

ARTIKELEN

In de schaduw, uit de schaduw

Oorsprong, aard en mogelijkheden van schaduwverkiezingen of exit polls

Jelke Bethlehem & Joop van Holsteyn

Er wordt in Nederland veel gepeild, zeker in de aanloop naar verkiezingen. Het peilen van toekomstig stemgedrag in de aanloop van verkiezingen is inherent lastig, omdat veel kiesgerechtigden vooraf niet weten of ze gaan stemmen, laat staan op welke partij. Over de kwaliteit van die peilingen is dan ook voortdurend discussie. Er zijn echter ook peilingen die niet voorafgaand aan verkiezingen worden gehouden, maar op de verkiezingsdag. Ze worden exit polls of ook wel schaduwverkiezingen genoemd. De steekproef ervan bestaat uit kiezers die daadwerkelijk het stemlokaal hebben bezocht en hun stem hebben uitgebracht. In dit artikel willen we de eigen aard en de nuttige en interessante rol van exit polls benadrukken. Exit polls vormen een belangrijk instrument om kort na sluiting van de stembussen een nauwkeurige prognose van de uitslag te maken. Ten tweede kan een exit poll nader inzicht bieden in electorale winst en verlies, en daarmee ongefundeerde speculatie tegengaan. De verzamelde gegevens vormen immers een empirische bron voor een eerste analyse van uitslag en kiesgedrag. Alles bij elkaar is de exit poll een relatief eenvoudig te organiseren en aantrekkelijk ingrediënt voor een uitslagenavond.

Relevantie voor practitioners: verwarring van preverkiezingspeilingen met exit polls doet allicht geen recht aan de in termen van prognose veelal hogere kwaliteit van exit polls. We willen (a) uitleggen waarin de exit poll verschilt van preverkiezingspeilingen en wat de belangrijkste keuzes zijn bij het opzetten van een dergelijke peiling. We laten (b) zien dat een goed opgezette exit poll nauwkeurig is en meerwaarde heeft op een uitslagenavond. We geven (c) praktische voorbeelden en gaan in op de exit poll zoals in Leiden georganiseerd bij de raadsverkiezingen van 21 maart 2018.

1 Peilingen, peilingen, en nog eens peilingen ...

Opiniepeilingen, ontkom er maar eens aan! Vooral in de aanloop naar verkiezingen zijn er vele peilingen, meer nog naarmate de verkiezingsdag nadert. Voormalig politicus André Rouvoet sprak zelfs van 'een ware terreur van opiniepeilers in ons land' (Witteveen, 2006: 14). En een enkeling meent dat we in de moderne westerse representatieve democratie leven in een zogenoemde *mobocracy*, waarin het politieke proces in het algemeen en het verkiezingsproces in het bijzonder gedomineerd worden door peilingen (Robinson, 2002). Zo ver is het (nog?) niet,

maar het is waar: politici, media en burgers willen welhaast voortdurend weten hoe de politieke wind vandaag en morgen waait, zeker met verkiezingen in zicht.

Om huidige en toekomstige politieke krachtsverhoudingen in beeld te krijgen, vragen opiniepeilers in de regel aan een selectie van kiesgerechtigden of ze gaan stemmen, en zo ja, op welke partij, lijst of kandidaat. Voor veel Nederlandse burgers is die vraag lastig te beantwoorden, bijvoorbeeld omdat ze nog niet weten of ze gaan stemmen. En als ze zeggen te gaan stemmen, dan weten ze vervolgens steeds vaker nog niet op welke partij. Bij de Tweede Kamerverkiezingen van 2017, bijvoorbeeld, besliste 40% van de kiezers in de laatste dagen voor de verkiezingsdag naar welke partij de stem zou gaan. Niet minder dan 15% hakte pas op die dag de electorale knoop door (Van Holsteyn & Irwin, 2018: 31). Er bestaan aldus bij het peilen van partijvoorkeuren en het daarop baseren van voorspellingen van de uitslag grote onzekerheden. Die zorgen voor discussie over de validiteit en nauwkeurigheid van peilingen als voorspellingen en daarmee over de waarde van peilingen voor politici, media en burgers.¹ Zo af en toe maakt dan ook de vraag of peilingen beperkt of verboden moeten worden deel uit van die soms heftige discussie (zie o.a. Blauw, 2017; Van Holsteyn & Irwin, 2006).

De discussie gaat veelal over peilingen *voorafgaand* aan verkiezingen. Er is echter een andere, enigszins alternatieve vorm van peilen, de zogenoemde schaduwverkiezing (niet te verwarren met bijvoorbeeld verkiezingen onder middelbare scholieren, die eveneens soms onder deze noemer plaatsvinden). In de Engelstalige literatuur spreekt men van *exit poll*, die in de schaduw van reguliere peilingen staat en ‘a special case’ van opiniepeilen betreft (o.a. Moon, 1999: 134). Zo’n schaduwverkiezing of *exit poll* – we gebruiken in het navolgende beide termen als elkaars synoniem – is een peiling die *op de dag* van de verkiezingen plaatsvindt. Bij deze peiling bestaat de doelgroep uit alle personen die feitelijk hun stem hebben uitgebracht in een stembureau. Alle opgekomen kiesgerechtigden krijgen – of een steekproef eruit krijgt – in of in de directe nabijheid van het stembureau een vragenlijst voorgelegd. Die vragenlijst kan uit één enkele vraag bestaan: op welke partij hebt u zojuist gestemd? Vaak bevat de vragenlijst nog enkele andere vragen; meer dan één A4’tje beslaat de vragenlijst echter zelden of nooit.

Waarom schaduwverkiezingen? Daar zijn meerdere redenen voor (zie ook o.a. Frankovic, 2008) en in deze bijdrage willen we de relevantie van die redenen schetsen en het belang van exit polls beargumenteren en illustreren. In de eerste plaats levert een schaduwverkiezing interessante informatie voor de media op. Die kunnen namelijk snel na het sluiten van de stembureaus een prognose van de uitslag geven. Het officiële tellen van de stemmen kan in Nederland uren duren, zeker zolang de stemcomputer taboe is – meer onder jongere dan oudere kiezers

1 In dit licht moet de bijeenkomst ‘Peilingoproer’ worden gezien, die in januari 2017 plaatsvond en waarbij ‘ruim 100 peilers, journalisten en wetenschappers’ zich bogen over de vraag: ‘Hoe kunnen we opiniepeilingen en het gebruik daarvan verbeteren?’ Zie www.universiteitleiden.nl/nieuws/2017/01/peilingoproer-hoe-beter-om-te-gaan-met-peilingen.

(Loeber, 2018: 28) – en het handmatig tellen zich hardnekkig handhaaft.² In Nederland worden uitslagen soms pas diep in de nacht bekend en dan meestal nog onvolledig ook. Daar willen en kunnen de media niet op wachten. Schaduwverkiezingen zitten echter zo in elkaar, dat ze bijna onmiddellijk na het sluiten van de stembureaus een prognose leveren. Dat is interessant voor de media die 's avonds al verslag wensen te doen over de verkiezingen en de uitkomst ervan. Die informatie is zeker zo interessant voor politici, die nagelbijtend uren moeten wachten op de werkelijke uitslag. En op basis van de gegevens van de schaduwverkiezing kan een uitslagenavond worden 'opgeleukt' door middel van in de regel goede prognoses van de uitslag.

Er is meer. Op het officiële stembiljet kan slechts één partij, of althans één kandidaat van een partijlijst, worden gekozen. Daarmee houdt het op. Schaduwverkiezingen bieden extra mogelijkheden tot informatieverzameling. Naast de vraag op welke partij men heeft gestemd, kan worden gevraagd naar het stemgedrag bij vorige verkiezingen evenals naar zaken als geslacht, leeftijd en opleiding van de kiezer. Zo genereert een schaduwverkiezing extra gegevens. Analyse daarvan biedt inzicht in onderliggende electorale bewegingen. Waar kwam de winst van een partij vandaan en waar ging het verlies naartoe, onder welke groepen kiezers deed een partij het opmerkelijk goed of juist belabberd? Die informatie is van belang bij het duiden van de verkiezingsuitslag en voorkomt dat journalisten en politici op de verkiezingsavond speculeren zonder degelijke empirische grondslag. De schaduwverkiezing helpt aldus bij het empirisch gefundeerd verlenen van betekenis aan een verkiezingsuitslag.

In landen waar manipulatie van de verkiezing vermoed of gevreesd wordt, kan een schaduwverkiezing trouwens een instrument zijn om een officiële verkiezing te controleren. Als de uitslag van een schaduwverkiezing in een stembureau sterk afwijkt van de latere officiële uitslag in dat bureau, dan is dat geen sluitend bewijs, maar een serieuze aanwijzing dat er gesjoemeld kan zijn bij het tellen van de stemmen of het bepalen van de officiële uitslag. Het is geen toeval dat de Organisatie voor Veiligheid en Samenwerking in Europa (OVSE) bij het beoordelen van het eerlijke, democratische gehalte van het verkiezingsproces bij herhaling gebruik maakt van de gegevens van schaduwverkiezingen om officiële uitslagen dan wel observaties van rapporteurs te valideren.

In Nederland worden schaduwverkiezingen vooral – maar onregelmatig – gehouden bij verkiezingen voor de Tweede Kamer en Provinciale Staten, om de genoemde redenen. Bij de raadsverkiezingen van 21 maart 2018 waren er eveneens enkele schaduwverkiezingen. In opdracht van de NOS organiseerde onderzoeksbureau Ipsos schaduwverkiezingen in Amsterdam, Emmen, Enschede, Rotterdam, Utrecht en Weert (zie www.ipsos.com/nl-nl/de-exitpolls-van-2018). En

2 Dat er misschien anders, efficiënter, zou kunnen worden geteld, is een idee dat al erg oud is: reeds in 1896 pleitte jhr. A.F. de Savornin Lohman voor het gebruik van enigerlei vorm van stemmachines (zie De Beaufort e.a., 2018: 344).

de Universiteit Leiden organiseerde een schaduwverkiezing in Leiden, waarover zo dadelijk meer.

We gaan in de volgende paragrafen in op de essentie van de exit poll, op de voornaamste keuzes bij het opzetten ervan, en op enkele praktische voorbeelden,³ om verwarring tussen deze peilingen en preverkiezingspeilingen weg te nemen en de ‘ware aard’ van de exit poll te tonen. Verwarring leidt niet alleen tot misverstanden, maar doet evenmin recht aan aard en (potentieel) nut van exit polls, in het bijzonder de hoge kwaliteit van de meeste exit polls als prognoses. We geven aan waarin de exit poll verschilt van de preverkiezingspeiling. We laten zien dat een goed opgezette exit poll een nauwkeurige voorspeller is. En dat allerlei interessante en relevante analyses mogelijk zijn met de verzamelde gegevens. Daarmee is een exit poll, die met relatief bescheiden middelen te organiseren is (ook en wellicht juist op lokaal niveau), een nuttige toevoeging aan de activiteiten op een uitlagenavond en het begrip van kiezers en kiesgedrag.

2 De eerste schaduwverkiezing

Zoals vaker het geval is bij succesvolle uitvindingen: er zijn uiteenlopende versies van het ontstaan van de schaduwverkiezing als peilingvariant. Velen houden het er stellig op dat de allereerste *exit poll* in november 1967 onder verantwoordelijkheid van Warren Mitofsky plaatsvond in Amerika (o.a. Moore, 1992: 255; Scheuren & Alvey, 2008). Dat was echter zeker niet de eerste schaduwverkiezing. Als de Amerikanen al willen claimen de uitvinders van deze alternatieve manier van peilen te zijn, dan kunnen ze hooguit wijzen op een in 1964 min of meer per ongeluk gehouden peiling die de trekken had van wat schaduwverkiezingen zouden worden. Toen had journaliste Ruth Clark namelijk geen zin om huizen langs te gaan om mensen naar hun politieke voorkeur te vragen, maar besloot zij zich te posteren bij een stembureau en daar de mensen naar hun keuze te vragen (Frankovic, 2008: 574-575).

Wij houden het er liever op – niet vanwege chauvinisme, maar met de kalender van 1967 in de hand – dat de schaduwverkiezing een Nederlandse uitvinding is (Bethlehem, 2017; vgl. de rubriek Factcheck in *de Volkskrant*, 15 maart 2017). Op 15 februari 2017 was het precies vijftig jaar geleden dat de eerste schaduwverkiezing werd georganiseerd, te Utrecht (zie ook Van Dam & Beishuizen, 1967; Dooremalen & Van ’t Hul, 2013). Dat gebeurde in het kader van de Tweede Kamerverkiezingen van die dag. Bedenker en ‘uitvinder’ was de socioloog Marcel van Dam, toen jonge onderzoeker en later onder meer Kamerlid, staatssecretaris, minister en mediaman in diverse verschijningsvormen. Van Dam vroeg (beter: liet medewerkers vragen) aan stemmers in vier stembureaus in Utrecht om een korte vragenlijst (zie figuur 1) in te vullen in het kader van zijn ‘verkiezingsonderzoek’. Op grond van de verkregen informatie kon hij de verkiezingsuitslag voor de hele stad

3 Deze praktische voorbeelden dienen primair ter illustratie. We hebben niet de ambitie om een overzicht te presenteren van alle ooit in Nederland gehouden exit polls.

Jelke Bethlehem & Joop van Holsteyn

Figuur 1 De vragenlijst van de eerste schaduwverkiezing, 15 februari 1967

**WILT U HET HOKJE ACHTER HET
JUISTE ANTWOORD ROOD MAKEN?
(ER ZIJN VIER VRAGEN).**

VRAAG 1 Op welke partij hebt U zojuist gestemd?

K. V. P.	<input type="checkbox"/>	Boerenpartij	<input type="checkbox"/>
P. v. d. A.	<input type="checkbox"/>	G. P. V.	<input type="checkbox"/>
V. V. D.	<input type="checkbox"/>	Lib. Volkspartij	<input type="checkbox"/>
A. R. P.	<input type="checkbox"/>	Noodraad	<input type="checkbox"/>
C. H. U.	<input type="checkbox"/>	Landsbelangen	<input type="checkbox"/>
C. P. N.	<input type="checkbox"/>	Machiela	<input type="checkbox"/>
P. S. P.	<input type="checkbox"/>	C. D. U.	<input type="checkbox"/>
S. G. P.	<input type="checkbox"/>	Democr. '66	<input type="checkbox"/>
Lijst Voogd	<input type="checkbox"/>	Partij v. Ongehuwd.	<input type="checkbox"/>

Blanco (ongeldig)

VRAAG 2 Op welke partij hebt U de laatste keer, dus bij de gemeenteraads-verkiezingen, 1 juni 1966, gestemd?

Was nog te jong om te stemmen

K. V. P.	<input type="checkbox"/>	Boerenpartij	<input type="checkbox"/>
P. v. d. A.	<input type="checkbox"/>	G. P. V.	<input type="checkbox"/>
V. V. D.	<input type="checkbox"/>	C. D. U.	<input type="checkbox"/>
A. R. P. (+S. G. P.)	<input type="checkbox"/>	Partij v. Ongehuwd.	<input type="checkbox"/>
C. H. U.	<input type="checkbox"/>	Wereldburgers	<input type="checkbox"/>
C. P. N.	<input type="checkbox"/>	Volksw. en Refer.	<input type="checkbox"/>
P. S. P.	<input type="checkbox"/>		

Blanco (ongeldig) }
Niet gestemd }

VRAAG 3 In welke leeftijdsgroep hoort U?

21 tot en met 30 jaar	<input type="checkbox"/>
31 tot en met 40 jaar	<input type="checkbox"/>
41 tot en met 50 jaar	<input type="checkbox"/>
51 tot en met 65 jaar	<input type="checkbox"/>
66 jaar en ouder	<input type="checkbox"/>

VRAAG 4 Tot welk Kerkgenootschap rekent U zich?

Rooms Katholiek	<input type="checkbox"/>
Nederlands Hervormd	<input type="checkbox"/>
Gereform. Gezindten	<input type="checkbox"/>
Andere	<input type="checkbox"/>
Geen	<input type="checkbox"/>

**WILT U HET FORMULIER INGEVULD
IN DE BUS DOEN?
DANK U!**

Utrecht en wellicht de Randstad als geheel nauwkeurig voorspellen, zo meende hij, en dit mocht hij zelfs 's avonds op televisie uitleggen in het NTS Journaal van 15 februari 1967.

De vragenlijst van de eerste schaduwverkiezing bevatte vier vragen. Bij de eerste vraag moesten deelnemers aangeven op welke partij ze 'zojuist' hadden gestemd; een vraag als deze is de essentie van elke schaduwverkiezing. Bij de tweede vraag ging het om het stemgedrag bij vorige verkiezingen, hier de raadsverkiezingen van 1 juni 1966. Er kunnen twee redenen zijn om te vragen naar dergelijk eerder stemgedrag:

- Er kan worden nagegaan of en hoe kiezers van partijkeuze zijn veranderd. Aan welke partijen zijn verliezende partijen kiezers kwijtgeraakt? Waar hebben winnende partijen kiezers gehaald? Wat deden nieuwe kiezers?
- Mocht de groep deelnemers in een stembureau niet representatief zijn, bijvoorbeeld als gevolg van selectieve non-respons, dan kun je de antwoorden op de tweede vraag gebruiken om te proberen te corrigeren; de uitslag van de vorige verkiezingen is immers bekend.

Bij de derde vraag gaven deelnemers hun leeftijd aan en bij de vierde vraag tot welk kerkgenootschap zij zich rekenden; in 1967 vond kiesgedrag nog in hoge mate volgens verzuilde patronen plaats en was religie een dominante factor in het electorale keuzeproces (zie o.a. Irwin & Van Holsteyn, 1989). De onderzoekers

gebruikten overigens twee soorten vragenformulieren: witte en gele. Mannen kregen een wit en vrouwen een geel formulier. Zo werd, zonder (extra) vraag te stellen, het geslacht van de kiezer vastgesteld.

Utrecht telde in 1967 in totaal 164 stembureaus; voor de schaduwverkiezing werden slechts vier bureaus geselecteerd. Die vier vormden geen *random* of aselechte steekproef uit het totaal, maar waren doelgericht zo gekozen dat verschuivingen bij eerdere verkiezingen overeenkwamen met verschuivingen in de Randstad als geheel. Verder zorgden de onderzoekers ervoor dat alle sociale klassen in het onderzoek vertegenwoordigd waren. En er mochten in de omgeving van de stembureaus geen 'storende' invloeden zijn van bijvoorbeeld kloosters en bejaardentehuizen, met hun al te bijzondere bevolking aan kiezers. De steekproef was aldus doelbewust en met het oog op het gebruik ervan representatief in opzet of *design*.

In elk geselecteerd stembureau werden *alle* stemmers uitgenodigd mee te doen aan het onderzoek. Er mochten echter geen rijen of files ontstaan in de bureaus. Dat gebeurde in de praktijk trouwens niet, omdat ook het officiële stemmen enige tijd in beslag nam: registreren als kiezer, uitvouwen stembiljet, hokje rood maken, opvouwen stembiljet, in de stembus gooien. Dat zorgde ervoor dat het onwaarschijnlijk was dat tegelijkertijd meerdere kiezers na het 'echte' stemmen de uitgang van het stemlokaal zochten. Dat ging nu min of meer een voor een en kiezer na kiezer kon aldus worden gevraagd aan het verkiezingsonderzoek mee te doen.

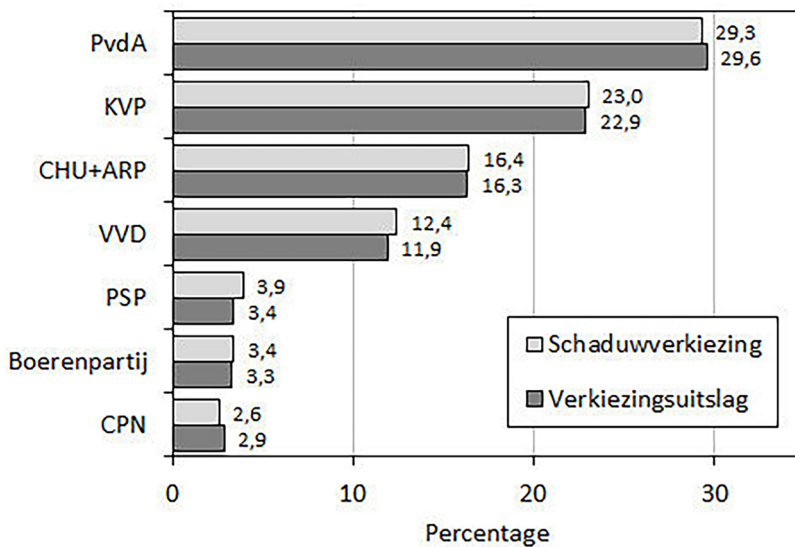
In totaal namen 3408 kiezers deel aan de Utrechtse schaduwverkiezing. Dit komt neer op 85% van alle kiezers van de vier stembureaus. Een hoge respons. Figuur 2 vergelijkt de uitslag van de schaduwverkiezing (gebaseerd op 4 stembureaus) met de officiële uitslag voor de stad Utrecht (164 stembureaus). En inderdaad: de resultaten of prognoses van de schaduwverkiezing liggen heel dicht bij de werkelijke uitslag. Het grootste verschil bedraagt niet meer dan een half procentpunt, wat zeer gering moet worden geacht, zelfs zonder een harde, objectieve maatstaf voor wat 'groot' of 'klein' is in termen van verschillen. De schaduwverkiezing bewees zich aldus direct als relatief eenvoudig te organiseren, nauwkeurig voorspellend nieuw instrument in de gereedschapskist der opiniepeilers.

3 Ingrediënten van een schaduwverkiezing

Wat 'maakt' een schaduwverkiezing? Het navolgende beschrijft cruciale ingrediënten: de doelpopulatie, de steekproef, de vragenlijst, de dataverzameling, het maken van prognoses, en tot slot de nauwkeurigheid van die prognoses.

3.1 De doelpopulatie

De zogeheten doelpopulatie is de doelgroep van de peiling: de groep mensen van wie de onderzoeker of peiler het stemgedrag wil weten/meten. Bij een schaduwverkiezing zijn dat burgers die hun stem hebben uitgebracht. Zij hebben gestemd. Merk op dat deze doelpopulatie daarmee anders is dan die van peilingen vooraf-

Figuur 2 *Schaduwverkiezing en verkiezing in Utrecht, 15 februari 1967*

gaand aan verkiezingen. Bij preverkiezingspeilingen bestaat de doelpopulatie immers uit alle kiesgerechtigden, dat wil zeggen stemmers en niet-stemmers samen. Reguliere peilingen 'gaan over' kiesgerechtigden, schaduwverkiezingen over de subgroep die daadwerkelijk van het kiesrecht gebruik heeft gemaakt, of althans de gang naar het stemlokaal heeft gemaakt.

Schijn kan echter bedriegen. Het idee dat schaduwverkiezingen betrekking hebben op *alle* kiezers is niet geheel juist. Er zijn immers kiesgerechtigden die weliswaar stemmen, maar niet langskomen in enig fysiek stembureau, waar schaduwverkiezingen plaatsvinden. Denk aan de volgende gevallen:

- *Stemmen per post.* Bij sommige verkiezingen is het mogelijk per post te stemmen, denk aan kiezers die in het buitenland wonen en hun stembiljet opsturen (zie verder Ellis e.a., 2007). Deze kiezers zullen niet in een schaduwverkiezing terechtkomen.
- *Eerder stemmen.* Bij sommige verkiezingen, in de Verenigde Staten alsook in andere landen, bestaat de mogelijkheid om voorafgaand aan de eigenlijke verkiezingsdag te stemmen. Stembureaus gaan eerder open in de hoop dat het dan bijvoorbeeld op de dag van de verkiezingen minder druk is. Als er alleen een schaduwverkiezing is op de dag van de verkiezing zelf, dan zitten dergelijke vroege stemmers er niet bij.
- *Stemmen bij volmacht.* Als stemmers niet in staat of bereid zijn om persoonlijk te stemmen, kunnen ze besluiten een andere persoon te machtigen namens hen te stemmen (zie verder Van der Kolk, 2014). Die persoon brengt dan meer dan één stem uit. In de schaduwverkiezing moet dit niet over het hoofd worden gezien. Maar het is lastig om iedere kiezer te vragen of hij meer dan

alleen de eigen stem heeft uitgebracht. Het is daarnaast de vraag of de gemachtigde de eventuele extra vragen correct kan invullen. Zo weet die gemachtigde mogelijk niet wat de volmachtgever de vorige keer heeft gestemd.

- *Ergens anders stemmen.* Het is niet zelden mogelijk om te stemmen in een ander dan het 'eigen' stembureau. In Nederland kan het zonder meer in elk stembureau in de eigen gemeente en zelfs buiten die gemeente (bij landelijke verkiezingen) als daarvoor een kiezerspas is aangevraagd. Bij de berekening van de uitkomsten van een schaduwverkiezing wordt er echter van uitgegaan dat iedereen in zijn eigen stembureau stemt, een aanname die in de loop der tijd enigszins kwetsbaar is geworden.

Door bovenstaande en wellicht nog andere uitzonderingen zullen niet alle stemmers beschikbaar zijn voor de schaduwverkiezing. Gelukkig zijn deze groepen – althans bij verkiezingen in Nederland – meestal niet zo groot, zodat de uitkomst van de schaduwverkiezing niet ernstig of wezenlijk wordt aangetast.

3.2 De steekproef

In principe zou de hele doelpopulatie van alle stemmers in alle stembureaus kunnen worden gevraagd wat ze 'zoiest' hebben gestemd. Dat is prohibitief kostbaar en tijdrovend, en overigens niet nodig. Onderzoekers ondervragen een selectie uit de populatie. Voor een nauwkeurige prognose moet die steekproef dan wel een goede afspiegeling zijn van de doelpopulatie; de steekproef moet representatief zijn. Dit streven naar representativiteit kan worden geoperationaliseerd via loting: de *random* of aselechte steekproef. Om zo'n steekproef te trekken is een volledig dekkend steekproefkader nodig: een lijst van alle eenheden (hier: mensen) in de doelpopulatie. Er bestaat echter geen lijst van personen die 'zoiest' hun stem hebben uitgebracht. Daarom wordt bij schaduwverkiezingen geopteerd voor een andere, gerichte aanpak: niet een selectie van kiezers, maar van stembureaus. In de geselecteerde stembureaus, die in een weloverwogen combinatie een representatief beeld dienen te geven van de doelpopulatie, wordt vervolgens aan alle kiezers hun keuze gevraagd. Tabel 1 bevat ter illustratie de aantallen geselecteerde stembureaus voor enkele recente schaduwverkiezingen.

Onze tabel bevat geen gegevens voor de periode 1967-2015. Dat is primair vanwege het feit dat het presenteren van alle schaduwverkiezingen geen doel op zichzelf is, maar ook omdat er relatief weinig bekend is over exit polls tussen 1967 en 2015, waarschijnlijk mede omdat door technische ontwikkelingen het einde van de schaduwverkiezing nabij leek en daardoor de animo ervoor afnam. Begin jaren negentig maakte de stemcomputer opgang in Nederland. Daarmee was de echte uitslag snel na het sluiten van stembureaus beschikbaar, en verviel de behoefte aan snelle prognoses. Wel waren er nog (soms telefonische) exit polls, onder meer voor het analyseren van winst of verlies van partijen. Na voortdurende discussie over en kritiek op de veiligheid en betrouwbaarheid van stemcomputers besloot staatssecretaris Ank Bijleveld in 2007 de goedkeuring van alle stemmachines die nog werden gebruikt in te trekken. Vanaf die verkiezingen werd er weer gestemd

Jelke Bethlehem & Joop van Holsteyn

Tabel 1 *Selectie van stembureaus*

Peiling	Jaar	Onderzoeker	Stembureaus totaal	Stembureaus geselecteerd
Tweede Kamer, Utrecht	1967	Marcel van Dam	164	4
Provinciale Staten	2015	Ipsos	9.266	43
Tweede Kamer, landelijk	2017	Ipsos	9.300	43
Gemeenteraad, Amsterdam	2018	Ipsos	560	16
Gemeenteraad, Emmen	2018	Ipsos	79	12
Gemeenteraad, Enschede	2018	Ipsos	78	15
Gemeenteraad, Rotterdam	2018	Ipsos	367	15
Gemeenteraad, Utrecht	2018	Ipsos	180	15
Gemeenteraad, Weert	2018	Ipsos	26	10
Gemeenteraad, Leiden	2018	Universiteit Leiden	53	8

met papier en rood potlood' (De Beaufort e.a., 2018: 345). Dit betekende een herleving van de interesse in en het belang van vroege prognoses en leidde tot meer exit polls. Zo waren er bij de verkiezingen voor de Tweede Kamer in 2010 en 2012 exit polls uitgevoerd door Synovate respectievelijk Ipsos Synovate.

Tabel 1 toont dat in alle gevallen een (zeer) klein deel van alle stembureaus deel uitmaakt van de steekproef, ook bij een groot totaal aan stembureaus, zoals bij Statenverkiezingen (2015) en Kamerverkiezingen (2017). Het lijkt wellicht, gezien de beoogde representativiteit, voor de hand te liggen deze in het onderzoek mee te nemen stembureaus via een aselechte steekproef uit alle stembureaus te loten. Dit gebeurt echter bewust niet. Meestal is geen sprake van een 'blinde', maar juist van een bewuste, doelgerichte selectie, waarbij het streven is een representatieve steekproef samen te stellen; er wordt door de onderzoekers een doelgerichte, naar opzet representatieve steekproef gekozen. Dat kan in dit geval ook. De uitslagen van de vorige verkiezingen zijn immers beschikbaar. Die kunnen worden gebruikt om een selectie van stembureaus te maken die in combinatie een goed beeld geven van die vorige verkiezingen. Daarbij wordt overigens wel verondersteld dat de bevolking in die stembureaus bij de actuele verkiezingen ook weer in hoge mate lijkt op de hele bevolking.

In Utrecht werden in 1967 de stembureaus al zo gekozen, dat verschuivingen bij eerdere verkiezingen overeenkwamen met de verschuivingen in de Randstad (en Utrecht). Bij de Tweede Kamerverkiezingen 2017 moest er een goede spreiding over stedelijke en landelijke gebieden zijn. Verder moest de uitslag van de vorige verkiezingen in de geselecteerde stembureaus samen overeenkomen met de landelijke uitslag bij die verkiezingen. Bij de schaduwverkiezingen in Leiden in 2018 is er nadrukkelijk op gelet dat de uitslag bij de raadsverkiezingen in 2014 in de

acht geselecteerde stembureaus samen overeenkwam met de uitslag in Leiden als geheel (met 68 stembureaus).

Dat een welbewuste keuze van stembureaus wordt gemaakt, roept vervolgens de vraag op hoeveel stembureaus er dan in een exit poll moeten worden meegenomen. In tabel 1 is te zien dat het aantal stembureaus per exit poll varieert, van 4 tot 43. Er is dan ook geen simpele, sluitende formule waarmee je het aantal benodigde stembureaus kunt berekenen. De keuze is veel meer een beredeneerde afweging. Met weinig stembureaus zijn de organisatiekosten minder. Je hebt bijvoorbeeld minder enquêteurs nodig. Maar de prognoses zijn dan wel wat minder stabiel. Een bijzondere gebeurtenis in een enkel stembureau kan een relatief grote invloed hebben op de cijfers. Worden juist veel stembureaus gebruikt, dan is de prognose stabiel. Maar de kosten zijn aanzienlijk hoger, en de marginale opbrengst in termen van stabiliteit voor elk volgend op te nemen bureau is gering. Wat overigens inzicht kan bieden, is het uitvoeren van een (herhaalde) simulatie met de uitkomsten van de vorige verkiezingen, zoals ook is gebeurd bij de Leidse schaduwverkiezing van 2018.

In elk geselecteerd stembureau wordt in beginsel aan alle stemmers gevraagd mee te doen aan de schaduwverkiezing. Dan mag echter de zogenaamde doorloopsnelheid in het 'echte' stembureau niet te hoog zijn. Als stemmers meer tijd nodig hebben voor het invullen van de vragenlijst van de schaduwverkiezing dan voor het uitbrengen van hun stem, dan kunnen er opstoppingen ontstaan, of lopen mensen simpelweg door omdat zij geen zin hebben op hun beurt te wachten om aan het onderzoek mee te doen. In zo'n geval kan (vooraf) worden besloten een steekproef van stemmers in het stembureau te vragen deel te nemen aan de schaduwverkiezing, bijvoorbeeld via een systematische steekproef. Dan wordt iedere zoveelste (zeg: derde) stemmer gevraagd mee te doen. Een systematische steekproef met een willekeurig beginpunt kan een representatief beeld opleveren van het stemgedrag binnen het betreffende stembureau. Recente voorbeelden van dergelijke steekproeven van stemmers in de stembureaus kunnen we buiten Nederland vinden. Bij de schaduwverkiezing voor de Britse parlementsverkiezingen in 2015, bijvoorbeeld, werden 141 stembureaus geselecteerd uit 50.000 stembureaus. In elk geselecteerd stembureau werd vervolgens een systematische steekproef getrokken, die tussen de honderd en tweehonderd deelnemers moest opleveren. Bij gebruikmaking van zo'n systematische steekproef is het overigens van cruciaal belang dat de onderzoekers zich strikt houden aan het selectieschema. Ze mogen zich niet bewust of onbewust laten leiden door persoonlijke voorkeuren of afkeuren – de volgende derde stemmer is een barse oudere man die ik liever oversla, ik vraag wel die vriendelijk ogende jongedame die na hem komt. Dat kiezen op eigen voorkeuren leidt ertoe dat de respondenten in een stembureau niet meer representatief zijn voor alle stemmers in dat bureau.

3.3 *De vragenlijst*

De mogelijkheden voor het stellen van vragen zijn beperkt. Het mag immers niet te veel tijd en moeite kosten – burgers komen 'even' stemmen en worden daarbij

enigszins overvallen met een verzoek om aan een onderzoek mee te doen. De vragenlijst moet dan kort zijn, om de ‘overlast’ zo veel als mogelijk te beperken.

De vragenlijst van de eerste schaduwverkiezing telde vier vragen (zie figuur 1), terwijl de kleur van het vragenformulier het geslacht van de deelnemer aangaf. Daarmee boden de verzamelde gegevens mogelijkheden voor analyse, bijvoorbeeld van verschuivingen tussen partijen en het stemgedrag van nieuwe stemmers. Verder kon de relatie tussen stemgedrag en leeftijd respectievelijk kerkgenootschap worden bekeken (zie verder Van Dam & Beishuizen, 1967). Bij de schaduwverkiezingen van Ipsos kregen de stemmers slechts één vraag, uiteraard naar de partij waarop men zojuist had gestemd. Daarmee was een prognose van de uitslag mogelijk, maar mogelijkheden tot analyse ontbraken, evenals de mogelijkheid om de uitslag van de steekproef achteraf zo nodig te corrigeren (wegen). Bij de Leidse schaduwverkiezing waren er relatief veel vragen, namelijk zeven (op een enkel eenzijdig bedrukt A4'tje): stemgedrag raadsverkiezingen 2018, stemgedrag raadsverkiezing 2014, factoren die het stemgedrag vooral bepaalden (lokale of landelijke), stemgedrag referendum, geslacht, leeftijd en opleiding. Dat leverde ruime analysemogelijkheden op, waarvan enkele voorbeelden volgen.

Er kan in de vragenlijst voor de schaduwverkiezing sprake zijn van gevoelige vragen. Kiezers willen misschien liever niet (eerlijk) zeggen op welke partij ze hebben gestemd. Ze neigen mogelijk naar een sociaal wenselijk antwoord: ze noemen een andere partij of weigeren aan te geven op welke partij is gestemd. Dergelijke antwoorden komen eerder voor als een interviewer de vragen *face-to-face* stelt (zie o.a. Kreuter e.a., 2008). Daarom is het beter om geen interviewer in te schakelen en de deelnemers de vragenlijst zelf te laten invullen en, bij een papieren vragenlijst, in een afgesloten schaduwstembus te deponeren. Utrecht had in 1967 zelfs gezorgd voor echte schaduwstembokjes!

3.4 Dataverzameling en respons

De ratio van een schaduwverkiezing is om ‘direct’ na sluiting van de stembureaus een prognose te geven van de uitslag. Dat kan alleen als gegevens tot (vlak voor) de sluiting worden verwerkt en vervolgens direct een voorspelling wordt berekend. In 1967 ging het nog niet zo snel. Na sluiting van de Utrechtse stembureaus werden de gegevens door de politie naar een computercentrum van de IBM in Rijswijk gebracht. Daar berekende een mainframe – pc’s waren er nog niet – de prognose en vervolgens werd die doorgegeven aan de NTS, de voorganger van de NOS, in Hilversum. Zo ging er enige tijd overheen, maar was er desalniettemin uren voor de werkelijke verkiezingsuitslag toch al een – naar achteraf bleek uitstekende – prognose.

Steekproefonderzoek, zoals peilingen, heeft zo goed als altijd last van *unit-non-respons*: er komen geen gegevens beschikbaar van sommige eenheden (personen) in de steekproef. Redenen kunnen grofweg zijn: geen contact, een weigering deel te nemen, en niet in staat zijn om te antwoorden op vragen. Non-respons kan leiden tot vertekening, als respondenten anders zijn dan non-respondenten. Naarmate het responspercentage lager is, is de kans op deze (forse) effecten groter.

Alleen al om die reden moet worden gestreefd naar een zo hoog mogelijke respons (zie echter Stoop, 2005). En al is het een wat aparte peiling, ook bij een *exit poll* kan non-respons optreden, in het bijzonder als gevolg van weigering. Die weigering kan diverse redenen hebben: men heeft geen tijd, men wil de privacy beschermen, et cetera.

Anders dan bij de meeste preverkiezingspeilingen, waar responspercentages tegenwoordig zeer laag zijn – de 50% wordt vaker niet dan wel gehaald – en bij telefonische onderzoeken niet zelden onder de 10% liggen, is de respons bij schaduwverkiezingen cijfermatig geen echt probleem. Daar was en is de respons hoog. De eerste schaduwverkiezing kwam tot 85% respons; vijftig jaar later was de respons bij de schaduwverkiezing voor de Statenverkiezingen met 82% nauwelijks lager. Bij de schaduwverkiezingen in Leiden was de respons met ruim 66% wel iets lager, maar nog altijd beduidend hoger dan de respons bij verreweg de meeste preverkiezings- en overige peilingen.

Er zijn verklaringen voor deze opvallend hoge responspercentages. In de eerste plaats is er een verband tussen stemmen en meedoen aan (politieke) peilingen. Stemmers zijn meer geneigd mee te doen dan niet-stemmers, en het zijn per definitie stemmers die worden gevraagd mee te doen aan de schaduwverkiezing. In de tweede plaats is het bij een slimme logistieke opstelling en aanpak letterlijk en figuurlijk lastig de onderzoekers te ontlopen. De onderzoekers zijn in de gelegenheid iedere stemmer in de geselecteerde stembureaus persoonlijk aan te spreken en te verzoeken deel te nemen. Het is dan voor velen moeilijk om nee te zeggen.

3.5 Prognoses

Prognoses zijn gebaseerd op onderliggende modellen. Die modellen zijn op hun beurt gebaseerd op veronderstellingen. De nauwkeurigheid van de prognose is daarmee afhankelijk van de juistheid van die veronderstellingen.

In Utrecht werd in 1967 een simpel model gebruikt: de verandering in het percentage stemmen voor een partij is overal (in elk stembureau) hetzelfde. Een consequentie van deze modelveronderstelling is dat de verandering in dit percentage bij alle stembureaus samen gelijk is aan de verandering in alleen de stembureaus in de steekproef. Als $P_{66}(\text{Landelijk})$ het landelijke percentage voor deze partij is bij de vorige verkiezingen in 1966, $P_{67}(\text{Landelijk})$ het landelijke percentage in 1967, en $P_{66}(\text{Steekproef})$ en $P_{67}(\text{Steekproef})$ de overeenkomstige percentages voor de steekproef, dan impliceert deze veronderstelling:

$$P_{66}(\text{Landelijk}) / P_{67}(\text{Landelijk}) = P_{66}(\text{Steekproef}) / P_{67}(\text{Steekproef}).$$

Hiervan kunnen $P_{66}(\text{Landelijk})$ en $P_{66}(\text{Steekproef})$ worden gehaald uit de uitslagen van de voorgaande verkiezingen; $P_{67}(\text{Steekproef})$ komt uit de schaduwverkiezing. De onbekende is dan $P_{67}(\text{Landelijk})$, de te voorspellen verkiezingsuitslag. Die wordt berekend door omwerking van de veronderstelling:

$$P_{67}(\text{Landelijk}) = P_{67}(\text{Steekproef}) \times P_{66}(\text{Landelijk}) / P_{66}(\text{Steekproef}).$$

Jelke Bethlehem & Joop van Holsteyn

Deze eenvoudige formules zijn toegepast in Utrecht, resulterend in een prognose (zie figuur 2) voor Utrecht als geheel die erg dicht in de buurt lag van de werkelijke uitslag.

Het is denkbaar dat zo'n Utrechts model te simpel is, bijvoorbeeld als dit niet geldt in de hele doelpopulatie, maar slechts in delen daarvan. Dan ware het beter om verschillende modellen te hanteren voor onderscheiden delen van de populatie; electorale ontwikkelingen op het platteland kunnen bijvoorbeeld anders zijn dan in stedelijke gebieden, en daardoor een ander model behoeven. In dat geval zouden twee (of meer) modellen kunnen worden gebruikt. In dit licht is het, gezien de fragmentatie en volatiliteit van het hedendaagse Nederlandse electoraat (zie o.a. Tillie e.a., 2016; Van der Meer e.a., 2018), jammer dat onduidelijk is welke modellen zijn toegepast bij de schaduwverkiezing voor de verkiezingen voor de Provinciale Staten en voor de Tweede Kamer. Ons is althans geen documentatie bekend met in dezen relevante informatie.

Hiervoor is al aangegeven dat non-respons een verstorende factor kan zijn. Het is van belang hiervoor, zo mogelijk, te corrigeren. Daarbij zijn weegvariabelen nodig, dat wil zeggen variabelen die in de peiling zijn gemeten en waarvoor de verdeling over *alle* stemmers in de stembureaus bekend is. In de schaduwverkiezing in Utrecht zijn vier hulpvariabelen gemeten: stem bij de vorige verkiezingen, geslacht, leeftijd en kerkgenootschap. Voor geslacht, leeftijd en kerkgenootschap is echter de verdeling over alle stemmers onbekend. Als weegvariabelen zijn deze dan ook onbruikbaar. Wel is de stem bij de vorige verkiezingen bruikbaar: de verdeling over de partijen bij deze verkiezingen is immers per stembureau bekend.

Dat lijkt mooi, want 'stem bij de vorige verkiezingen' is in beginsel een krachtige weegvariabele, omdat deze – in 1967 beduidend sterker dan nu, in de 21ste eeuw – gecorreleerd is met de stem bij de navolgende, actuele verkiezingen. Echter, de voorgaande verkiezingen kunnen alweer enige tijd geleden hebben plaatsgevonden, jaren geleden misschien. Weet men dan nog wat men *toen* heeft gestemd? Dit is niet vanzelfsprekend. Onderzoek heeft uitgewezen dat de herinneringsvraag naar vorig stemgedrag bepaald niet de meest valide en betrouwbare informatie oplevert. Mensen weten gewoon niet meer op welke partij ze de vorige keer hebben gestemd. Of ze weten het, maar willen het niet zeggen. Er zitten dus fouten in de antwoorden op dit soort vragen, waardoor ze uiteindelijk minder praktisch bruikbaar zijn voor een weging of ijking van uitkomsten. Om die reden is bij de schaduwverkiezing in Leiden geen gebruik gemaakt van een corrigerend weegmodel met daarin opgenomen het kiesgedrag bij de voorgaande verkiezing. Voor het berekenen van de prognose zijn simpelweg de antwoorden op de vraag naar het stemgedrag 'zou juist' gebruikt.

3.6 Nauwkeurigheid

Er kan sprake zijn van een aselechte steekproef van stembureaus waarbinnen iedere stemmer in de steekproef komt. Het betreft dan een zogenaemde clustersteekproef, te onderscheiden van de tweetrapssteekproef, waarbij een steekproef van stembureaus wordt getrokken en vervolgens uit elk geselecteerd stembureau

een steekproef van stemmers. Als er dan geen enkele selectieve non-respons optreedt, dan kunnen zuivere, valide schattingen worden berekend van de percentages kiezersaanhang voor elke afzonderlijke partij of lijst (zie voor de theoretische onderbouwing o.a. Bethlehem, 2009). De praktijk volgt echter slechts zelden deze handboekenrichtlijnen.

Schaduwverkiezingen genereren schattingen waarvan achteraf kan worden nagegaan hoe nauwkeurig deze zijn geweest. We kunnen dan ook achteraf nagaan hoe nauwkeurig de schatting is geweest, bijvoorbeeld op basis van de maat die het 'gemiddelde absolute verschil' wordt genoemd. Eerst wordt voor elke partij de waarde bepaald van het (absolute) verschil tussen het percentage in de prognose en het percentage in de uitslag. Vervolgens wordt het gemiddelde van al die afzonderlijke verschillen berekend. Een resulterend laag 'gemiddeld absoluut verschil' geeft aan dat prognose en uitslag sterk gelijken; idealiter is dit verschil 0,00, als de prognose exact overeenkomt met de uitslag. Voor elke waarde boven de 0,00 – en daarvan zal in de praktijk altijd sprake zijn – moeten we ons afvragen of deze nog als een goede schatting mag worden beschouwd. Welk gemiddeld absoluut verschil is aanvaardbaar? Op die vraag is geen objectief eenduidig antwoord mogelijk; een gouden standaard om schattingen op hun juistheid te beoordelen bestaat simpelweg niet. We hebben hier, enigszins arbitrair, gekozen voor een drempelwaarde van 1,00.⁴ Dit betekent dus dat we een exit poll als goed beoordelen als het gemiddelde verschil niet meer bedraagt dan 1 procentpunt.

Voor alle schaduwverkiezingen in tabel 1 hebben we dit gemiddelde absolute verschil berekend en gepresenteerd in tabel 2.

De prognose van de allereerste schaduwverkiezing was bijzonder nauwkeurig. De gemiddelde afwijking per partij bedroeg in 1967 namelijk slechts 0,27 procentpunt. Recente voorbeelden bij landelijke verkiezingen uit 2015 en 2017, uitgevoerd door Ipsos, waren eveneens erg nauwkeurig, al zijn de afwijkingen van 0,49 en 0,41 groter dan die van de eerste schaduwverkiezing, die in nauwkeurigheid nu eenmaal moeilijk valt te verbeteren of zelfs te evenaren. De zes meest recente lokale schaduwverkiezingen van Ipsos van 2018 hadden, met enige spreiding, een gemiddeld absoluut verschil van ongeveer 0,8 procentpunt. Dat is weer meer dan in 2015 en 2017, maar het is en blijft een goede schatting, met een gemiddeld absoluut verschil onder de 1,0. Bij de schaduwverkiezing van de Universiteit Leiden was het gemiddelde absolute verschil 0,88, een waarde in de buurt van de gemiddelde verschillen van de zes peilingen van bureau Ipsos. Dat is een uitstekend resultaat, mede gezien de onervarenheid van de studenten met het doen van (dit soort) onderzoek, en het benaderen en instrueren van respondenten in het bijzonder; ze deden het na een enkele instructie en training voor de eerste maal

4 Over deze keuze valt te twisten, bijvoorbeeld omdat binnen het Nederlandse kiesstelsel en in de Nederlandse politieke verhoudingen kleine cijfermatige verschillen aanzienlijke politieke gevolgen kunnen hebben, in de politieke praktijk en in de perceptie van politieke actoren. Omdat hier echter elke keuze uiteindelijk willekeurig is en de grens van 1 procentpunt althans intuïtief plausibel is, hanteren wij hier de door ons gekozen beoordelingsstandaard.

Tabel 2 *Het gemiddelde absolute verschil tussen schaduwverkiezingen en uitslag*

Verkiezingen	Jaar	Peiler	Gemiddelde absolute verschil
Tweede Kamer, Utrecht	1967	Marcel van Dam	0,27
Provinciale Staten	2015	Ipsos	0,49
Tweede Kamer	2017	Ipsos	0,41
Gemeenteraad, Amsterdam	2018	Ipsos	0,88
Gemeenteraad, Emmen	2018	Ipsos	0,81
Gemeenteraad, Enschede	2018	Ipsos	0,69
Gemeenteraad, Rotterdam	2018	Ipsos	0,79
Gemeenteraad, Utrecht	2018	Ipsos	0,85
Gemeenteraad, Weert	2018	Ipsos	0,52
Gemeenteraad, Leiden	2018	Universiteit Leiden	0,88

(zie ook Bethlehem & Van Holsteyn, 2018) en konden gelden als weliswaar enthousiaste en welwillende, maar professioneel onervaren onderzoekers.

4 De Leidse schaduwverkiezing van 21 maart 2018

Op 21 maart 2018 waren er in de meeste gemeenten van Nederland raadsverkiezingen. Medewerkers van de Universiteit Leiden organiseerden die dag schaduwverkiezingen in Leiden, waarnaar al herhaaldelijk is verwezen. De organisatie was in handen van Jaap de Jong, Alexander Pleijter en Willem van Rooijen van de afdeling Journalistiek en Nieuwe Media (Faculteit der Geesteswetenschappen) en Jelke Bethlehem en Joop van Holsteyn van het Instituut voor Politieke Wetenschap (Faculteit der Sociale Wetenschappen). Het praktische enquêteren in de stembureaus werd gedaan door ongeveer twintig studenten, allen in geld noch studiepunten betaalde vrijwilligers. Doel van dit interdisciplinaire project, in samenspraak met de gemeente Leiden, was het praktisch toepassen van en oefenen met leerstukken uit de survey-methodologie en het genereren van berichtgeving in diverse varianten, in oude en nieuwe media. Oefening baart kunst, ook binnen de academie. Op de drukbezochte, levendige uitslagenavond van 21 maart genoten bestuurders, burgers en buitenlui van vroege prognoses en eerste analyses – al hing dat genot mede af van het voorspelde wel en wee van de eigen partij en politieke voorkeur. Nadere analyse was op basis van het geconstrueerde databestand mogelijk, maar daar leent een uitslagenavond zich minder voor. Hierna gaan we, ter illustratie van de voorafgaande meer algemene schets en als voorbeeld van hoe een en ander relatief eenvoudig op lokaal niveau georganiseerd zou kunnen worden, wat dieper in op enkele aspecten van de Leidse *exit poll*.

Figuur 3 Een van de Leidse stembureaus (foto: Jaap de Jong)

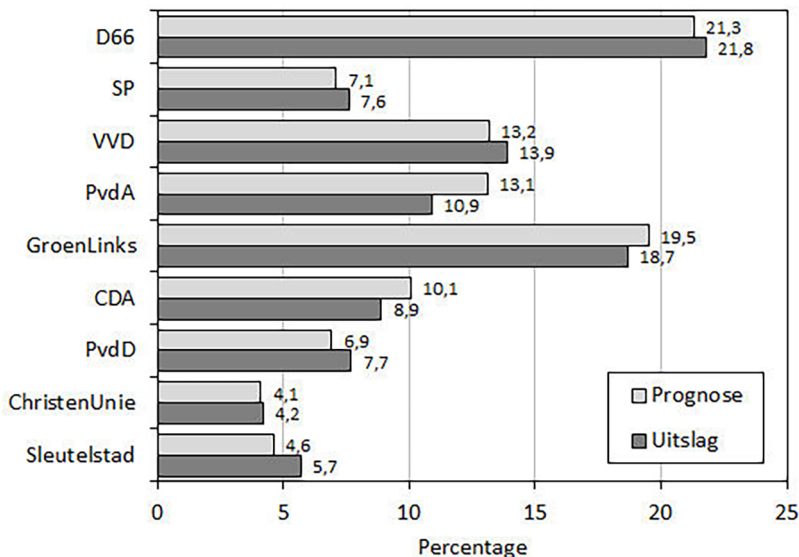
4.1 De opzet van de Leidse schaduwverkiezing

Er was besloten om acht stembureaus te selecteren. Dat aantal was een compromis: met minder stembureaus werd de prognose te kwetsbaar voor bijzondere gebeurtenissen in een enkel stembureau, met meer stembureaus zou de operatie groter en logistiek ingewikkelder worden en zou er ook meer menskracht nodig zijn. Deze acht bureaus waren zo gekozen, dat ze samen de uitslag bij de vorige gemeenteraadsverkiezingen van 2014 reproduceerden: de uitslag in de acht stembureaus was (vrijwel) hetzelfde als die in de in totaal 68 stembureaus van 2014.⁵ Via een wiskundig algoritme werd de combinatie van stembureaus bepaald die daar het beste aan voldeed. Het gemiddelde absolute verschil tussen de 8 stembureaus enerzijds en de 68 stembureaus anderzijds was 0,18.

In elk van de acht stembureaus werden twee of, bij bureaus met relatief veel te verwachten kiezers, drie studenten geposteerd als onderzoekers/enquêteurs. Ze richtten vroeg in de ochtend (voor 7.30 uur) een eigen tafel in, met daarop een 'schaduwverkiezingsstembus' (zie figuur 3), ergens bij of in de loop naar de uitgang van het stembureau en spraken vanaf die strategisch gekozen uitvalsbasis stemmers aan nadat die hun stem hadden uitgebracht. Ze waren de hele dag aanwezig, van opening (7.30 uur) tot sluiting (21.00 uur). De voorzitters van de geselecteerde stembureaus waren eerder al op de hoogte gebracht van de schaduwverkiezingen. De studenten waren herkenbaar aan hun exit poll-T-shirt en hadden

5 Ter voorbereiding en wetenschappelijke onderbouwing van de keuze van deze acht bureaus heeft de eerste auteur van deze bijdrage een aantal berekeningen gemaakt en in een memo vastgelegd; de lezer die hierin belang stelt, kan met de eerste auteur contact opnemen.

Jelke Bethlehem & Joop van Holsteyn

Figuur 4 *Prognose en uitslag in Leiden, 21 maart 2018*

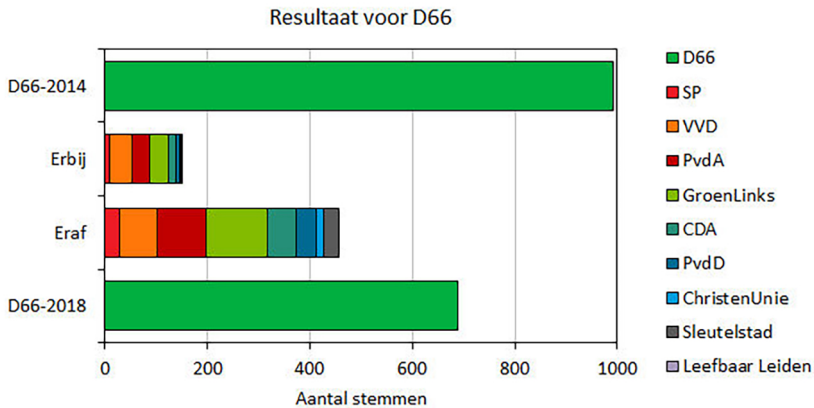
een brief bij zich waarin de burgemeester van Leiden aangaf dat hij toestemming voor de schaduwverkiezing had gegeven. Uiteindelijk vulden 5130 Leidse stemmers de vragenlijst in.

Gedurende de dag werden de geselecteerde stembureaus diverse malen door de leiders van het onderzoek bezocht om te informeren naar de gang van zaken en eventueel bij te sturen, en om ingevulde vragenlijsten op te halen. Op een centrale universitaire locatie werden deze verwerkt: eerst gescand tot pdf-bestanden, waarna via speciale software de antwoorden uit die bestanden werden gehaald. Met R-scripts werden prognoses berekend en analyses uitgevoerd.

4.2 De prognose

Een vroege en liefst correcte prognose, daar gaat het om bij een schaduwverkiezing. In Leiden werd deze prognose in twee stappen bekendgemaakt. Om ongeveer 21.30 uur was er een eerste, voorlopige prognose, gebaseerd op gegevens die tot 18.00 uur waren verzameld en vervolgens verwerkt. De definitieve prognose, gebaseerd op alle gegevens, werd gepresenteerd om ongeveer 22.30 uur. De voorlopige uitslag van de gemeente Leiden kwam trouwens pas diep in de nacht, als eerste via Twitter om 3.20 uur. Dat is rijkelijk laat voor een betekenisvolle, interessante uitslagenavond.

Figuur 4 vergelijkt de definitieve prognose met de uiteindelijke uitslag in Leiden. Duidelijk is te zien dat prognose en uitslag dicht bij elkaar zitten. De meeste verschillen liggen onder de 1 procentpunt. Het grootste verschil is dat voor de PvdA.

Figuur 5 Winst en verlies D66 bij de Leidse schaduwverkiezingen 2018

Die deelpoging is namelijk om voorspellingen om onduidelijke oorzaken 2,2 procentpunt hoger dan het werkelijk behaalde percentage stemmen. Het gemiddelde absolute verschil is overigens 0,88 (zie ook tabel 2). Dat is onder de 1 procentpunt die door ons (nogmaals: enigszins willekeurig) als kwaliteitseis is gesteld, en daarmee mag de conclusie zijn dat de prognose van de uitslag voor de Leidse gemeenteraadsverkiezingen van 21 maart 2018 nauwkeurig is geweest.

4.3 Analyse

De Leidse schaduwverkiezing kende zoals gezegd een relatief lange vragenlijst. Dat is een klein risico voor de respons, maar biedt interessante mogelijkheden voor analyse en kan leiden tot beter inzicht in het lokale verkiezingsproces. We geven twee voorbeelden.

Het eerste voorbeeld betreft een analyse van winst- en verliesstromen. Omdat in de schaduwverkiezing is gevraagd op welke partij deelnemers aan het onderzoek in 2018 en 2014 hebben gestemd, kunnen we nagaan waar de winst van een partij vandaan kwam (mensen die de vorige keer niet op die partij hebben gestemd en nu wel) en waar het verlies van een partij naartoe ging (mensen die de vorige keer op die partij hebben gestemd en nu niet meer).

Volgens de officiële uitslag van de Leidse raadsverkiezingen was D66 de grote verliezer. De partij bleef weliswaar de grootste in de raad, maar viel terug van 12 naar 9 zetels in een gemeenteraad van 39 zetels. Het simpele gegeven '3 zetels verlies' verhult echter onderliggende kiezersstromen, van én naar de partij. Uit de gegevens van de schaduwverkiezingen blijkt dat D66 zowel verloren als gewonnen heeft. Alleen verloor D66 meer kiezers dan zij wist te winnen. Dit uiterst dynamische proces is in figuur 5 zichtbaar gemaakt, op basis van gegevens van kiezers die beide stemvragen (voor 2018 en 2014) hebben beantwoord.

De bovenste balk of staaf in figuur 5 geeft de uitkomst voor D66 weer bij de raadsverkiezingen in 2014. Bijna 1000 respondenten stemden toen op D66. De onderste balk toont de situatie voor 2018. De aanhang van D66 is gezakt tot onder de 700 stemmen. Dat is het resultaat van twee gecombineerde electorale processen:

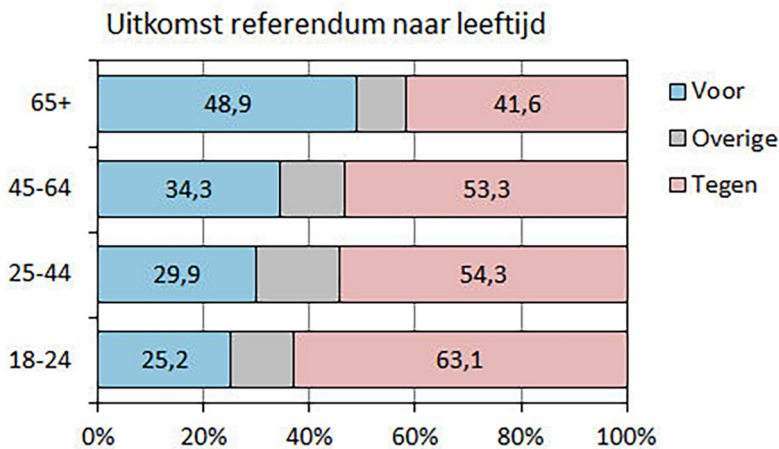
- D66 verloor bijna de helft van de aanhang (zie de staaf ‘Eraf’) van 2014. Veel kiezers van toen blijken te zijn overgestapt naar GroenLinks, maar ook verliest D66 stevig aan PvdA, VVD en CDA.
- Er is een bescheiden toestroom van nieuwe D66-kiezers (zie de staaf ‘Erbij’). Dat zijn vooral kiezers die de vorige keer nog op VVD, PvdA en GroenLinks stemden.

Dit ene voorbeeld laat al zien dat het stemproces en de einduitkomst in termen van zetelwinst of -verlies bepaald niet eendimensionaal of rechtlijnig zijn. Het is niet zo dat een partij alleen maar wint of alleen maar verliest. De verliezer D66 verliest inderdaad behoorlijk aan winnaar GroenLinks, maar winnaar GroenLinks ziet op hetzelfde moment kiezers de oversteek maken naar D66! Veel minder weliswaar, maar toch. Het is, kortom, uiteindelijk het resultaat van allerlei deels tegengestelde electorale bewegingen die onder de streep een positief of negatief saldo aan stemmen en zetels opleveren. Als de uitslag het gerecht is dat op de uitslagenavond kant-en-klaar op tafel komt, dan biedt de schaduwverkiezing een kijkje in de keuken van de dynamische verkiezingsdag.

Het tweede voorbeeld, nogmaals slechts ter illustratie van de potentie van de exit poll, gaat over het landelijke referendum dat gelijktijdig met de raadsverkiezingen werd georganiseerd. Kiezers konden aangeven of ze voor of tegen de Wet op de inlichtingen- en veiligheidsdiensten (Wiv) – veelal de Sleepwet genoemd – waren. In de Leidse schaduwverkiezing is aan deelnemers gevraagd of ze gestemd hebben voor het referendum. En als ze hebben gestemd, konden ze aangeven of ze voor of tegen hebben gestemd. Omdat ook hun leeftijd is opgegeven, kunnen we nagaan of er verband is tussen voor/tegen stemmen bij het referendum en leeftijd. Figuur 6 toont dit verband.

Er is een duidelijk patroon: naarmate (Leidse) stemmers jonger zijn, is de tegenstand tegen de Wiv groter. Jongeren hebben in meerderheid tegen gestemd. Bij jongeren tussen de 18 en 24 jaar is 61% tegen de wet waarover het referendum ging. En ook bij de respondenten tussen de 25 en 64 jaar is een krappe meerderheid tegen. Bij Leidse kiezers van 65 jaar en ouder zijn er echter meer voorstanders dan tegenstanders, al is er ook onder deze groep geen meerderheid aan voorstanders te vinden.

De twee gegeven voorbeelden laten zich eenvoudig met vele andere aanvullen, maar maken duidelijk dat er interessante nadere analyse mogelijk is met de gegevens uit een schaduwverkiezing. Een goede schaduwverkiezing is aldus in potentie een prognose-plus. Al zal eveneens duidelijk zijn dat nog meer en andere gegevens van omvattender kiezersonderzoek nodig zijn om diepgaandere analyses

Figuur 6 Stemgedrag referendum naar leeftijd

mogelijk te maken en een sluitende verklaring te vinden voor hedendaagse electorale processen in Nederland. Als het Nederlandse electoraat zich heden ten dage nog leent voor een dergelijke ambitieuze sluitende verklaring, maar dat is een vraag die we hier slechts opwerpen.

5 Slotopmerkingen

De vele reguliere opiniepeilingen dreigen de exit poll of schaduwverkiezing te overwoekeren en daar het zicht op te ontnemen. Dat zou jammer zijn, gezien de eigen aard en mogelijkheden die deze peilingvariant heeft. Wij hebben daarom het stiefkindje van de opiniepeilingen uit de schaduw gehaald en in het zonnetje gezet. En we kunnen concluderen dat deze schaduwverkiezingen of *exit polls*, ook op lokaal niveau, in de zo wisselvallige en beweeglijke electorale situatie van het begin van de 21ste eeuw prima kunnen functioneren en behoorlijk nauwkeurig als prognose zijn – wat ten aanzien van de reguliere preverkiezingspeilingen bepaald minder evident is. Een enigszins hilarische anekdote onderstreept die nauwkeurigheid van de schaduwverkiezingen. Op 7 mei 2015 waren er parlamentsverkiezingen in het Verenigd Koninkrijk. Preverkiezingspeilingen voorspelden een nek-aan-nekrace tussen Labour en de Conservatieven. Toen kwam om 22.00 uur de prognose, op basis van schaduwverkiezingen. Die was, tot grote verbazing van iedereen, heel anders dan wat op basis van eerdere peilingen verwacht werd. De schaduwverkiezing voorspelde een ruime overwinning voor de Conservatieven!

Voormalig leider van de Liberaal-Democraten Paddy Ashdown was een van de velen die zijn ongeloof in de uitkomst van de schaduwverkiezing niet onder stoe-

Jelke Bethlehem & Joop van Holsteyn

len of banken stak: *'If this exit poll is right I will publicly eat my hat.'*⁶ We weten het nu: de schaduwverkiezing had het bij het rechte eind. Twee dagen later kreeg Ashdown voor de camera's van de BBC een taartversie van een hoed ter consumptie aangeboden.⁷

Wij pleiten hier uiteraard niet voor het op grote schaal na verkiezingen opeten van al dan niet speciaal klaargemaakte hoeden. Wel zouden we in het bredere, levendige en soms felle debat over verkiezingspeilingen een scherp onderscheid willen aanbrengen tussen de reguliere preverkiezingspeilingen en de schaduwverkiezing. Die laatste variant heeft misschien een bescheiden rol op het druk bezette peilingtoneel en treedt ook uitsluitend op de verkiezingsdag zelf op, maar weet op dat moment, mits goed uitgevoerd – en dat kan met relatief weinig middelen en menskracht, zoals de Leidse schaduwverkiezing van 21 maart 2018 heeft laten zien –, de rol van voorspeller van de uitslag en bron van enig nader inzicht uitstekend waar te maken. Gegevens van schaduwverkiezingen zijn bruikbaar om een goede eerste indruk van de uitslag te krijgen en om het debat over die uitslag empirisch onderbouwd te bespreken en te duiden. In een tijd dat nepnieuws en *fake politics* hoogtij lijken te vieren, is dat geen geringe verdienste.

Literatuur

- Beaufort, F. de, P. van den Berg, R. de Jong, H. van der Kolk & P. van Schie, *Tussen geschiedenis en grondrecht. De ontwikkeling van het Nederlandse kiesrecht vanaf 1795*, Amsterdam: 2018.
- Bethlehem, J.G., *Applied survey methods. A statistical perspective*, Hoboken, NJ: 2009.
- Bethlehem, J.G., *Vijftig jaar exitpolls!*, Stuk Rood Vlees, 24 mei 2017, gevonden op <http://stukroodvlees.nl/vijftig-jaar-exitpolls/>.
- Bethlehem, J.G. & J.J.M. van Holsteyn, 'Leidse studenten peilden als ware profs', *Mare*, 41/24, 29 maart 2018, gevonden op www.mareonline.nl/archive/2018/04/04/opinie-leidse-studenten-peilden-als-ware-profs.
- Blauw, S., *Moeten peilingen vlak voor de verkiezingsdag verboden worden?*, 7 maart 2017, gevonden op <https://decorrespondent.nl/6341/moeten-peilingen-vlak-voor-de-verkiezingsdag-verboden-woorden/92278759474-59340399>.
- Dam, M.P.A. van & J. Beishuizen, *Kijk op de kiezer*, Amsterdam: 1967.
- Dooremalen, T. & L. van 't Hul, 'Interview: de sociologie in het wereldbeeld van Marcel van Dam', *Sociologie Magazine*, 2013/3, gevonden op <https://sociologiemagazine.nl/artikel/interview-de-sociologie-het-wereldbeeld-van-marcel-van-dam>.
- Ellis, A., C. Navarro, I. Morales, M. Gratschew & N. Braun, *Voting from abroad. The International IDEA Handbook*, Stockholm: 2007.
- Frankovic, K.A., 'Exit polls and pre-election polls', W. Donsbach & M.W. Traugott (eds.), *The SAGE Handbook of public opinion research*, Thousand Oaks: 2008, p. 570-579.

6 Zie www.youtube.com/watch?v=Y_l7-l2XnsY.

7 Zie bijvoorbeeld www.dailymail.co.uk/news/article-3074212/Paddy-Ashdown-given-chocolate-hat-eat-Question-Time-following-comments-ridiculing-exit-poll-accurately-predicted-Lib-Dem-collapse-huge-SNP-gains.html.

- Holsteyn, J.J.M. van & G.A. Irwin, 'Wijsheid en waan. Opiniepeilingen en het politieke proces', C.C. van Baalen e.a. (eds.), *De waan van de dag. Jaarboek Parlementaire Geschiedenis 2006*, Amsterdam: 2006, p. 55-66.
- Holsteyn, J.J.M. van & G.A. Irwin, 'Wie het laatst kiest, kiest...', T. van der Meer, H. van der Kolk & R. Rekker (eds.), *Aanhoudend wisselvallig: Nationaal Kiezersonderzoek 2017*, Amsterdam: 2018, p. 30-39.
- Irwin, G.A. & J.J.M. van Holsteyn, 'Decline of the structured model of electoral competition', H. Daalder & G.A. Irwin (eds.), *Politics in the Netherlands. How much change?*, Londen: 1989, p. 21-41.
- Kolk, H. van der, *Over het aantal volmachtstemmen. Een verkennend rapport voor de Kiesraad over het gebruik van volmachtstemmen in gemeenten bij de Tweede Kamerverkiezingen van 2012, de gemeenteraadsverkiezingen van 2014 en de Europese verkiezingen van 2014*, Enschede: 2014.
- Kreuter, F.S., S. Presse & R. Tourangeau, 'Social desirability bias in CATI, IVR, and Web surveys. The effects of mode and question sensitivity', *Public Opinion Quarterly*, 2008/5, p. 847-865.
- Loeber, L., 'De eerlijkheid van verkiezingen volgens Nederlandse kiezers', T. van der Meer, H. van der Kolk & R. Rekker (eds.), *Aanhoudend wisselvallig. Nationaal Kiezersonderzoek 2017*, Amsterdam: 2018, p. 22-29.
- Meer, T. van der, H. van der Kolk & R. Rekker (eds.), *Aanhoudend wisselvallig. Nationaal Kiezersonderzoek 2017*, Amsterdam: 2018.
- Moon, K., *Opinion polls. History, theory and practice*, Manchester: 1999.
- Moore, D.W., *The superpollsters. How they measure and manipulate public opinion in America*, New York: 1992.
- Robinson, M., *Mobocracy. How the media's obsession with polling twists the news, alters elections, and undermines democracy*, Roseville, CA: 2002.
- Scheuren, F.J. & W. Alvey (eds.), *Elections and exit polling*, Hoboken, NJ: 2008.
- Stoop, I., *The hunt for the last respondent. Nonresponse in sample surveys*, proefschrift Universiteit Utrecht, Den Haag: 2005.
- Tillie, J., J.J.M. van Holsteyn, H. van der Kolk & C.W.A.M. Aarts, *Rumoer. Nederlandse kiezers en politiek 1998-2012*, Amsterdam: 2016.
- Witteman, P., 'Zwevende politici. Paul Witteman in gesprek met André Rouvoet over opportunisme in Den Haag', Amsterdam: 2006.