



SRG E: Transactieconcept SRG: bepaling van de optimale vertragingsvoet

Frank Pijpers

Lex Dekkers

Thomas Slager

samenvatting *In het kader van vernieuwing van de Statistiek Re-integratie door Gemeenten (SRG), waarbij de halfjaarlijkse uitvraag is veranderd naar een meer geautomatiseerde afslag op maandbasis van de gemeentelijke registraties, is nagegaan of het mogelijk is om binnen de SRG administratief vertragde informatie terug te leggen met dezelfde vertragingsvoet die binnen de Bijstandsuitkeringen statistiek (BUS) gehanteerd wordt.*

Gebleken is dat er nog onvoldoende data beschikbaar is om definitief een betrouwbare bepaling te kunnen doen van de optimale vertragingsvoet voor de SRG om te komen tot een SRG statistiek op transactiebasis. Voor de meest recente maanden kan nog geen definitief transactiecijfer worden samengesteld omdat de noodzakelijke registratiegegevens nog niet bestaan. Ook bestaat er een grote mate van administratieve vertraging; de totale aantallen mutaties die op de registratie bestanden moeten worden toegepast om te komen tot een transactiebestand zijn nog zeer groot.

De belangrijkste aanbeveling op dit moment is om relevante data te blijven verzamelen tot er genoeg beschikbaar is om maandtransactiecijfers samen te kunnen stellen voor diverse keuzes van vertragingsvoet. Daarnaast is het van belang om te monitoren of de grote administratief vertragde stromen afnemen naarmate de gemeenten hun maandregistraties beter in hebben kunnen regelen, en in samenwerking met gemeenten te zoeken naar oorzaken en maatregelen wanneer er geen verbetering plaatsvindt.

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

projectnummer 300821-3

15 decemberr 2015

kennisgeving De in dit rapport weergegeven opvattingen zijn die van de auteur(s) en komen niet noodzakelijk overeen met het beleid van het Centraal Bureau voor de Statistiek.

Inhoudsopgave

1. Achtergrond	4
2. Methode	5
3. Empirische bepaling van parameters	6
4. Conclusies en aanbevelingen	10

1. Achtergrond

In het kader van vernieuwing van de Statistiek Re-integratie door Gemeenten (SRG) is de halfjaarlijkse uitvraag veranderd naar een meer geautomatiseerde afslag op maandbasis van de gemeentelijke registraties, beginnend in januari 2015. Door deze verandering is de informatiestroom meer gaan lijken op die van de Bijstandsuitkeringen statistiek (BUS). Om te voldoen aan interne eisen rondom kwaliteit, reproduceerbaarheid, en uniformiteit van statistieken binnen hetzelfde domein, zowel als om tegemoet te komen aan de voornaamste gebruiker van deze statistiek (ministerie SZW) is het wenselijk om nauw aan te sluiten bij de verwerkingsprocessen rond de BUS. Eén van de verwerkingsprocessen is het terugleggen van administratief vertraagde informatie die in de volgende maanden wordt geleverd in de gemeentelijke bestanden. Dit resulteert in zogenaamde 'transactiebestanden' waarin voor de BUS informatie is teruggelegd tot drie maanden. Om ook actuele cijfers te kunnen leveren, worden voor de drie laatste maanden transactiecijfers geraamd. De notitie "Een nieuw publicatieconcept voor maandcijfers van bijstands-uitkeringenstatistiek" (dd. 07-02-2012) geeft de achtergronden van deze werkwijze met transactiecijfers.

In een aparte notitie "De verdragingsvoet in het nieuwe publicatieconcept van bijstandsuitkeringenstatistieken (dd. 16-07-2013)" wordt beschreven hoe is gekomen tot de keuze van een verdragingsvoet van 2 maanden voor de BUS, die in een later stadium verhoogd is naar 3 maanden na wetwijzigingen die een dergelijke langere verdragingsvoet rechtvaardigen. Ook voor de SRG zal worden nagegaan of het invoeren van een dergelijk transactieprincipe zinvol is en wat in dat het aantal maanden is waarin vertraagde informatie geleverd wordt (we noemen dit de verdragingsvoet). Hoewel het wellicht wenselijk is om met eenzelfde verdragingsvoet ook binnen de SRG administratief vertraagde informatie terug te leggen, is het belangrijk om na te gaan of de data zelf dat ook toelaten.

Deze analyse volgt in grote lijnen dezelfde aanpak die ook is gebruikt om de optimale BUS verdragingsvoet vast te stellen. Echter, op het moment van uitvoeren van deze analyse is er slechts voor 6 maanden (januari t/m juni 2015) registratie data beschikbaar.

2. Methode

Er is een modelmatige aanpak gekozen, om te schatten welke fractie van corrigerende informatie na een bepaalde vertraging is verwerkt in de bestanden, en dus ook welke fractie er nog steeds niet is verwerkt. Overigens zullen niet alle correcties aanleiding geven tot veranderingen in tellingen die worden gedaan op basis van die bestanden. Voor de standaard Statline publicaties zijn eigenlijk alleen die correcties van belang, die wél aanleiding geven tot veranderingen in tellingen. De voorgestelde methode is hiervoor onverminderd van toepassing, en is ook gebruikt voor het bepalen van de optimale vertragingsvoet voor de BUS.

Om de schatting te doen moeten bepaalde aannames worden gemaakt over het proces van administratieve aanvullingen en correcties op bestanden met SRG voorzieningen gegevens. De aannames kunnen kort worden beschreven als volgt:

- Er is een eindig aantal aanvullingen of correcties dat nodig is voor een bepaald maandbestand van SRG voorzieningen gegevens.
- Per gegeven tijdsinterval wordt er een aantal correcties aangebracht dat een constante fractie is van de nog ongecorrigeerde verkeerde gegevens in de bestanden.

De combinatie van deze twee aannames leidt tot de beschrijving van de nog resterende correcties na een vertragingsvoet van t als een Poisson proces. Het tijdsverloop van elke willekeurige telling uit de bestanden, zoals bijvoorbeeld het aantal records is dan :

$$R_t = R_\infty - C e^{-\frac{t}{\tau}}$$

Hierbij is R de daadwerkelijke telling op basis van het bestand met vertragingsvoet t , en R_∞ de telling zoals die zou worden na een oneindig lange tijd wachten. C is het aantal correcties dat op $t=0$ nog moet worden toegepast, en τ is een karakteristiek tijdsinterval tussen correcties.

De parameters R_∞ , C en τ zijn vooralsnog onbekend en moeten worden bepaald op basis van bestandsgegevens. De belangrijkste hiervan voor het bepalen van de optimale vertragingsvoet is de parameter τ .

3. Empirische bepaling van parameters

De bepaling van de optimale verdragingsvoet hangt nauw samen met de karakteristieke tijd τ . Vanuit het oogpunt van tijdigheid van de statistiek is het wenselijk om een zo kort mogelijke verdragingsvoet te hanteren. Vanuit het oogpunt van nauwkeurigheid is het wenselijk om lang te wachten om zo veel mogelijk correcties te hebben toegepast. De balans hiertussen is om een tijd τ te gebruiken als verdragingsvoet. Een kortere verdragingsvoet betekent dat er onacceptabel veel correcties nog niet zijn ingebracht, en een langere verdragingsvoet is niet voldoende lonend in aantallen correcties om de publicatievertraging te rechtvaardigen.

Anders dan bij de BUS, zijn er voor de SRG als maandstatistiek nog slechts een beperkt aantal maanden beschikbaar, op basis waarvan de parameter τ bepaald zou kunnen worden. Bovendien is het niet onwaarschijnlijk dat in deze maanden het maandproces nog onvoldoende goed is ingeregeld bij gemeenten en er dus een groter aantal wijzigingen en correcties wordt doorgevoerd dan over één of enkele jaren het geval zal zijn. Dat kan ook een invloed hebben op de schatting voor τ en daarmee de schatting voor de optimale verdragingsvoet.

In tabel 1 zijn voor drie verslagperioden (januari t/m maart 2015) de transactiecijfers bepaald, met teruglegging van informatie uit registratie bestanden van 1, 2, of 3 maanden later. Voor februari respectievelijk januari is er zelfs informatie teruggelegd uit registratie bestanden van 4, of 5 maanden later.

Tabel 1: SRG-gegevens (aantallen voorzieningen) waarin verstraagde informatie is verwerkt met een verdragingsvoet van 1 t/m 5 maanden

Transactiecijfers	#Teruglegmaanden				
	1	2	3	4	5
Lopende voorzieningen ultimo verslagperiode					
201501	207369	208084	208719	208649	208697
201502	207883	209076	209236	209374	
201503	210750	211315	211642		

Gezien de slechts heel beperkte hoeveelheid informatie kan er op dit moment nog niet definitief vastgesteld worden wat de optimale verdragingsvoet is. Daarvoor is het nodig om voor een heel jaar transactiecijfers samengesteld te kunnen hebben.

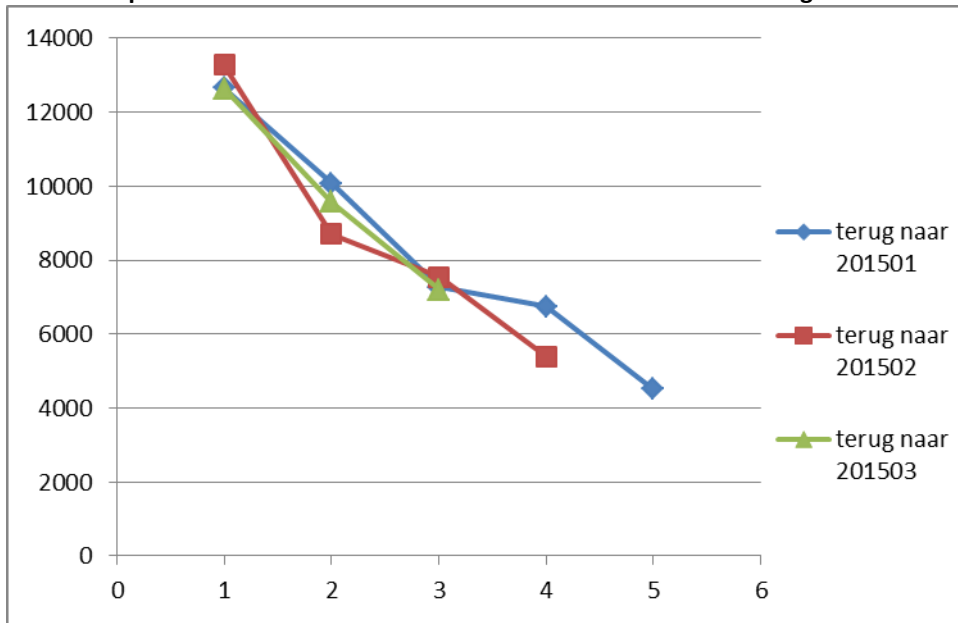
In tabel 2 wordt aangegeven wat de administratieve instroom en uitstroom is voor ieder van de 3 transactie maanden, januari t/m maart 2015. Het gaat hier dus om voorzieningen waarvoor in opvolgende maanden (maximaal 5 maanden later) gegevens aangepast zijn in de registratie van de gemeenten, zodanig dat de betreffende voorziening in de transactiemaand niet geteld is maar wel geteld had moeten worden (instroom), of wel geteld is, maar niet geteld had moeten zijn (uitstroom). Het eindtotaal is het totaal aantal mutaties dat moet worden uitgevoerd om tot een transactiebestand te komen voor een bepaalde maand, op basis van de gegevens van de 'donor'-maand die 1, 2, of meer maanden volgend is op de maand waarvoor het transactiebestand wordt samengesteld.

Tabel 2: SRG-gegevens (aantallen voorzieningen) uitgesplitst naar vertraagde in- respectievelijk uitstroom vanuit opvolgende maanden, voor ieder van de 3 transactiem maanden (januari t/m maart 2015). Ook aangegeven is het eindtotaal aan mutaties, en het netto effect op de standcijfers (verschil instroom-uitstroom).

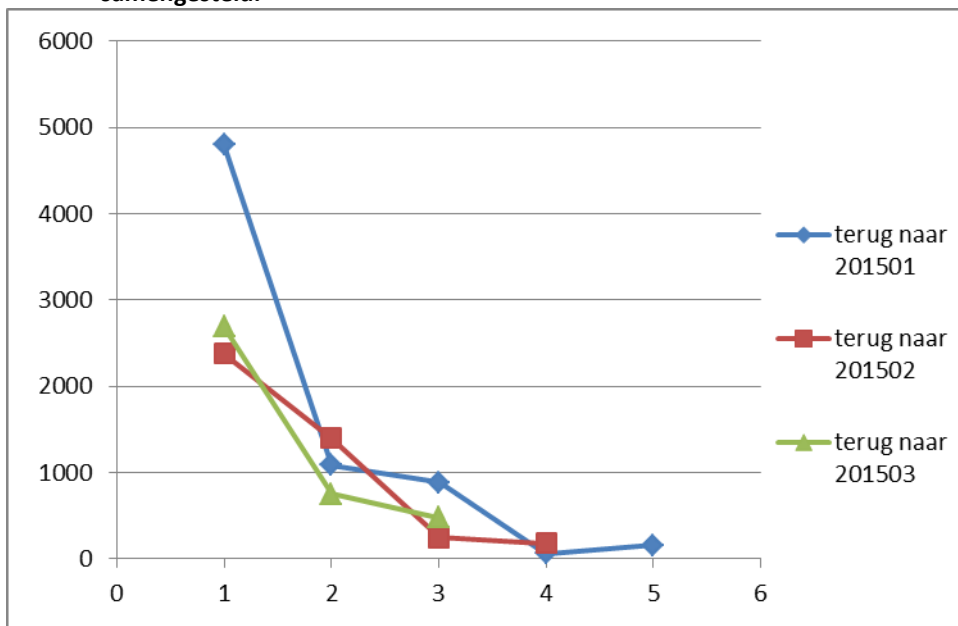
Administratief teruggelegde in- en uitstroom 201501						
	201502	201503	201504	201505	201506	Eindtotaal
Administratief vertraagde instroom teruggelegd	8725	5576	4078	3408	2329	24116
Van voorziening einddatum voorziening aangepast	3928	4490	3193	3346	2173	17130
Eindtotaal	12653	10066	7271	6754	4502	41246
Verschil	4797	1086	885	62	156	6986
Administratief teruggelegde in- en uitstroom 201502						
Rijlabels	201503	201504	201505	201506	Eindtotaal	
Administratief vertraagde instroom teruggelegd	7827	5053	3892	2789	19561	
Van voorziening einddatum voorziening aangepast	5445	3653	3646	2606	15350	
Eindtotaal	13272	8706	7538	5395	34911	
Verschil	0	2382	1400	246	183	4211
Administratief teruggelegde in- en uitstroom 201503						
Rijlabels	201504	201505	201506	Eindtotaal		
Administratief vertraagde instroom teruggelegd	7658	5166	3834	16658		
Van voorziening einddatum voorziening aangepast	4971	4412	3356	12739		
Eindtotaal	12629	9578	7190	29397		
Verschil	0	0	2687	754	478	3919

In figuur 1 worden de eindtotalen in een grafiek weergegeven als functie van de tijdspanne (in maanden) tussen de doelmaand voor het transactiem maand en de donormaanen. Hieruit valt op dat voor alle drie de doelmaanden de lijnen nagenoeg bovenop elkaar vallen en een op het oog lineair dalende trend vertonen. Extrapolatie van deze lijn lijkt er op duiden dat na 7 maanden alle mutaties verwerkt zouden moeten zijn. Dit is echter een heel ander gedrag dan de exponentiële daling die het model gebruikt. Als echter op dezelfde manier de netto cijfers worden afgebeeld (zie figuur 2), waar het model voor is bedoeld, dan lijkt het door het model voorspelde exponentiële gedrag wel gerealiseerd te worden.

Figuur1: Totaal aantal mutaties komend vanuit 1, 2, 3 of meer maanden volgend op een bepaalde maand waarvoor het transactiebestand wordt samengesteld.



Figuur2: Netto wijzigingen in standcijfers door informatie komend vanuit 1, 2, 3 of meer maanden volgend op een bepaalde maand waarvoor het transactiebestand wordt samengesteld.



Op basis van de transactie aantallen zijn de verschillen $R_2 - R_1$, $R_3 - R_1$, en waar mogelijk ook $R_4 - R_1$ en $R_5 - R_1$ berekend. Gebruikmakend van de bovenstaande vergelijking zijn deze gelijk aan:

$$R_2 - R_1 = C \left(e^{-\frac{1}{\tau}} - e^{-\frac{2}{\tau}} \right)$$

$$R_3 - R_1 = C \left(e^{-\frac{1}{\tau}} - e^{-\frac{3}{\tau}} \right)$$

$$R_4 - R_1 = C \left(e^{-\frac{1}{\tau}} - e^{-\frac{4}{\tau}} \right)$$

$$R_5 - R_1 = C \left(e^{-\frac{1}{\tau}} - e^{-\frac{5}{\tau}} \right)$$

De ratio van deze twee verschillen wordt vervolgens gebruikt om de parameter C te elimineren:

$$\frac{R_{n+1} - R_1}{R_n - R_1} - 1 = e^{-\frac{1}{\tau}}$$

waarbij de index n de waarde 2, 3, of 4 aan kan nemen. Over alle beschikbare waarden voor deze ratio kan het gemiddelde worden berekend, waaruit een schatting voor de parameter τ te bepalen valt. Hieruit volgt dat $\tau \approx 0,8$ maanden.

Deze waarde voor τ houdt in dat als een verdragingsvoet van 3 maanden wordt gehanteerd, volgens het model de transactiecijfers daarna netto met minder dan 2,5% zullen verschuiven.

4. Conclusies en aanbevelingen

Er is nog onvoldoende data beschikbaar om definitief een betrouwbare bepaling te kunnen doen van de optimale verdragingsvoet voor de SRG, om te komen tot een SRG statistiek op transactiebasis. Uit de nieuwe registratie is er voor de 6 maanden van januari t/m juni 2015 data beschikbaar, terwijl het wenselijk is om minimaal 1,5 tot 2 jaar (18 tot 24 maanden) aan data beschikbaar te hebben voor analyse.

Alvorens tot publicatie over kan worden gegaan, speelt er ook mee dat bij het invoeren van publicatie op transactie basis, er een raming moet kunnen worden gemaakt voor de transactiecijfers voor de meest recente maanden om actuele informatie te kunnen publiceren. Voor de meest recente maanden kan namelijk nog geen definitief transactiecijfer worden samengesteld omdat de noodzakelijke registratiegegevens nog niet bestaan. Voor het maken van een betrouwbare raming van transactiecijfers moeten ook minimaal 18 tot 24 maanden aan zowel register als transactie data van goede kwaliteit beschikbaar zijn.

Op dit moment zijn de totale aantallen mutaties die op de registratie bestanden moeten worden toegepast om te komen tot een transactiebestand nog zeer groot: van de grootte orde van 5% tot 10% van de totale aantallen voorzieningen, voor ieder van de 2 tot 3 opvolgende maanden op een bepaalde maandperiode. Hoewel het netto-effect op de standcijfers voor SRG voorzieningen van al deze mutaties veel kleiner is dan dit, is de grootte en het verloop met verdragingsvoet van de mutatie-aantallen zodanig dat er gereede twijfel is dat de aannames om tot een schatting van een karakteristieke tijdschaal te komen incorrect zijn.

Er zijn aanwijzingen dat de huidige aantallen mutaties gedomineerd worden door het administreren van een grote backlog of een zeer dynamische populatie waar de gemeentelijke administratieprocessen nog niet goed op toegerust zijn, en niet door het invoeren van een (eindig) aantal correcties. In het bijzonder zijn er veel voorzieningen met begindata van de administratief vertraagde instroom die meer dan een half jaar vóór een maand vallen waarover een transactiebestand moet worden samengesteld. Na goed inregelen van het maand administratiesysteem zouden vooral deze mutaties in aantal sterk af moeten nemen of zelfs geheel moeten verdwijnen.

De belangrijkste aanbeveling op dit moment is om relevante data te blijven verzamelen tot er genoeg beschikbaar is om minimaal 1 jaar van maandtransactiecijfers samen te kunnen stellen voor diverse keuzes van verdragingsvoet, waarvoor minimaal 1,5 jaar aan maand-registratiecijfers beschikbaar moet zijn.

Daarnaast is het van belang om te monitoren of de grote administratief vertraagde stromen afnemen naarmate de gemeenten hun maandregistraties beter in hebben kunnen regelen, en in samenwerking met gemeenten te zoeken naar oorzaken en maatregelen wanneer er geen verbetering plaatsvindt.