventueel tot op een betrouwbaarheidsinterval van 0,8% aangepast.

Tabel 1. Voorbeeld van het inpassen van het driemaandsgemiddelde nov-jan.

	EBB	AR- kwartaalcijfer	Interpolatie AR	EBB- ingepast*
4 ^e kwartaal 2004	0%	-1,5% (flash)		-0,7%
Nov-Jan Dec-Feb	0%		-1,6% -1,7%	-0,8%
1 ^e kwartaal 2005		-1.8% (voorspelling)		

* Het cijfer wordt aangepast tot op 0,8% van de flash of het geïnterpoleerde cijfer.

3. Resultaten

De methode voor inpassen zoals hierboven beschreven is toegepast vanaf het eerste kwartaal 2004. In de periode voor 2004 gelden de reeds gepubliceerde cijfers. Tabel 2 geeft een overzicht van de inpassingscorrecties die gedaan zijn op basis van de in dit document beschreven methode. Het vrouwelijke deel van de werkzame beroepsbevolking is in 2004 niet is aangepast, met uitzondering van het laatste kwartaal. De werkzame manlijke beroepsbevolking is met enkele tienduizenden aangepast.

Tabel 2. Mutatie van de werkzame beroepsbevolking, duizend personen.

	Oorspronkelijk	Inpassingscorrectie	Ingepast
Vrouwen			
1e kw 2004	2910	0	2910
2e kw 2004	2950	0	2950
3e kw 2004	2997	0	2997
4e kw 2004	2959	-20	2939
Mannen			
1e kw 2004	4108	-54	4054
2e kw 2004	4115	-36	4079
3e kw 2004	4167	-40	4127
4e kw 2004	4106	-22	4084

ⁱ S. Vries, 2003, Relatie tussen banen, werkzame personen en werkzame beroeps-bevolking, SET 2003-4, CBS

ⁱⁱ J. Durbin en S. Koopman, Time series analysis by State Space methods, Oxford University Press 2001

ⁱⁱⁱ R. Milot, 2003, Maandelijkse geïntegreerde statistiek over de werkzame en werkloze beroepsbevolking: gebruikte methode, SET 2003-7, CBS.