

De voorspelkracht van stemprogramma's

Loek Groot

**Dr. L.F.M. Groot (SISWO/ Instituut voor Maatschappijwetenschappen en ASSR/
Amsterdam School of Social Research)**

Op 10 juni 2004 zijn de Europese parlementsverkiezingen. Waarschijnlijk zullen weer vele stemprogramma's op internet beschikbaar komen om de kiezer van een stemadvies te voorzien. Tot nu toe is er weinig onderzoek gedaan naar hoe stemprogramma's hebben gewerkt en welke lessen daaruit getrokken kunnen worden. In dit artikel wordt onderzocht hoe het vermogen van kieswijzers om de politieke voorkeuren te duiden kan worden beoordeeld. Het gaat dan om de verhouding tussen het aantal vragen en de voorspelkracht van een stemprogramma, ofwel om de vraag hoe zo efficiënt mogelijk zoveel mogelijk informatie van respondenten kan worden verkregen. Ook wordt een zo kort en eenvoudig mogelijk stemprogramma, de Stenvork, gepresenteerd en inhoudelijk gemotiveerd. Vervolgens wordt beargumenteerd dat een stemprogramma niet moet bestaan uit een uniforme lijst vragen die iedereen moet beantwoorden, maar interactief aan verschillende deelnemers verschillende sets vragen voorlegt. Tenslotte wordt een recept gegeven voor een nieuw stemprogramma, bestaande uit vier onderscheiden sets vragen, dat kan worden gebruikt voor de Europese parlementsverkiezingen van 2004 (van de redactie).

Onder de *kwaliteit* van een stemprogramma zal worden verstaan het vermogen om juiste stemadviezen te geven, waarbij juist gelezen moet worden als 'in overeenstemming met de inhoudelijke politieke voorkeuren van de respondent'. Onder de *voorspelkracht* wordt verstaan het percentage stemadviezen dat overeenkomt met het daadwerkelijke stemgedrag van de kiezer. Het verschil tussen de kwaliteit en de voorspelkracht ligt voornamelijk hierin dat ook niet-inhoudelijke beweegredenen (het imago van een partij of lijsttrekker, strategische overwegingen, etc.) een rol kunnen spelen bij het stemgedrag. Een juist stemadvies hoeft dus niet overeen te komen met de keuze in het stemhokje. Bij een juist stemadvies gaat het altijd om het individuele niveau: een confrontatie van het gegeven advies met de politieke voorkeur van de kiezer in kwestie. Bij de voorspelkracht is naast het individuele niveau - heeft het stemprogramma goed voorspeld wat de kiezer gaat stemmen - ook het geaggregeerde niveau van belang: in hoeverre komt de virtuele verkiezingsuitslag op basis van het stemprogramma overeen met de opiniepeilingen van dat moment of de uiteindelijke verkiezingsuitslag. In dit artikel zal het accent voornamelijk liggen op de voorspelkracht, maar ook de kwaliteit van het stemadvies zal zijdelings aan de orde komen.

Een stemadvies vereist eerst een politieke voorkeursduiding van de respondent. Stemprogramma's¹ baseren die politieke voorkeursduiding op de antwoorden die respondenten geven op de verschillende vragen, stellingen, issues of dilemma's. Het stemprogramma fungeert

¹ Bij de verkiezingen in 2002 en 2003 waren verschillende stemprogramma's beschikbaar: Stemwijzer, Stembox, Stemtest, Stenvork, Wij kiezen partij voor u, Wat Stem Je, Ken je Lijsttrekker, Waar stem ik op, StemmingMakerij, Polderwensen, Geen woorden en de Issuevergelijker.

hierbij in de rol van politiek expert terwijl de positie van de respondent het beste kan worden vergeleken met die van een zwevende kiezer, die nog niet weet wat hij zal gaan stemmen, dan wel met de politieke leek, die niet weet voor wat hij kan stemmen. De vergelijking van een stemprogramma met een politiek expert wordt gemaakt omdat het stemprogramma, in tegenstelling tot de kiezer, bij elke vraag exact op de hoogte is welke posities de verschillende politieke partijen innemen. De kiezer geeft bij elke vraag of stelling prijs welke positie hij inneemt en op basis daarvan wordt bepaald welke partij hiermee het beste overeenkomt. In principe geldt dat als een stemprogramma alle relevante informatie benodigd voor de politieke voorkeursduiding van de respondent boven tafel zou kunnen krijgen, het dan een 100%-juist stemadvies zou moeten kunnen geven. Het laatste is vergelijkbaar met de politiek expert die met zekerheid een juist stemadvies kan geven als hij maar alle opvattingen, vooroordelen en sympathieën van een kiezer zou kennen.

In de praktijk echter moet het stemadvies worden gegeven op basis van een beperkte hoeveelheid informatie, verkregen uit de antwoorden op een beperkt aantal vragen. Het aantal vragen varieert sterk, afhankelijk van het stemprogramma. Hierbij geldt dat een politieke voorkeurstest waarbij twintig vragen moeten worden beantwoord en waarbij in één op de tien gevallen het stemgedrag goed wordt voorspeld of een juist stemadvies wordt gegeven, slechter is dan een test met slechts tien vragen en een score van één op twee. Immers, als met minder vragen toch een betere politieke voorkeursduiding (in de zin van een juiste voorspelling van het stemgedrag, van belang voor de voorspelkracht, of een juist advies, van belang voor de kwaliteit) wordt gegeven, dan is de keus niet moeilijk. Maar is een politieke voorkeurstest waarbij twintig vragen moeten worden beantwoord met een score van één op twee beter of slechter dan een test met slechts tien vragen en een score van één op drie?² Hier is de keuze moeilijker, want de eerste test heeft weliswaar een betere score, maar er moeten ook veel meer vragen, het dubbele, worden beantwoord.

De positie die in dit artikel wordt ingenomen is dat niet straffeloos meer vragen mogen worden gesteld, zonder dat dit op de een of andere manier ten goede komt aan een betere politieke voorkeursduiding. Het uitgangspunt is dus dat de voorspelkracht van een stemprogramma positief is gerelateerd aan het percentage juiste voorspellingen van het stemgedrag en negatief aan het aantal vragen. Immers, hoe meer vragen er worden gesteld, hoe meer informatie er voor het stemtestprogramma beschikbaar komt om de politieke voorkeur van de respondent te duiden. Dat de lat hoger komt te liggen naarmate het aantal vragen toeneemt ligt voor de hand, maar hoeveel hoger? Dat is de vraag die centraal staat in het eerste deel, maar die zoals zal blijken niet zo gemakkelijk kan worden beantwoord zonder vereenvoudigende veronderstellingen te maken.

In het tweede deel wordt de Stemvork behandeld. Bij de Stemvork geeft de deelnemer slechts een minieme hoeveelheid informatie prijs, op basis waarvan de politieke voorkeur moet worden geduid. Als men het motto zou huldigen dat de kiezer altijd gelijk heeft en dat derhalve een stemadvies alleen juist is als het overeenkomt met de partijkeuze van de kiezer, dan geldt het

² Een mogelijke, zeer verdedigbare positie, is dat de meer uitgebreide test beter is louter vanwege het hogere percentage juiste voorspellingen of het gemiddeld betere stemadvies. Bij het meedoen aan zo'n test is niet zozeer het aantal vragen van belang, maar of daarmee een goed stemadvies wordt verkregen. Tijd speelt dan nauwelijks een rol. Weer een andere positie is dat mensen eraan mee doen voor hun lol, als spelletje, of als tijdverdrijf. In dat geval is het een vorm van spelbederf om leuke stemtesten te diskwalificeren alleen omdat ze niet goed genoeg voorspellen of adviseren.

vermogen van de Stemvork om de politieke voorkeur juist te duiden als een ondergrens voor meer uitgebreide stemprogramma's zoals de *StemWijzer* (met 30 vragen) of *Wij kiezen partij voor u* (met ruim 40 vragen). Deze onttrekken aan de respondent een mening over allerlei issues en moeten dus in staat zijn een veel hogere score goede stemadviezen te realiseren dan de Stemvork die de politieke voorkeursduiding slechts op één enkel informatie-item moet baseren. Omdat de Stemvork operationeel is geweest bij de laatste twee Tweede Kamerverkiezingen in 2002 en 2003, is ook empirisch onderzocht of het mogelijk is met slechts enkele vragen tot een redelijke politieke voorkeursduiding in termen van links of rechts te komen. Dat blijkt het geval. Vervolgens wordt beargumenteerd dat een hoger percentage juiste voorkeursduidingen kan worden verkregen door ook andere elementen dan standpunten van partijen en de mening daarover van kiezers in beeld te brengen. Het gaat dan om de politieke kennis van de deelnemer, het achterhalen van strategische keuzen die de kiezer wil maken en zijn voor- of afkeur van bepaalde politici.

1. Het stemprogramma als opiniepeiler

Ten tijde van de verzuiling zou een eenvoudig, maar zeer effectief stemprogramma er als volgt uitzien, met tussen haakjes de voorspelling van de politieke kleur: Bent u praktiserend katholiek (KVP), gereformeerd (ARP) of hervormd (CHU)? Zo niet, behoort u tot een lagere sociale klasse (PvdA) of tot de midden- en hogere klasse (VVD)? In het midden van de jaren vijftig zou deze stemtest ruim 70 procent van het stemgedrag juist hebben voorspeld, de reden waarom verkiezingen in die tijd ook wel werden vergeleken met volkstellingen (zie Holsteyn (1994)). Dit voorbeeld laat zien dat de voorspelkracht van een stemprogramma sterk afhangt van de houdbaarheid van de onderliggende, meestal impliciete, theorie. Dit voorbeeld heeft de theorie dat iemand politieke voorkeur kan worden afgeleid uit diens geloof en bij ongelovigen uit diens sociale klasse. Het is wel zeker dat anno 2003 de loutere achtergrondkenmerken van de respondent een veel lagere voorspelkracht hebben voor het stemgedrag. Met de teloorgang van de verzuiling gaat het niet meer zozeer om de achtergrondkenmerken, maar om de politieke voorkeuren van de respondenten (en het imago van partijen en politici). Dit is precies wat stemprogramma's in kaart proberen te brengen. Stemprogramma's als de *Stemwijzer*, de *Stembox*, de *Stemtest* en de *Kieshulp* gaan er echter impliciet of expliciet van uit dat de politieke voorkeur een soort simpele optelsom is van de politieke voorkeuren op deelterreinen, zoals door de respondenten geopenbaard in de antwoorden op de stellingen of vragen. Feitelijk wordt dan de theorie gehanteerd dat de kiezer gezien moet worden als vergaarbak van opinies, oordelen en meningen die bij elkaar kunnen worden opgeteld.³ De enige nuancering hierbij is dat bepaalde vragen of deelterreinen dubbel kunnen worden gewogen in de bepaling van het stemadvies.⁴

³ Bijvoorbeeld, als de deelnemer aan een stemprogramma aangeeft voorstander te zijn van de gekozen minister-president en tegenstander van gasboringen in de Waddenzee, dan levert dat pluspunten op voor D66 en minpunten voor de VVD.

⁴ Een ideale kiezer zou bij elke vraag precies die optie kiezen die correspondeert met het verkiezingsprogramma van de partij waartoe hij behoort. Aangezien de ideale kiezer niet bestaat is het uiterst onwaarschijnlijk dat een stemadvies uit de bus rolt waarbij de respondent het voor 100 procent eens is met een bepaalde partij. Dit verklaart ook waarom stemtest- en stemadviesprogramma's de feedback geven in de vorm van: u bent het voor 60 procent eens met partij A, voor 55 procent met partij B en voor 80 procent oneens met partij X. En

De verhouding tussen het stemgedrag en de meningsvorming op deelreinen is echter dermate gecompliceerd dat het vooralsnog moeilijk toepasbaar is bij het formuleren van een stemadvies. De simpele optelling van scores op afzonderlijke issues, met eventueel dubbele weging van bepaalde thema's, is dan ook niet meer dan een slag in de lucht. Wat ontbreekt is een onderliggende theorie hoe politieke voorkeuren zijn gestructureerd of waardoor ze worden bepaald. In het gehanteerde voorbeeld in sectie 3 wordt verondersteld dat zowel de kiezers als de partijen op een eendimensionaal links-rechts spectrum geplaatst kunnen worden.⁵

Het is niet uitgesloten dat in de toekomst stemprogramma's belangrijke, additionele informatie kunnen verschaffen naast de gangbare opiniepeilingen onder een representatieve steekproef. De zetelverdeling afgeleid uit de antwoorden op de gebruikelijke vraag bij opiniepeilingen 'Wat zou u gaan stemmen als er nu verkiezingen zouden zijn?' is grilliger en onbetrouwbaarder naarmate de groep zwevende kiezers groter wordt. Zwevende kiezers weten nog niet wat ze zullen gaan stemmen, ze zweven immers. Bij opiniepeilingen zal dan ook, afhankelijk van de waan van de dag, de ene keer partij A en de andere keer partij B een flinke winst of verlies boeken. Echter, de politieke voorkeuren van zwevende kiezers op deelreinen kunnen net zo uitgesproken zijn als van de niet-zwevende kiezers. Het prototype van de zwevende kiezer is namelijk niet zozeer iemand die de ene week tegen het recht op abortus is en de andere week mordicus voor. Een zwevende kiezer is bijvoorbeeld iemand die overweegt uit strategische motieven op een andere, grote, partij te stemmen, of twijfelt tussen twee partijen.⁶ Omdat de politieke voorkeursduidingen van stemprogramma's zijn gebaseerd op de tamelijk stabiele opvattingen die het electoraat heeft over allerlei kwesties hebben stemprogramma's, in tegenstelling tot opiniepeilingen, geen last van een groot aandeel zwevende kiezers. Stemprogramma's hebben dus het voordeel boven opiniepeilingen dat naast de verdeling van de politieke voorkeuren ook verschuivingen in de onderliggende *opvattingen* van het electoraat vrij nauwkeurig in kaart kunnen worden gebracht. Met een goed uitgebalanceerd stemprogramma zou, bij een representatieve steekproef van voldoende omvang, een overzicht gegeven kunnen worden van het kiezerspotentieel van de verschillende politieke partijen. Bovendien, juist de steeds groter wordende groep (zwevende) kiezers die steeds verder afbreuk doet aan de betrouwbaarheid van de opiniepeilingen kan met behulp van een gedegen stemprogramma goed worden gepeild in haar politieke voorkeuren.

Omdat zowel de kwaliteit⁷ als de voorspelkracht van stemprogramma's nog veel te wensen overlaten is het nog lang niet zover. Illustratief hiervoor is Tabel 1, waaruit blijkt dat de werkelijke verkiezingsuitslag sterk afwijkt van de virtuele verkiezingsuitslag op basis van de versterkte

het is ook duidelijk dat de optie om bepaalde issues zwaarder te laten wegen dan andere van belang wordt. Als er alleen maar ideale kiezers zouden zijn, dan zou dat achterwege kunnen blijven. In de praktijk zullen kiezers het soms wel en soms niet eens zijn met de standpunten van de partij waarop ze stemmen. Maar als respondenten bepaalde afwijkende standpunten innemen, dan is het voor het juiste stemadvies belangrijk om te weten of ze dat doen op onderdelen die voor bepaalde partijen of de kiezer zelf van uitermate belang zijn, of juist van gering belang.

⁵ Dit overeenkomstig de dominante visie binnen politicologisch onderzoek naar de situering van kiezers en politieke partijen (zie van der Brug, 1997, voor een bewijsvoering vanuit verschillende invalshoeken).

⁶ Voor een kiezer die zich nauwelijks voor politiek interesseert daarentegen zijn niet alleen de eigen standpunten, voorzover hij die heeft, onduidelijk, maar ook welke partij welke standpunten deelt. Op deze groep wordt nader ingegaan in de laatste sectie.

⁷ Zie Groot (2002) voor een kritische evaluatie van de kwaliteit van de Stembijzetter 2002.

stemadviezen aan een kwart van het electoraat door de Stemwijzer 2003. Het verschil tussen de werkelijke en de virtuele uitslag bedraagt, gesommeerd over alle partijen, maar liefst 87%. Het CDA was met 29% de grootste partij, maar slechts 2% van de deelnemers aan de stemwijzer heeft CDA als stemadvies gekregen. Omgekeerd, de LPF en Leefbaar Nederland horen onder het Stemwijzer-electoraat bij de drie grootste partijen met 18 en 17%, terwijl ze bij de verkiezingen slechts 5% of zelfs de kiesdrempel niet haalden.

Tabel 1. De werkelijke en de virtuele verkiezingsuitslag volgens de Stemwijzer 2003 (in %)

| Partij | Werkelijk | Virtueel |
|-----------------------|-----------|----------|
| CDA | 29 | 2 |
| PvdA | 28 | 17 |
| VVD | 19 | 14 |
| Socialistische Partij | 6 | 8 |
| Lijst Pim Fortuyn | 5 | 18 |
| GroenLinks | 5 | 10 |
| D66 | 4 | 4 |
| ChristenUnie | 2 | 3 |
| SGP | 1 | 7 |
| Leefbaar Nederland | 0 | 17 |
| Totaal | 100 | 100 |

Aangenomen dat deze grote discrepantie maar ten dele kan worden toegeschreven aan dat het deelnemersveld van de stemwijzer niet representatief is voor het electoraat als geheel, of dat de virtuele uitslag is ‘vervuild’ doordat deelnemers niet serieus of meerdere malen de test invullen, dan kan het niet anders dan dat kiezers op heel andere dan puur inhoudelijke redenen hun politieke voorkeur bepalen. Het antwoord van de Stemwijzer 2003 op de vraag waarom het verschil zo groot is tussen de uitgebrachte adviezen en de verkiezingsuitslag is het volgende:

De verstrekte adviezen kunnen op geen enkele manier vergeleken worden met bij voorbeeld opiniepeilingen, laat staan de werkelijke verkiezingsuitslag. Bij opiniepeilingen staat de vraag centraal: 'wat zou u stemmen, wanneer er vandaag verkiezingen zouden zijn'. De StemWijzer vraagt, los van de waan van de dag, een oordeel over dertig onderwerpen waarover partijen van mening verschillen. Vervolgens krijg je een stemadvies in beeld voor de partij die het dichtst bij je staat, als het om inhoud van de politiek gaat. Het is en blijft een advies, geen verplichting vanzelfsprekend. In de praktijk gelden voor kiezers heel andere redenen om op een bepaalde partij of persoon te stemmen. Het imago van de lijsttrekker speelt vaak een grote rol, de sociale, economische achtergrond van de kiezer, voor sommigen de emotie op verkiezingsdag, voor anderen de onberedeneerde weerzin tegen of juist voorkeur voor een bepaalde politicus. Kortom, lang niet iedereen kiest om inhoudelijke redenen, de bestaansgrond van de StemWijzer, voor een bepaalde partij. De rangorde Verstrekte Adviezen van de StemWijzer laat wel heel goed zien tussen welke partijen de strijd om de zwevende kiezer zich afspeelt (rubriek FAQ van de Stemwijzer 2003).

Hier worden minimaal twee controversiële claims gemaakt. Ten eerste de pretentie van *stemwijzer* dat het stemadvies, als het om de inhoud van de politiek gaat, aangeeft welke partij het dichtst bij je staat. Dit zou betekenen dat bijvoorbeeld 29% van alle kiezers, namelijk het CDA-electoraat, andere meningen zijn toegedaan dan in het CDA partijprogramma staan (zie de eerste rij in Tabel 1). Zij horen eigenlijk, als het om inhoud gaat, bij andere partijen thuis en moeten derhalve andere

motieven hebben gehad om op het CDA te stemmen. Ten tweede wordt in de laatste zin de claim gelegd dat de rangorde heel goed laat zien tussen welke partijen de strijd om de zwevende kiezer zich afspeelt. Volgens Tabel 1 zou dat voornamelijk gaan om de PvdA, LPF en VVD en bijgevolg zou het CDA niet meedoen in de strijd om de zwevende kiezer. Of deze claim kan worden waargemaakt en waarop ze is gestoeld is volstrekt onduidelijk. In ieder geval laat Tabel 1 zien dat het nog lang niet zover is dat stemprogramma's kunnen concurreren met opiniepeilingen met het in kaart brengen van het stemgedrag. Voorlopig worden stemprogramma's alleen nog maar gebruikt voor het genereren van stemadviezen. Bovendien, aan het eind van dit artikel zal ik betogen dat deze adviezen alleen waarde hebben voor politieke leken. Tevens zal ik een recept geven voor een interactief stemprogramma met vier onderscheiden sets vragen waarmee een veel hogere score juiste politieke voorkeursduidingen mag worden verwacht dan met de huidige stemprogramma's.

2. De voorspelkracht van een stemprogramma

Om te kunnen beoordelen hoe goed stemprogramma's kunnen voorspellen wat iemand zal gaan stemmen bij de eerstvolgende Tweede Kamer verkiezingen is het nuttig om eerst eens retrospectief te beginnen. We kunnen mensen vragen wat ze op 22 januari 2003 hebben gestemd. De opdracht voor een stemprogramma wordt dan door middel van vragen zo goed mogelijk te voorspellen, retrospectief, wat iemand de laatste keer heeft gestemd. Bij zo een retrospectieve voorspelling is de verkiezingsuitslag weliswaar bekend, maar niet wie welke partij heeft gekozen. Het zal blijken dat door de zaken in retrospectief te plaatsen zekere onder- en bovengrenzen kunnen worden afgeleid voor de kans op een juiste voorspelling. Terugvertaald naar toekomstvoorspellingen houdt dit het volgende in: hoe goed zouden we aan de vooravond van de verkiezingen kunnen voorspellen wat iemand gaat stemmen als we de verkiezingsuitslag al wisten. In de loop van het verhaal zullen nog enkele vereenvoudigende veronderstellingen worden gemaakt om iets zinnig te kunnen zeggen over de voorspelkracht van stemprogramma's. Een van die veronderstellingen is dat tamelijk triviale, rechttoe-rechtaan vragen kunnen worden gesteld die niets aan duidelijkheid te wensen overlaten.

Als startpunt dient de uitslag van de Tweede Kamer verkiezingen in 2003, waarbij negen partijen de kiesdrempel hebben gehaald. Zonder überhaupt één vraag te stellen is de absolute ondergrens voor de kans op een goede voorspelling dus sowieso $1/9$, ofwel 11%. Van die negen partijen zijn er zes kleine en drie wat grotere. Als zonder enige aanvullende informatie dan toch een voorspelling moet worden gemaakt van het stemgedrag, dan is de beste gok natuurlijk te zeggen dat iedereen op het CDA stemt. Bij de laatste verkiezingen zou deze uitspraak dan een voorspelkracht (noem dit p_0 , waarbij de subscript 0 aangeeft dat er nul vragen gesteld zijn aan de respondenten) hebben gehad van 29%, namelijk 44/150, het zetelaantal van de CDA ten opzichte van het totaal. Dit kan als een tweede ondergrens worden gezien, wat minder absoluut dan de eerste ondergrens omdat het aan de vooravond van de verkiezingen van 2003 nog niet duidelijk is welke partij de grootste wordt, met welk aandeel en met welk verschil. Vooral dat laatste, het verschil met de runner(s) up, is belangrijk. Als het verschil tussen de grote partijen klein is, dan maakt het voor de

kans op een goede voorspelling niet zo veel uit of de voorspelling luidt ‘u stemt op partij A’ of ‘u stemt op partij B’.

Veronderstel vervolgens dat aan een willekeurige respondent één vraag mag worden gesteld. De kunst is dan natuurlijk die vraag te stellen waarmee de kans op een hit, een goede voorspelling, zoveel mogelijk stijgt. Een redelijk goede vraag zou zijn: ‘Stemt u links of rechts?’. We gaan er gemakshalve van uit dat er geen enkel misverstand bestaat over de termen links en rechts en welke partijen daartoe behoren. In het ene geval zegt de respondent: ‘ik stem links’, wat dus betekent PvdA, GroenLinks, SP of D66. De kans dat de respondent dit antwoord geeft is $f = L/T = 65/150 = 43\%$, want de linkse partijen hebben gezamenlijk $L = 65$ van de $T = 150$ zetels behaald. Met deze informatie erbij wordt de beste voorspelling natuurlijk ‘u stemt PvdA’ want de PvdA is de grootste linkse partij, met een aandeel van 42 zetels (noem dit L_g , het zetelaantal van de grootste Linkse partij) in het linkse totaal van 65. De kans dat goed wordt voorspeld wordt dan $p_{L_g} = L_g/L = 42/65 = 65\%$.

In het andere geval, met kans $1 - f = (150-65)/150 = 57\%$ is de partij van de respondent rechts en daarvoor komen in beginsel vijf partijen (CDA, VVD, LPF, CU en SGP) voor in aanmerking. De beste keus als voorspelling, bij gebrek aan meer informatie, is dan maar de grootste partij te nemen. De kans op een goede voorspelling p_{R_g} wordt $R_g/(T - L) = 44/(150-65) = 52\%$, met $R_g = 44$ het zetelaantal van de grootste Rechtse partij, het CDA.

Uit de beide voorspelkansen blijkt dat het gunstiger is om iemand te treffen wiens partij in het linker kamp zit, omdat het zetelaandeel van de grootste linkse partij in het zeteltotaal van links (p_{L_g}) hoger is dan het aandeel van de grootste rechtse partij in het zeteltotaal van rechts (p_{R_g}). De overall voorspelkans p_1 wordt een gewogen gemiddelde van de goede voorspelkansen in het gunstige en het ongunstige geval:

$$(1) p_1 = f p_{L_g} + (1 - f) p_{R_g}$$

Toegepast op de verkiezingsuitslag van 2003 is gemakkelijk te berekenen dat $p_1 = 28\% + 29\% = 57\%$. Dit komt, niet toevallig, precies overeen met het gezamenlijke zetelaandeel van de grootste linkse partij en de grootste rechtse partij in het totaal, want:

$$(2) p_1 = f p_{L_g} + (1 - f) p_{R_g} = \frac{L}{T} \frac{L_g}{L} + \frac{(T - L)}{T} \frac{R_g}{(T - L)} = \frac{L_g + R_g}{T}$$

Ten opzichte van de situatie dat niets bekend was over de politieke voorkeursrichting (links of rechts) van de respondent is er een aanzienlijke winst geboekt in de voorspelkans. Aangezien $p_0 = R_g/T$ gelijk is aan het zetelaandeel van de grootste partij, is de toename van de goede voorspelkans gelijk aan:

$$(3) p_1 - p_0 = \frac{(L_g + R_g)}{T} - \frac{R_g}{T} = \frac{L_g}{T}$$

De voorspelkans door het stellen van de vraag waardoor bekend wordt of de respondent links of rechts zit op het politieke spectrum neemt dus toe met het zetelaandeel van de grootste linkse partij in het totaal. Uit het getallenvoorbeeld blijkt dat de toename in de kans op een goede voorspelling met 28%-punten stijgt van 29% naar 57%, wat precies overeenkomt met het zetelaandeel van de PvdA in het totaal.

De uitwerking van de toename in de goede voorspelkans bij een tweede vraag is rechttoe-rechtaan. De vraag zou kunnen zijn: ‘Is de partij van uw voorkeur binnen het linkse of rechtse kamp het meest extreem?’ Stel de respondent heeft op vraag 1 geantwoord ‘rechts’ en antwoordt op vraag 2 ontkennend. Dan is de kans groot dat de respondent stemt op het CDA. Is het antwoord op vraag 2 echter bevestigend, dan kan het LPF of VVD zijn, afhankelijk welke van de twee als meest rechts wordt gekwalificeerd. Neem aan dat het de LPF is. De kans op een goede voorspelling bij twee vragen wordt nu:

$$(4) p_2 = \frac{(L_g + R_g + L_e + R_e)}{T} = p_1 + \frac{L_e + R_e}{T}$$

met

L_e = het zetelaantal van de meest linkse partij

R_e = het zetelaantal van de meest rechtse partij

Aangenomen dat de LPF de meest rechtse partij was, met $L_e = 8$ en SP de meest linkse partij, met $L_e = 9$, dan stijgt de voorspelkans volgens formule (4) met ruim 11%-punten van 57% naar 69%.

De antwoorden op deze vragen laten niets aan duidelijkheid te wensen over, terwijl dit bij politieke voorkeursvragen van een stemprogramma wel het geval is. In plaats van een ondergrens, zoals bij p_0 , moeten p_1 en p_2 dus beschouwd worden als bovengrenzen voor de goede voorspelkans. Waar het in deze exercitie om gaat is of bovenstaande methodiek bruikbaar is om in te schatten welke percentage juiste voorkeursduidingen een programma zou kunnen halen met een beperkte set vragen of issues.

Stel dat het electoraat van de partijen een bepaald, van elkaar verschillend, profiel hebben wat betreft politieke issues. Het GroenLinks-profiel luidt bijvoorbeeld: voor strenge milieumaatregelen, voor een ruimhartig asielbeleid, voor inkomensnivellering, voor een multiculturele samenleving en voor een ontspannen arbeidsmarkt.⁸ Andere partijen hebben weer andere profielkenmerken, maar kunnen deels ook overlappend zijn met het GroenLinks-profiel. Stel dat alle partijen aan de hand van een set van tien issues kunnen worden geportretteerd in termen van (extreem) links of rechts. De tien issues corresponderen in een stemprogramma met evenzoveel vragen. Als aan de hand van de eerste vijf vragen kan worden bepaald of iemands politieke voorkeur links of rechts is, dan vormt deze subset van vijf issues het equivalent van vraag 1 hierboven. Met die vijf vragen zou dan een percentage goede stemadviezen moeten kunnen worden verstrekt van *maximaal* $p_1 = 57\%$. Immers, als de politieke voorkeursduiding in termen van links of rechts aan de hand van vijf issues in bijvoorbeeld maar 75% van de gevallen zou kloppen,

⁸ Bij een ontspannen arbeidsmarkt gaat het om issues zoals vrijstelling van de sollicitatieplicht voor bijstandsmoeders met jonge kinderen en voor oudere werklozen, de mogelijkheid tot het opnemen van

dan is de voorspelkracht gelijk aan $0,75 \cdot 0,57 = 43\%$. Als met de volgende subset van vijf issues of vragen kan worden bepaald of iemands linkse of rechtse politieke voorkeur extreem links of extreem rechts is, dan zou dit percentage juiste stemadviezen moeten kunnen stijgen tot maximaal $p_2 = 69\%$ (of met $0,75 \cdot 0,69 = 52\%$ als het slechts voor 75% lukt). In dit licht bezien is de score van de Stemwijzer met 30 vragen of issues, waarbij blijkens de NIPO-enquête in een derde van de gevallen het stemadvies overeenkomt met de politieke voorkeur, aan de lage kant. Opgemerkt dient te worden dat hier alleen een uitspraak wordt gedaan over de voorspelkracht van de Stemwijzer, maar niet over de kwaliteit (het percentage juiste adviezen, wat veel hoger of lager kan liggen).

Wat deze exercitie ook laat zien is dat het beste resultaat wordt verkregen door een stemprogramma interactief te werk te laten gaan: met de eerste set vragen wordt bepaald of de deelnemer zich links of rechts op het politieke spectrum bevindt, en afhankelijk daarvan krijgt de deelnemer een tweede set vragen voorgelegd om te bepalen hoe links of hoe rechts zijn opvattingen zijn. Als bijvoorbeeld uit de eerste set vragen blijkt dat de respondent voorstander is van een gelijkere inkomensverdeling en voor een multiculturele samenleving, dan kan vervolgens in de tweede set vragen worden bepaald bij welke partij in het linkse kamp de deelnemer thuis hoort door te achterhalen hoe ver de respondent wil gaan in het bereiken van die doelstellingen. Het is duidelijk dat zodra er sterke aanwijzingen zijn dat het een kiezer is met linkse voorkeuren, het zinloos is de vraag te stellen of de onroerend zaakbelasting moet worden afgeschaft. Alle linkse partijen (en waarschijnlijk ook bijna alle linkse kiezers) zijn daar tegen, dus deze vraag discrimineert niet tussen de linkse partijen.⁹ Omgekeerd is het van belang om te weten of de kiezer voor een zeer soepel immigratiebeleid is (GroenLinks), of juist tegen (SP), of iets daar tussenin (PvdA). Geen van de bestaande stemprogramma's gaat echter op deze manier interactief te werk.¹⁰ De voorgestelde interactieve opzet valt of staat met de vraag of het mogelijk is deelnemers aan de hand van een beperkt aantal issues in links- of rechts-georiënteerd in te delen. Voordat de volgende twee sets vragen worden toegelicht wordt aan de hand van de Stemvork in de volgende paragraaf aangetoond dat het inderdaad mogelijk is met een kleine set vragen de links-rechts politieke oriëntatie van kiezers te achterhalen.

3. De Stemvork

De Stemvork bevat slechts twee simpele vragen en poogt te discrimineren tussen maar liefst negen politieke partijen.¹¹ Net zoals in het stemhokje kan maar één keer iets worden aangekruisd. Hoewel

(betaald) zorgverlof voor ouders en *sabbatical leaves*, het recht om in deeltijd te gaan werken en de mogelijkheid om bijverdiensten toe te staan aan uitkeringsgerechtigden.

⁹ Dit is het grote nadeel van een uniforme set vragen die iedereen moet beantwoorden. Een groot deel van die vragen discrimineert niet binnen de linkse of binnen de rechtse partijen, terwijl met een kleine subset van vragen de deelnemer met grote waarschijnlijkheid in een van beide kampen kan worden ingedeeld.

¹⁰ Het zou kunnen dat een interactieve opzet meer voor de hand ligt wanneer het doel slechts het juist voorspellen met zo min mogelijk vragen is en niet zozeer het geven van een juist stemadvies. Bij een juist stemadvies kunnen 'tegenstrijdige' antwoorden binnen een set waardevolle informatie geven. Bovendien gaat de vierde set (zie paragraaf 4) met vragen over niet-inhoudelijke zaken ten koste van de kwaliteit van het stemadvies louter op basis van de inhoudelijke standpunten van de respondent.

¹¹ Op de website www.self-gov.org/quiz.html is 'Worlds Smallest Political Quiz' te vinden. Om te discrimineren tussen slechts twee partijen, Democraten en Republikeinen, bestaat deze toch nog uit tien

het niet realistisch is te veronderstellen dat kiezers over allerlei issues al een weloverwogen mening hebben, hebben ze vaak wel iets wat van veel grotere waarde is, namelijk een notie waar de verschillende partijen in essentie voor staan, bijvoorbeeld dat de SP staat voor radicale sociale rechtvaardigheid, terwijl de VVD wat grotere ongelijkheid wil tolereren om meer ruimte te geven aan eigen verantwoordelijkheid. De meeste stemprogramma's pakken maar in zeer beperkte mate de essentie van de verschillende partijen, niet zozeer omdat slechts een beperkt aantal issues worden opgenomen, maar vooral omdat afzonderlijke standpunten uit partijprogramma's worden gelicht. Daarmee verdwijnt het ideologisch profiel van een partij grotendeels uit het beeld. De deelnemer aan de Stemvork geeft slechts een minieme hoeveelheid informatie prijs, op basis waarvan de politieke voorkeur moet worden geded. Niettemin zal blijken dat met zo een simpele test al in redelijke mate kan worden bepaald of respondenten politiek links of rechts georiënteerd zijn, wat in het vervolg van belang is voor de opzet van een interactief stemprogramma. Hieronder staan de twee vragen, en daarna zal worden toegelicht waarom de Stemvork zo in elkaar steekt.

Vraag 1: Is één van de volgende onderwerpen voor u van doorslaggevend belang bij het stemmen op een politieke partij, waarbij u tevens de bijbehorende stelling voor uw rekening wilt nemen. Zo ja, kies dan 'Mee eens'. Zo nee, ga dan door naar vraag 2.

| | |
|--|------------|
| Sociale rechtvaardigheid: het inkomen dat een kwart miljoen euro per jaar te boven gaat mag gerust met een belastingtarief van ruim 70% worden belast | O Mee eens |
| Milieu: duurzaamheid gaat voor economische groei, de economie moet krimpen voor behoud van het milieu | O Mee eens |
| Ethische kwesties: de wettelijke regelingen die euthanasie, abortus en het homohuwelijk mogelijk maken dienen zo snel mogelijk te worden teruggedraaid | O Mee eens |
| Democratie: burgemeesters en de minister-president moeten rechtstreeks door de burgers worden gekozen, en zoveel mogelijk beslissingen worden per volksreferendum genomen | O Mee eens |
| Overheid: het aantal ambtenaren kan gemakkelijk met een kwart worden teruggebracht, zonder dat dit gevolgen heeft voor de dienstverlening | O Mee eens |
| Criminaliteit en immigratie: het wordt hoog tijd dat de problemen rond criminaliteit en immigratie met naam en toenaam worden genoemd en eindelijk eens worden aangepakt | O Mee eens |

Vraag 2: Welk van de volgende drie stellingen spreekt u het meest aan?

| | |
|-----------------------|--|
| <input type="radio"/> | De kwaliteit van gezondheidszorg en de toegang tot sociale zekerheid is voor iedereen hetzelfde en dient te worden betaald uit algemene belastingmiddelen |
| <input type="radio"/> | De duur en hoogte van de uitkeringen dienen bij voorkeur zoveel mogelijk afhankelijk te zijn van het arbeidsverleden en het laatst verdiende inkomen |
| <input type="radio"/> | De overheid hoeft alleen zorg te dragen voor een basisvoorziening in de gezondheidszorg en de sociale zekerheid. Het bovenminimale deel moeten de burgers zelf regelen |

vragen (wel of geen militaire dienstplicht, legalisering van drugs, landbouwsubsidies, minimumloon, ontwikkelingshulp, e.d.).

Bij vraag 1 corresponderen de onderwerpen en de daarbij behorende stellingen met achtereenvolgens de SP, GroenLinks, CU, D66, Leefbaar Nederland en de LPF. De volgorde van deze partijen komt overeen met de gangbare situering op een eendimensionale links-rechts continuüm met de SP uiterst links en de LPF uiterst rechts. De stellingen bij vraag 2 corresponderen, eveneens van links naar rechts, met PvdA, CDA en VVD.

Toelichting bij de eerste vraag

Het is de bedoeling dat bij de eerste vraag alle kiezers ‘stranden’ die een bepaald onderwerp van doorslaggevend belang vinden¹² en bovendien met betrekking tot dat thema een vrij radicaal standpunt onderschrijven.¹³ Ze zullen dus het hokje aankruisen van de partij die op dat onderwerp het meest uitgesproken is. Zo is GroenLinks marktleider op het terrein van milieu, de SP vindt haar aanhang vooral onder voorstanders van radicale inkomensnivellering uit hoofde van sociale rechtvaardigheid, de CU profileert zich vooral op ethische kwesties, de kroonjuwelen van D66 hebben betrekking op de staatsrechtelijke vernieuwing, volgens Leefbaar Nederland kan en moet het overheidsapparaat veel minder bureaucratisch worden en de LPF wordt vooral geassocieerd met de no-nonsense aanpak van wetsovertreders en het sluiten van de grenzen voor asielzoekers. Hiermee is niet gezegd dat deze partijen geen andere punten hebben dan bovengenoemde. Het zijn geen *single issue* partijen. Zo behandelt het verkiezingsprogramma *Eerste weg links* van de SP het hele spectrum van zorg, milieu en onderwijs tot en met voedselveiligheid en dierenwelzijn. Bij de eerste vraag wordt echter impliciet aangenomen dat de kleine of nieuwe politieke partijen voornamelijk worden geassocieerd met een enkel onderwerp waarbij ze bovendien een standpunt innemen dat nogal afwijkt van dat van de overige partijen.

Er wordt dus verondersteld dat dat elk van de zes partijen bij vraag 1 een duidelijk afwijkend standpunt inneemt op het thema waarmee ze worden geassocieerd. Wat de overige onderwerpen betreft, kunnen ze meer of minder van andere partijen verschillen. Een voorbeeld kan dit verduidelijken. Als er slechts drie partijen zijn die het over alles min of meer eens zijn behalve over issues P, Q, en R, waarbij partij 1 een duidelijk afwijkend standpunt inneemt ten opzichte van P, partij 2 ten opzichte van Q en partij 3 ten opzichte van R, dan wordt issue P de *core business* voor partij 1, Q voor partij 2 en partij 3 wordt marktleider op het terrein van R. Ondanks dat de drie partijen allen standpunten hebben ten aanzien van de issues A tot met O, worden toch de issues P, Q en R de handelsmerken van de partijen. Bij de eerste vraag van de Stemvork wordt gebruik gemaakt van het gegeven dat de zes kleine partijen zich op bepaalde terreinen uniek onderscheiden.

¹² Het is niet uitgesloten dat er kiezers zijn die het deels eens zijn met de ene en deels met een andere partij. Bijvoorbeeld kiezers die op sociaal-economische terreinen VVD-standpunten innemen, maar in afwijking van de VVD-lijn voorstander zijn van een veel strenger milieubeleid. Bij de Stemvork is met deze eigenaardigheid rekening gehouden door bij vraag 1 expliciet te vermelden dat het onderwerp voor de kiezer van doorslaggevend belang moet zijn. Als dat zo is, dan stranden ze bij vraag 1 als GroenLinks-stemmers. Als ze het niet van doorslaggevend belang vinden, maar het wel eens zijn met de bijbehorende stelling, dan zullen ze desondanks doorgaan naar vraag 2 van de Stemvork en alsnog terechtkomen bij de VVD.

¹³ Bij de eerste vraag gaat het er dus om te achterhalen of de respondent een van de onderwerpen van doorslaggevend belang vindt bij het bepalen van haar stemkeuze. Is dat het geval, dan komt de respondent automatisch (en terecht, want het onderwerp is van doorslaggevend belang) uit bij de partij die dit onderwerp met de bijbehorende stelling propageert. Dat klopt natuurlijk niet helemaal. Ook als de stellingen bij 1 niet doorslaggevend zijn voor het stemgedrag kan iemand toch uitkomen bij één van de zes partijen van vraag 1. Ook is het mogelijk dat de respondent meerdere stellingen onderschrijft. In dat geval moet zij de belangrijkste kiezen.

Dit is tevens de reden waarom CDA, PvdA en VVD niet zijn opgenomen bij vraag 1. Deze drie grote partijen nemen op de zes onderwerpen een veel minder radicaal standpunt in dan de kleine of nieuwe partij die op het desbetreffende onderwerp het meest uitgesproken is. Tegelijkertijd zijn er geen andere issues of beleidsterreinen waar de grote drie een radicale positie innemen die door geen van de overige partijen wordt gedeeld. Zou men bijvoorbeeld het CDA willen associëren met gezinspolitiek, de VVD met veiligheid en de PvdA met een kritische houding tegenover de NAVO, dan kan op elk van die terreinen een partij worden aangewezen met een op dat terrein soortgelijke of zelfs meer extreme stellingname, te weten te weten CU, LPF en de SP.

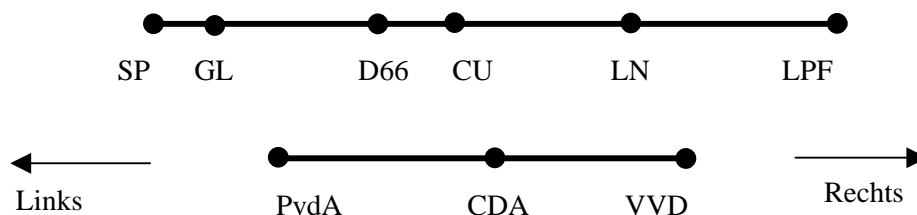
Toelichting bij de tweede vraag

Bij vraag 2 zijn de traditioneel grote partijen CDA, PvdA en VVD ondergebracht. In tegenstelling tot vraag 1 wordt van de respondent niet gevraagd of zij het met de stelling eens is, maar welke van de drie stellingen bij vraag 2 het meest aanspreekt. Het gaat er bij de tweede vraag niet zozeer om of de stelling precies aansluit bij de programma's van de desbetreffende partijen, maar om de deelnemer een keuze te ontlokken aan welke brede oriëntatie het sociaal zekerheidsstelsel moet voldoen. Dat is ook de reden waarom wordt gevraagd welke stelling het *meest* aanspreekt. In essentie gaat het om een keuze tussen drie ideaaltypen verzorgingsstaat, de liberale, sociaal-democratische en de corporatistische variant. De driedeling komt overeen met de indeling in *The Three Worlds of Welfare Capitalism* (1990) van Esping-Andersen, die ook van elk de grondtrekken heeft geschetst. De drie opties bij vraag 2 van de Stemvork representeren een keuze om de Nederlandse verzorgingsstaat te modelleren of hervormen naar een van de drie varianten. De gedachte is dat VVD-aanhangers zoveel mogelijk aan de markt willen overlaten, en dus ook het liefst het collectieve deel van het sociale zekerheidsstelsel willen beperken tot basisvoorzieningen. PvdA-aanhangers zien de arrangementen van de verzorgingsstaat niet alleen als collectieve voorzieningen om sociale risico's te dekken, maar ook als een instrument om gelijkheid en gelijke toegang tot primaire goederen zoals onderwijs en gezondheidszorg te bevorderen. Aan CDA-ers wordt een middenpositie toegedicht: sociale zekerheid dient uitgebreider te zijn dan een basis- of ministelsel, maar minder universeel dan in de ideaaltypische sociaal-democratische variant.

Het politieke spectrum

De kleine partijen onderscheiden zich niet alleen van de grote partijen door een min of meer extreme of afwijkende stellingname op een welbepaald terrein. Er is nog een andere eigenaardigheid waarop kan worden gewezen. Beschouw hiertoe de onderstaande twee balkjes.

Figuur 1. Het Nederlandse politieke landschap, gestileerd weergegeven op een eendimensionale links-rechts continuüm



De drie grote partijen zijn op de onderste links-rechts as geplaatst en de kleinere partijen op de bovenste. Het is in dit verband overigens niet zo belangrijk of de positionering van de partijen in een eendimensionale (links-rechts) ruimte of een meerdimensionale ruimte plaatsvindt. Waar het hier om gaat is de vraag in de nabijheid van welke grote partijen de kleine of nieuwe partijen zich situeren. Zoals gezegd veronderstelt de Stemvork dat het imago van GroenLinks verbonden is met milieuvraagstukken, van de SP met inkomensnivellering, etc. Daarnaast spreken de SP en GroenLinks zich in hun verkiezingsprogramma's ook uit over andere zaken dan milieu- of inkomenspolitiek. In principe is het mogelijk dat er een succesvolle milieupartij, zeg GroenRechts, zou zijn die verregaande milieumaatregelen wil nemen, maar voor het overige grosso modo de standpunten van de VVD deelt. Ook is het denkbaar dat een partij, zeg de Egalitair Liberale Partij, sociale rechtvaardigheid heel belangrijk vindt, maar net zoals de VVD zoveel mogelijk aan de markt en het particulier initiatief wil overlaten. Toch zien we die partijen niet ontstaan. Hier is niet van belang hoe dit komt, maar de vaststelling dat de kleine of nieuwe partijen zich op een bepaalde manier situeren ten opzichte van de grote partijen, namelijk dat kleine of nieuwe partijen zich daar op het politieke spectrum nestelen, waar de grote partij zit die het meest verwant is met het issue waarin de kleine of nieuwe partij marktleider is en grotendeels haar bestaansrecht aan ontleend. Dit ligt juist minder voor de hand. Het is enigszins onlogisch dat bijvoorbeeld GroenLinks of de SP zich situeren in de nabijheid van de PvdA. Kiezers die zwaar tillen aan het milieu of aan inkomensgelijkheid, worden immers deels al bediend door de PvdA, meer dan de VVD dat doet. Om kiezers van de PvdA weg te kapen zijn GroenLinks en SP gedwongen veel radicalere standpunten in te nemen op hun kernthema dan ze zouden moeten doen wanneer ze zich in de nabijheid van een grote partij situeren die een gering belang hecht aan het desbetreffende onderwerp.

Volgens deze stelling situeren de SP en GroenLinks zich in de nabijheid van de PvdA en de CU zal zich situeren in de nabijheid van het CDA, want het CDA is in ethische kwesties de aanverwante partij. De benaming 'klein rechts' voor de CU (en de SGP) is vanuit deze optiek dus een misser. Aangenomen dat het CDA zich situeert tussen de PvdA en VVD, zou 'klein midden' beter zijn. D66 is in deze optiek het minst verankerd op het politieke spectrum. Zij is een soort zwevende partij, omdat geen van de grote drie zich de afgelopen 35 jaar heeft uitgesproken voor staatsrechtelijke vernieuwing en versterking van de democratie zoals D66 sinds 1966 heeft gedaan. Zou bijvoorbeeld de VVD zich in het verleden wat meer gelegen hebben laten liggen aan staatsrechtelijke vernieuwing dan het CDA of de PvdA, dan zou D66 volgens de hierboven

geformuleerde stelling op andere terreinen dan de staatsrechtelijke vernieuwing langzamerhand zijn opgeschoven in de richting van de VVD.

De Stemvork 2002 was operationeel vanaf 1 januari 2002. In de periode tot aan zaterdag 16 maart 2002 zijn door de Stemvork 1496 uitslagen verstrekt. Probleem is echter dat de software niet zo ontwikkeld is dat deelnemers maar een keer hun stem kunnen uitbrengen. Het was mogelijk om met de Back en Forward toets steeds een andere keuze-optie uit de Stemvork aan te klikken om te kijken welke partij met welke keuze correspondeert.¹⁴ Deelnemers hebben dat ook veelvuldig gedaan, dus het aantal van 1496 uitslagen is veel hoger, ongeveer een factor vier, dan het daadwerkelijke aantal deelnemers. Om die reden is het ook niet zinvol een soort virtuele verkiezingsuitslag te genereren op basis van de 1496 uitslagen. Echter, nadat de deelnemer een keuze uit vraag 1 of anders uit vraag 2 heeft gemaakt en de uitslag is gegeven vraagt de Stemvork of de uitslag klopt, als volgt:

“Klopt de Stemvork? Ik wil het graag van u weten als de uitslag onjuist is.

Deze uitslag klopt niet.

Ik stem op

Alle informatie wordt mij anoniem toegestuurd.“

In totaal hebben 112 deelnemers van deze mogelijkheid gebruik gemaakt en aangegeven dat de uitslag niet klopt, omdat ze op een andere partij stemmen dan de Stemvork aangeeft op basis van de gemaakte keuze. Als het inderdaad klopt dat achter de 1496 uitslagen in werkelijkheid slechts $1496/4=374$ deelnemers schuilgaan, dan is de fractie fout verstrekte politieke voorkeursduidingen gelijk aan $112/374=0.30$. Dat wil zeggen dat op basis van deze ruwe berekening in 7 op de 10 gevallen de Stemvork op basis van de gemaakte keuze de politieke voorkeur juist voorspelt, maar ook in 3 van de 10 gevallen fout zit.¹⁵ Na 16 maart zijn alleen nog de foute voorspellingen geregistreerd, waarbij dus zowel bekend is wat de respondent zegt te stemmen en wat hij of zij volgens de Stemvork-test zou stemmen. In de laatste verkiezingsmaand, van 8 april tot 15 mei, waren dat er ruim 1500. Ruim 40% daarvan (662 stuks) betroffen de zes kleine of nieuwe partijen (zie Tabel 2A). Hetzelfde is gedaan bij de Stemvork voor de verkiezingen van 22 januari 2003, waarin alle foutieve uitslagen (879 stuks) tot aan 15 januari 2003 zijn verwerkt (zie Tabel 2B).

In Tabel 2A en B staat verticaal de (foutieve) uitslag van de Stemvork op de kleine of nieuwe partijen en horizontaal de partijen waarop de deelnemers zeggen zelf te stemmen. Onder de categorie ‘Overig’ vallen de kiezers van partijen die hier niet zijn opgenomen (SGP, De Groenen, seniorenpartijen, Partij van de Dieren).¹⁶ Bijvoorbeeld, in de eerste rij staan in totaal 125 deelnemers die volgens de Stemvork op de SP stemmen, maar in werkelijkheid stemmen op een

¹⁴ Na 16 maart is gepoogd dit te voorkomen door na de uitslag aan de deelnemer direct deze informatie te geven, in de vorm van een lijst waarmee de deelnemer direct kan zien welke politieke partij met welke stelling correspondeert.

¹⁵ Let wel: dit is onder de veronderstelling dat alle deelnemers waarbij de Stemvork fout zit ook reageren. Het ligt daarom voor de hand de foutscore van 30% als ondergrens te nemen, en bijgevolg de goedscore van 70% als bovengrens.

¹⁶ Het percentage ‘Overig’ valt bij de verkiezingen van 2003 licht hoger uit door de deelname van de Partij van de Dieren en het niet halen van de kiesdrempel door LN.

andere partij, voornamelijk GroenLinks (26%) of PvdA (25%). De 108 foutief geduide LPF-stemmers stemmen in werkelijkheid vooral VVD of CDA.

Tabel 2A. De verdeling van fout-voorspellingen over de partijen (vetgedrukt de twee hoogste percentages) bij de Stemvork 2002.

| | <i>SP</i> | <i>GL</i> | <i>PvdA</i> | <i>D66</i> | <i>CU</i> | <i>CDA</i> | <i>LN</i> | <i>VVD</i> | <i>LPF</i> | <i>Overig</i> | <i>Totaal</i> | <i>Aantal</i> |
|-----|-----------|-----------|-------------|------------|-----------|------------|-----------|------------|------------|---------------|---------------|---------------|
| SP | - | 26 | 25 | 10 | 3 | 14 | 2 | 2 | 14 | 3 | 100 | 125 |
| GL | 32 | - | 16 | 13 | 3 | 10 | 2 | 9 | 12 | 4 | 100 | 101 |
| D66 | 8 | 7 | 8 | - | 1 | 10 | 2 | 23 | 32 | 8 | 100 | 154 |
| CU | 2 | 6 | 0 | 4 | - | 33 | 2 | 6 | 21 | 27 | 100 | 52 |
| LN | 4 | 7 | 11 | 11 | 0 | 16 | - | 27 | 17 | 7 | 100 | 122 |
| LPF | 4 | 3 | 12 | 14 | 2 | 29 | 3 | 31 | - | 4 | 100 | 108 |

Tabel 2B. De verdeling van fout-voorspellingen over de partijen (vetgedrukt de twee hoogste percentages) bij de Stemvork 2003.

| | <i>SP</i> | <i>GL</i> | <i>PvdA</i> | <i>D66</i> | <i>CU</i> | <i>CDA</i> | <i>LN</i> | <i>VVD</i> | <i>LPF</i> | <i>Overig</i> | <i>Totaal</i> | <i>Aantal</i> |
|-----|-----------|-----------|-------------|------------|-----------|------------|-----------|------------|------------|---------------|---------------|---------------|
| SP | - | 20 | 39 | 9 | 1 | 10 | 1 | 8 | 3 | 9 | 100 | 157 |
| GL | 22 | - | 32 | 18 | 0 | 5 | 2 | 9 | 3 | 8 | 100 | 102 |
| D66 | 12 | 6 | 17 | - | 1 | 7 | 4 | 32 | 10 | 10 | 100 | 162 |
| CU | 4 | 2 | 7 | 2 | - | 36 | 4 | 9 | 5 | 31 | 100 | 55 |
| LN | 6 | 6 | 9 | 7 | 0 | 9 | - | 46 | 12 | 5 | 100 | 145 |
| LPF | 3 | 2 | 16 | 9 | 1 | 13 | 3 | 45 | - | 7 | 100 | 258 |

De foutief aan de CU toegedichte stemmen zouden volgens de hierboven geformuleerde verwachting vooral bij het CDA thuis horen, wat inderdaad wordt bevestigd door de verdeling in de beide tabellen (resp. 33 en 36%). De categorie overig bij de CU bevat in Tabel 2A 14 stemmen, waarvan 13 aangeven in werkelijkheid op de SGP te stemmen (in Tabel 2B behoren 15 van de 17 stemmen 'Overig' tot de SGP). Mensen die veel voor het milieu over hebben of sterk voorstander zijn van inkomensnivellering en volgens de Stemvork dus thuishoren bij GroenLinks of SP blijken vaak op de ander of PvdA te stemmen. De SP en GroenLinks zijn dus een soort satellieten die om de PvdA heen draaien. Voorts blijken deelnemers die volgens de Stemvork bij D66 horen op grond van de aangegeven voorkeur voor referenda en gekozen burgemeesters in werkelijkheid in meerderheid LPF of VVD kiezers (in Tabel 2A) of VVD en PvdA kiezers (in Tabel 2B), maar nauwelijks CDA-kiezers. Achter de voorkeursduiding LN, op grond van de keuze dat het met de overheidsbureaucratie wel minder kan, gaan in werkelijkheid vaak VVD of LPF stemmers schuil. De door de Stemvork als LPF gekwalificeerde deelnemers blijken in werkelijkheid vooral VVD en in veel mindere mate CDA of PvdA stemmers.

Beide tabellen laten zien dat waar de Stemvork de deelnemers een politieke voorkeur toeschrijft die niet overeenkomt met de werkelijke voorkeur, desondanks een ruime meerderheid wel juist is geplaatst op het politieke links-rechts spectrum zoals afgebeeld in Figuur 1. Met andere woorden, er is een (onbekend) percentage waarin de Stemvork meteen goed zit; waar de Stemvork een foutieve uitslag geeft behoort niettemin een ruime meerderheid wel tot het verwachte (linkse of rechtse) kamp. In aansluiting op hetgeen in de vorige paragraaf is uiteengezet laten deze empirische gegevens zien dat het zeer goed mogelijk is om aan de hand van een beperkt aantal stellingen respondenten te situeren op een links-rechts spectrum. Als met vijf welgekozen vragen of stellingen de respondenten kunnen worden ingedeeld in links of rechts georiënteerd, dan had daarmee

maximaal een score van 57% goede politieke voorkeursduidingen kunnen worden behaald (wat overeenkomt met de kans p_1 in sectie 3). Om deze score verder te verhogen moet een volgende set vragen proberen te onderscheiden of linkse kiezers extreem of gematigd links zijn en idem dito, met een andere set vragen, voor de rechtse kiezers (het stemprogramma gaat dus interactief te werk: de antwoorden bij de eerste set bepalen welke tweede set wordt voorgelegd). Daarmee kan de score maximaal worden verhoogd tot 69% (gelijk aan p_2).

Hierbij passen wel enige kanttekeningen. Het kan zijn dat de resultaten van de Stemvork niet maatgevend zijn voor wat met inhoudelijke vragen rond bepaalde issues kan worden bereikt als het zo is dat de onderscheiden onderwerpen en stellingen van partijen bij vraag 1 vooral appelleren aan de beeldvorming van partijen waartoe kiezers zich aangetrokken voelen. Dit hoeft niet naadloos aan te sluiten bij de inhoudelijke partijlijn. Ten tweede is het de vraag of de links-rechts tegenstelling volstaat om partijen en kiezers te situeren. Als uit politicologisch onderzoek zou blijken dat voor een adequate situering van kiezers en politieke partijen naast de links-rechts dimensie bijvoorbeeld ook de morele dimensie (met vragen rond euthanasie, homohuwelijk, therapeutisch klonen, genetische manipulatie, etc.) of de tegenstelling liberalisme versus communautarisme van belang is (zie bijvoorbeeld Pellikaan en de Keijser (1998) en Pellikaan (2001)), dan moet het stemprogramma zich daaraan aanpassen door ook hierover stellingen voor te leggen. Op die manier kan de achterliggende filosofie van een stemprogramma nauw aansluiten bij en zich verantwoorden door gedegen politicologisch onderzoek naar de ruimtelijke situering van kiezers en partijen.

4. De structuur van een nieuw stemprogramma

In paragraaf 2 is beargumenteerd dat een stemprogramma interactief moet zijn¹⁷ en dat, aangenomen dat de tegenstelling tussen links en rechts de dominante conflictdimensie is in de politiek, eerst bepaald moet worden of een respondent links of rechts is en vervolgens of ze hierin gematigd of extreem is. Laten we voor de eenvoud van de argumentatie aannemen dat op deze interactieve wijze voor 60% van de deelnemers een juist politieke voorkeursduiding wordt verkregen. Als het stemprogramma wordt voorgelegd aan een representatieve steekproef van het electoraat, dan is dit percentage nog te laag om zo een virtuele verkiezingsuitslag van het stemprogramma te laten concurreren met dat van opiniepeilingen. Om het percentage juiste voorkeursduidingen verder te verhogen moet het standpunt worden losgelaten dat het bij een stemadvies louter om de inhoud van de politiek gaat. Dit standpunt kan men innemen, zoals de *Stemwijzer* doet, maar het is een keuze, geen noodzakelijkheid. De *Stemtest* van het elektronische magazine *Writers Block* bijvoorbeeld begint met te vragen welke coalitie wordt geprefereerd en wie minister-president moet worden en eindigt met een vraag waarbij de respondenten aan partijen een betrouwbaarheidsscore kunnen geven. Bij de laatste verkiezingen zou deze informatie uiterst relevant zijn geweest. Als de aan de hand van de eerste sets vragen als tamelijk links gekwalificeerde kiezers aangeven dat ze graag een coalitie CDA-PvdA willen en Bos als minister-

¹⁷ Afhankelijk van de eerdere antwoorden krijgen respondenten andere vragen voorgeschoteld in plaats van dat iedereen een uniforme lijst vragen moet afwerken.

president, dan is de kans groot dat ze toch PvdA gaan stemmen terwijl als ze dit niet aangeven gewoon GroenLinks of SP zullen stemmen. Een respondent voor wie het twijfelachtig is of deze behoort tot CDA of VVD kan aan de hand van de voorkeuren voor de coalitie (CDA-VVD) en de premier (Zalm) veel beter worden ingedeeld dan wanneer het advies alleen op basis van de politieke inhoud moet worden gegeven. Met andere woorden, met dit soort aanvullende informatie, hoewel het niet behoort tot de standpunten van partijen, kan het percentage juiste voorkeursduidingen verder worden verhoogd. Met hoeveel hangt af van de mate waarin dit soort overwegingen, variërend van strategisch stemmen met het oog op een bepaalde coalitie tot de betrouwbaarheid van lijsttrekkers of partijen, een rol spelen bij het bepalen van de politieke voorkeur.

Een laatste suggestie om de voorspelkracht te verbeteren en in het verlengde daarvan de virtuele verkiezingsuitslag is om rekening te houden met de kennis van politieke zaken van de respondent. Om dit te achterhalen moet een (vierde) set vragen worden toegevoegd. Het punt is dat het te pretentius is van veel stemprogramma's dat zij op basis van zeer beperkte informatie beter de politieke voorkeur zouden kunnen duiden dan de respondent zelf, als deze bijvoorbeeld politieke ontwikkelingen in de media goed volgt of politiek actief is. De respondent die ongeveer weet wat het ideologische profiel is van politieke partijen en goed op de hoogte is welke standpunten zijn partij bij belangrijke issues heeft, zal een afwijkend stemadvies doodleuk naast zich neer leggen, omdat hij meer waarde zal hechten aan het eigen oordeel. De kans dat wat extra informatie, in dit geval het stemadvies, de politieke voorkeur zal wijzigen, is klein (en bovendien, waarschijnlijk kleiner dan het percentage foutieve stemadviezen). Bij respondenten die zichzelf als politieke leken beschouwen kan het stemadvies echter, zeker als het gebaseerd is op een goed uitgebalanceerd programma, de belangrijkste factor zijn bij de partijkeuze: zij weten niet veel van politiek en zullen bijgevolg veel waarde hechten aan het stemadvies. Uiteindelijk betekent dit dat het stemadvies alleen van belang is voor leken in de politiek, hetgeen is weergegeven in onderstaande tabel. Om die reden zijn de Europese parlamentsverkiezingen van 2004, waarbij de kiezers veelal niet weten wat precies de inzet is, een goede gelegenheid om zo een stemprogramma te lanceren.

Tabel 3. De aanbeveling hangt af van de politieke kennis.

| <i>Politieke kennis</i> | <i>Stemadvies vs. eigen voorkeur</i> | <i>Aanbeveling</i> |
|-------------------------|--------------------------------------|--------------------|
| Goed | Eensluidend | Eigen voorkeur |
| | Verschillend | Eigen voorkeur |
| Redelijk | Eensluidend | Eigen voorkeur |
| | Verschillend | Eigen voorkeur |
| Slecht | Eensluidend | Eigen voorkeur |
| | Verschillend | Stemadvies |

Waar het stemadvies en de eigen politieke voorkeur eensluidend zijn is er geen probleem. Bij het optreden van een verschil zou het een stemprogramma sieren als het in de gevallen van politiek redelijk tot goed ingevoerde respondenten zou antwoorden met het stemadvies: u weet voldoende

van politiek, dus volg uw eigen politieke voorkeur.¹⁸ Alleen bij deelnemers die in de vierde set vragen hebben laten blijken weinig kennis te hebben van politiek is een goed stemprogramma in de positie een stemadvies in de letterlijke zin van het woord te geven op basis van de antwoorden op de eerste drie sets vragen. Dit mag op het oog een mager resultaat lijken, maar niet vergeten moet worden dat deze groep toch een tamelijk groot aandeel in het totale electoraat kan hebben. Verwacht mag worden dat de door hen aangegeven politieke voorkeuren bij een opiniepeiling ook veel minder constant en betrouwbaar zijn dan die van de beter geïnformeerde kiezers. Juist bij deze groep kiezers kunnen stemprogramma's hun nut bewijzen, niet alleen voor die kiezers zelf, maar ook als aanvullende informatie op opiniepeilingen. Dit effect treedt des te sterker op naarmate een groter deel van dat deel van het electoraat te rade gaat bij stemprogramma's om hun politieke voorkeur te bepalen.¹⁹ Tenslotte, het ligt voor de hand om issues of beleidsterreinen (ook als ze in de eerste drie sets zijn opgenomen) waar de respondent aangeeft kennis van zaken te hebben extra gewicht te geven in de bepaling van het stemadvies. Het idee hierachter is dat kiezers niet gauw op een partij zullen stemmen die heel andere standpunten heeft op terreinen waar zij zichzelf als expert beschouwen (bijvoorbeeld omdat ze in de gezondheidszorg of het onderwijs werken).

5. Conclusie

Het spreekt voor zich dat er nog een lange weg is te gaan voordat stemprogramma's voldoende betrouwbaar zijn, hetzij om een redelijk percentage juiste stemadviezen te geven, hetzij om het stemgedrag te voorspellen. Hiervoor is met name onderzoek nodig, zowel onder een representatieve steekproef als onder politieke experts, naar de kwaliteit en voorspelkracht van de verschillende stemprogramma's die tot nu toe het licht hebben gezien. Op deze manier kan veel informatie boven water komen over wat welgekozen vragen zijn en welke de plank misslaan. In de bijdrage 'Een kritische evaluatie van de Stemwijzer 2002' (Groot 2003) zijn een aantal criteria zoals validiteitseisen, robuustheid en discriminerend vermogen naar voren gebracht ter waarborging van de kwaliteit van stemprogramma's. De resultaten van het onderzoek onder politieke experts, tezamen met de in acht te nemen criteria en de in dit artikel voorgestelde getrapte, interactieve structuur van een stemprogramma zou mogelijk kunnen leiden tot een voldoende betrouwbaar stemprogramma in de toekomst (de eerstvolgende testcase is de Europese parlamentsverkiezingen op 10 juni 2004). Als die kritische grens wordt bereikt, dan zal het in de eerste plaats minder vaak voorkomen dat het stemadvies in strijd is met de eigen politieke voorkeur van redelijk tot goed ingeburgerde respondenten in de politiek. In de tweede plaats kunnen leken meer waarde hechten aan het stemadvies, omdat het niet alleen is gebaseerd op de politieke inhoud, maar ook omdat andere aspecten worden meegenomen. Dan treedt een zelfversterkend proces in werking: goed geïnformeerde kiezers zien hun voorkeur bevestigd, slecht geïnformeerde kiezers zullen steeds meer waarde hechten aan het advies. Het denkbeeldige eindpunt van dit proces is dat

¹⁸ Dit laat onverlet dat deze groep natuurlijk wel kennis mag nemen van het stemadvies dat resulteert uit hun antwoorden.

¹⁹ Ter indicatie: In 1998 verstreekte de Stemwijzer ruim 12 duizend adviezen; bij de verkiezingen in 2002 en in 2003 zijn bij elk meer dan twee miljoen adviezen gegeven, wat overeenkomt met een kwart van het electoraat.

meedoen aan een stemprogramma het stemmen zelf is, met andere woorden dat de uitslag van het stemprogramma, het stemadvies, tegelijk de stem is. Dan heeft de niet langer virtuele verkiezingsuitslag van het stemprogramma ook de opiniepeiling naar de kroon gestoken.

Literatuur

Brug, W. van der, 1997, *Where's the Party?: Voters' Perceptions of Party Positions*, Proefschrift Universiteit van Amsterdam.

Groot, L.F.M., 2003, Een kritische evaluatie van de StemWijzer 2002, *b en m: tijdschrift voor politiek, beleid en maatschappij*, jg. 30, nr. 1, 20-30.

Holsteyn, J. van, 1994, *Het woord is aan de kiezer: een beschouwing over verkiezingen en stemgedrag aan de hand van open vragen*, Proefschrift Universiteit Leiden: DSWO Press.

Pellikaan, H. en J. de Keijser, 1998, 'Kamerleden in de publieke ruimte'. In Bovens et al., *Nieuwe tegenstellingen in de Nederlandse politiek*, Boom, 165-186.

Pellikaan, H., 2001, *De noodzaak van een meetlat voor ruimtelijke modellen*, paper gepresenteerd op de Nederlandse Kring voor Wetenschap der Politiek.