



Universiteit
Leiden
The Netherlands

Kiezersopkomst in Nederland- Een empirisch onderzoek

Jaarsma, B.; Schram, A.; Winden, F. van; Linssen, G.

Citation

Jaarsma, B., Schram, A., Winden, F. van, & Linssen, G. (1986). Kiezersopkomst in Nederland-: Een empirisch onderzoek. *Acta Politica*, 21: 1986(1), 39-55. Retrieved from <https://hdl.handle.net/1887/3452627>

Version: Publisher's Version

License: [Leiden University Non-exclusive license](#)

Downloaded from: <https://hdl.handle.net/1887/3452627>

Note: To cite this publication please use the final published version (if applicable).

Onderzoek

Kiezersopkomst in Nederland

Een empirisch onderzoek*

Bert Jaarsma, Arthur Schram, Frans van Winden en
Guido Linssen

1. Inleiding

Er is de laatste decennia zowel bij politicologen als bij economen een groeiende interesse in het verklaren van stemgedrag met behulp van meer of minder geavanceerde onderzoeksmethoden. Deze interesse betreft onder meer het verklaren van opkomstgedrag (het al dan niet verschijnen bij de stembus) met behulp van economische, demografische en andere categoriale variabelen. Om een indruk te geven van het onderzoek dienaangaande, zullen wij beginnen met een overzicht van een viertal buitenlandse en twee Nederlandse studies die zich met dit probleem beziggehouden hebben. Na vermelding van de conclusies en onze kritiek op deze studies komen we op ons eigen onderzoek, dat in de volgende paragrafen gepresenteerd zal worden.

Als eerste studie noemen we het artikel van Frey (1971), waarin een positief verband wordt gelegd tussen opkomst en inkomensniveau. Wij zullen later zien, dat het zich beperken tot één verklarende variabele – hier het inkomen – een ernstig tekort is, dat de werkelijke verbanden kan verduisteren.

Volgens Bennet en Orzechowski (1983) gaan ambtenaren eerder naar de stembus dan de rest van de bevolking, ook als voor een aantal andere variabelen gecorrigeerd wordt. Zij corrigeren echter slechts voor één of twee variabelen tegelijk, zodat ook hun conclusies met de nodige voorzichtigheid betracht dienen te worden. Slechts een multivariate analyse, waarin alle verklarende variabelen tegelijkertijd een rol kunnen spelen, kan het afzonderlijke effect van een variabele, gegeven de waarde van de overige variabelen, aan de oppervlakte brengen.

Dit wordt mooi aangetoond door Wolfinger en Rosenstone (1980), die een multivariaat (probit-)model presenteren, ter verklaring van opkomstpercentages in de Verenigde Staten. Zij concluderen dat opleiding en leeftijd de belangrijkste verklarende variabelen zijn, terwijl ook het ambtenaar-zijn een (positieve) invloed heeft. Frey en Pommerehne (1982) con-

cluderen dit laatste ook, en voeren er theoretische argumenten voor aan.

Het grootste gebrek in het werk van Wolfinger en Rosenstone, schuilt in het feit dat ze slechts één verkiezing onderzoeken. Het moge duidelijk zijn, dat een dergelijke momentopname een vertekend beeld van de werkelijkheid kan geven – ons onderzoek zal dit ook illustreren – en dat de voorkeur gegeven moet worden aan een serie van verkiezingen, als de data aanwezig zijn.

Wolfinger en Rosenstone wijzen wel terecht op het feit dat een lineair model in het algemeen niet geschikt is om opkomstgedrag te verklaren. De voorkeur dient gegeven te worden aan een zogeheten logit- of probit-model, en wel om een aantal redenen. Op de eerste plaats is de afhankelijke variabele, het al of niet gaan stemmen, dichotoom voor ieder individu. Men kan dan de kans om te gaan stemmen proberen te verklaren. In een lineair model hoeft de geschatte kans echter niet tussen 0 en 1 te liggen, bij een logit- of probit-specificatie is dit wel het geval. Op de tweede plaats is het effect van een variabele op de opkomstkans voor ieder individu gelijk bij een lineair model. Bij een logit(probit-)model is dit effect het grootst voor diegenen die buiten deze variabele om een opkomstkans van 50% hebben, en het kleinst voor mensen die anders zeker wel of zeker niet zouden zijn gaan stemmen. Deze eigenschap van logit- en probit-modellen lijkt meer in overeenstemming met de realiteit dan het constante effect in geval van een lineair model¹. Zolang de data niet geconcentreerd zijn in extreme waarden, blijft de keuze tussen logit- en probit-modellen tamelijk arbitrair. Wie meer te weten wil komen over deze modellen, wordt verwezen naar Amemiya (1981) en Maddala (1983).

Laten wij nu eens gaan kijken naar de Nederlandse stand van zaken. Op de eerste plaats dient dan opgemerkt te worden dat de nationale kiezersonderzoeken de onderzoeker inmiddels sinds 1971 voorzien van een aangename hoeveelheid goed bruikbare data. In *Acta Politica* 1983/2 ('In het spoor van de kiezer') worden enkele studies gepresenteerd die op deze data gebaseerd zijn. Twee ervan gaan in op de verklaring van opkomstgedrag met behulp van demografische variabelen, te weten Schmidt (p. 139-155) en Rosenthal (p. 229-241).

Schmidt vindt een positief verband tussen opkomstgedrag en achtereenvolgens inkomen, leeftijd en frequentie van kerkbezoek. Hij constateert een (gering) negatief verband met urbanisatiegraad en geen verband met opleiding en geslacht. Verder vindt hij een positieve invloed van een aantal variabelen die indicatoren zijn van politieke belangstelling. Naast afzonderlijke kruistabellen voor deze variabelen, presenteert Schmidt ook een multivariaat lineair model. Alles past hij afzonderlijk toe op de verkiezingsjaren 1971, 1972, 1977, 1981 en 1982.

Wij zullen hier niet verder ingaan op het stuk van Schmidt. Vermeld wordt slechts dat hij als afhankelijke variabele niet de opkomst, doch een zelfgeconstrueerde stemindex neemt, met als argument dat de helft van de niet-stemmers om min of meer toevallige redenen thuisgebleven zijn, en daarom niet zo zinvol als niet-stemmers te kenmerken zouden zijn. Dit probleem verdwijnt echter indien men, zoals bij een logit- of probit-model, met opkomstkans werkt. Zolang iemand niet met een kans van 1 gaat stemmen (en dit zal nooit voorkomen), zijn de 'toevallige thuisblijvers' in het model opgenomen. Waar het om gaat is dat sommige individuen een grotere opkomstkans hebben dan anderen.

Rosenthal presenteert in zijn paragraaf over opkomst een analyse met betrekking tot één jaar. Hij beschouwt het verschil in opkomstgedrag tussen actieven en niet-actieven en maakt ook onderscheid tussen een hoge en een lage inkomenspositie. Zijn conclusie is dat ambtenaren eerder gaan stemmen dan werklozen, met name werklozen met lage inkomens.

De analyse van Rosenthal is statistisch nauwelijks onderbouwd. Hij gaat niet verder dan het presenteren van enkele opkomstpercentages. Bovendien baseert hij zich slechts op één jaar, hetgeen zoals gezegd een vertekend beeld kan geven. Overigens betreft hij bij zijn analyse van partijkeuze wel andere verkiezingsjaren in zijn onderzoek. Tot slot vinden wij een andere indeling in sociale groepen wenselijk dan die van Rosenthal. Deze indeling zullen we later presenteren.

Uit het bovenstaande moge duidelijk blijken, dat wij een multivariaat logit-(of probit-)model wenselijk achten voor een onderzoek naar opkomstgedrag. Daarbij dient bij voorkeur gebruik gemaakt te worden van data betreffende een aantal verkiezingen. Geïnspireerd door de conclusies van andere auteurs, en op grond van een eerdere eigen analyse (waaronder een theoretische analyse in Van Winden (1983)), kozen wij voor de variabelen inkomen, opleiding, sociaal-economische status (objectief, subjectief), leeftijd, frequentie van kerkbezoek en organisatorische status. Deze laatste variabele, die door ons gebruikt wordt als een indicator van politieke belangstelling, geeft aan of de respondent al dan niet lid is van een vakbond of politieke partij².

In de volgende paragraaf zullen de variabelen nader gespecificeerd worden, en zullen de eerste resultaten gepresenteerd worden. In paragraaf 3 volgen enkele verdere analyses. Het artikel besluit met een aantal conclusies in paragraaf 4.

2. Presentatie van het model en de eerste resultaten

In een logit-model wordt de kans y dat een individu zal gaan stemmen geacht op de volgende wijze een functie te zijn van een vector van onafhankelijke variabelen X :

$$y = F(X\beta) = 1/(1 + e^{-X\beta})$$

met

- y : de kans dat een individu zal gaan stemmen;
- X : een vector van onafhankelijke variabelen;
- β : een vector van onbekende coëfficiënten;
- $F(\cdot)$: de logistische verdelingsfunctie.

Er zij opgemerkt dat vergelijking (1) afgeleid kan worden met behulp van een zogenaamd 'random utility model', waarbij verondersteld wordt dat een individu zal gaan stemmen indien hij/zij een hoger nut toekent aan het uitbrengen van een stem dan aan het zich onthouden van stemmen (zie Amemiya (1981)).

De vector X bestaat uit een constante term en de in de vorige paragraaf genoemde variabelen. Deze variabelen worden omgezet in dummy-variabelen, die een categoriale indeling per variabele weergeven. Per variabele kunnen dan een aantal dummies (één voor iedere categorie van een variabele) onderscheiden worden. Als deze dummies in de logit-vergelijking (1) opgenomen zouden worden, zou dit echter tot multicollineariteit leiden, omdat ieder individu altijd tot precies één van de categorieën behoort. Daarom zal ieder individu steeds een 1 hebben voor één dummy en een 0 voor de overige dummies met betrekking tot een variabele. De som over alle dummies zal dan voor iedereen gelijk zijn aan 1 (zie voor dit alles bijvoorbeeld Amemiya (1981)). Om deze multicollineariteit te ontlopen, wordt per variabele één dummy uit de regressievergelijking weggelaten. Deze weggelaten categorieën worden gerepresenteerd door de constante term, iemand die tot een dergelijke weggelaten categorie behoort heeft een 0 voor de dummies inzake de betreffende variabele in de regressievergelijking.

Opgemerkt wordt nog dat bij de sociaal-economische status onderscheid wordt gemaakt tussen de sociale klasse waartoe de respondent meent te behoren (subjectieve sociale klasse), en de categorie waarin de respondent objectief ingedeeld kan worden (objectieve sociale klasse). De gekozen indeling is als volgt, waarbij de weggelaten dummy met een * is aangegeven.

Inkomen

1. nul-mimumum*
2. minimum-modaal
3. modaal-2 x modaal
4. meer dan 2 x modaal

Subjectieve sociale klasse

1. hogere (midden)klasse*
2. middenklasse
3. hogere arbeidersklasse
4. arbeidersklasse

Frequentie van kerkbezoek

1. geen godsdienst*
2. nooit
3. een paar keer per jaar
4. alleen op hoogtijdagen
5. elke week

Opleiding

1. alleen lager onderwijs*
2. enige jaren voortgezet onderwijs
3. middelbaar onderwijs
4. hoger onderwijs

Objectieve sociale klasse

1. ambtenaren*
2. particuliere werknemers
3. zelfstandigen
4. zonder beroep (huisvrouwen)
5. werklozen³
6. gepensioneerden/WAO'ers

Er wordt een dummy-variabele betreffende organisatorische status toegevoegd, die 1 is als de respondent lid is van een politieke partij of een vakbond, en anders 0. De variabele leeftijd wordt niet gecategoriseerd. De constante term vertegenwoordigt nu dus een individu die in de categorieën 1 van alle variabelen is ingedeeld.

De vector X in de regressievergelijking (1) bestaat uit 20 elementen, te weten: een 1 voor de constante term, drie inkomens-, opleidings- en subjectieve sociale categorie-dummies, vier objectieve sociale categorie- en religie-dummies, een dummy betreffende organisatorische status, en de leeftijd. Op de variabele leeftijd na, betreft dit dus steeds dichotome (0, 1) variabelen.

De data die wij gebruikt hebben om vergelijking (1) te schatten, zijn afkomstig van de Nederlandse nationale kiezersonderzoeken 1971, 1972, 1977, 1981 en 1982.

Een opmerking dient nog gemaakt te worden over het feit dat de opkomst zoals die in de steekproeven naar voren komt steeds hoger is dan de werkelijke opkomst. Dit verschijnsel doet zich veelal voor bij verkiezingsonderzoeken. Onderzoeken van Katosh en Traugott (1981) en Sigelman (1982) suggereren dat dit de werkelijkheidswaarde van de te trekken conclusies niet aantast. Ook Schmidt (1983, p. 141) vindt in zijn eerder genoemd stuk aanwijzingen in die richting.

Tabel 1 geeft de 'maximum likelihood' schattingen voor de coëfficiënten β in (1) weer, voor 1977, 1981 en voor de geaggregeerde data van de vijf

Tabel 1: De geschatte logit-vergelijkingen

Jaar	1977	1981	geaggregeerd
Waarnemingen	729	1151	5253
Constante	4.157 (1.718)xx	1.804 (0.957)x	0.756 (0.398)x
<i>Inkomen</i>			
minimum-modaal	-0.238 (0.397)	0.308 (0.301)	0.182 (0.126)
modaal-2 x modaal	-0.280 (0.464)	1.076 (0.407)xx	0.438 (0.144)xx
meer dan 2 x modaal	-0.119 (0.886)	0.926 (0.543)x	0.509 (0.244)xx
<i>Opleiding</i>			
voortgezet	0.069 (0.351)	-0.397 (0.396)	0.230 (0.126)x
middelbaar	-0.175 (0.829)	-0.862 (0.514)x	0.435 (0.204)xx
hoger	0.558 (0.836)	-0.204 (0.595)	0.580 (0.243)xx
<i>Subj. Soc. Cat.</i>			
middenklasse	-1.125 (1.101)	-0.250 (0.533)	-0.397 (0.239)x
hogere arbeider	-0.629 (1.195)	0.514 (0.709)	-0.651 (0.260)xx
lagere arbeider	-1.681 (1.134)	-0.643 (0.582)	-0.900 (0.269)xx
<i>Obj. Soc. Cat.</i>			
werknemer	-1.261 (1.058)	0.359 (0.396)	-0.411 (0.187)xx
zelfstandige	-1.734 (1.209)	0.909 (0.806)	-0.543 (0.269)xx
zonder beroep	-0.888 (1.065)	0.613 (0.431)	-0.173 (0.196)
gepens./WAO	-0.933 (1.66)	1.155 (0.684)x	-0.419 (0.257)
<i>Frequ. kerkbezoek</i>			
nooit	-1.177 (0.433)xx	0.094 (0.335)	-0.284 (0.137)xx
paar keer/jaar	-0.601 (0.436)	0.101 (0.355)	-0.124 (0.143)
hoogtijdagen	0.079 (0.526)	0.327 (0.422)	0.171 (0.167)
wekelijks	-0.121 (0.455)	0.919 (0.401)xx	0.551 (0.137)xx
Org. Status	1.041 (0.400)xx	0.494 (0.303)	0.800 (0.123)xx
Leeftijd/100	1.933 (1.123)x	0.580 (0.998)	2.804 (0.397)xx
Pseudo-R ²	0.13	0.09	
<i>Jaar</i>			
1972			0.224 (0.136)x
1977			0.838 (0.181)xx
1981			0.948 (0.164)xx
1982			0.279 (0.147)x
Pseudo-R ²			0.11

In tabel 1 staan per kolom de schattingen voor de regressiecoëfficiënten β van vergelijking (1). Tussen haakjes staan de standaardafwijkingen van deze coëfficiënten vermeld. Een 'x' geeft aan dat de betreffende coëfficiënt significant verschilt van 0 bij een onbetrouwbaarheidsniveau van 10%, een 'xx' geeft hetzelfde aan bij 5%.

jaren. De resultaten voor de overige jaren kunnen desgewenst van de auteurs verkregen worden. Op het samenvoegen van de data komen we later nog terug. Laten we eerst op de wijze van Schmidt kijken naar de resultaten voor 1977 en 1981.

Een vergelijking van de coëfficiënten voor 1977 en 1981 laat zien dat de verschillen per jaar groot kunnen zijn. Op de constante term na, is er geen variabele die in beide jaren een significante coëfficiënt heeft.

In 1977 is de coëfficiënt voor godsdienstigen die niet naar de kerk gaan significant negatief (op 5%-niveau), waaruit geconcludeerd kan worden dat deze groep een lagere opkomstkans heeft dan de basis-(weggelaten-) categorie, in dit geval degenen zonder godsdienst. Verder zijn in 1977 alleen de coëfficiënten behorende bij organisatorische status en leeftijd significant.

Voor 1981 zijn de resultaten geheel anders. Nu zijn er alleen significante coëfficiënten te vinden bij de dummies behorende bij de twee hoogste inkomenscategorieën, die met betrekking tot het middelbaar onderwijs⁴, de sociaal-economische categorie van gepensioneerden en WAO'ers, en de wekelijkse kerkgangers.

Een vergelijking van alle (dus ook de niet-significante) coëfficiënten in 1977 en 1981 leert dat 13 van de 20 coëfficiënten van teken verschillen voor deze twee verkiezingen.

Op dezelfde wijze hebben wij ook de resultaten voor de overige verkiezingsjaren bekeken. Alleen de voornaamste conclusies worden vermeld. Op de eerste plaats is er niet één categorie van een variabele die voor alle vijf jaren significant (zelfs op 10%-niveau) van nul verschilt⁵. Nemen we als (tamelijk willekeurig) criterium, dat dit tenminste in drie van de vijf jaren het geval moet zijn, dan voldoen (behalve de constante term) de coëfficiënten betreffende de gepensioneerden, de wekelijkse kerkgangers, de organisatorische status en leeftijd hieraan. De coëfficiënt voor gepensioneerden en WAO'ers is significant negatief in 1972 en 1982 (hetgeen betekent dat de kans dat een gepensioneerde zal gaan stemmen kleiner is dan voor een ambtenaar met dezelfde indeling bij de overige variabelen). Deze coëfficiënt is echter significant positief in 1981 (hetgeen het omgekeerde betekent uiteraard).

Deze resultaten laten zien dat conclusies die op één jaargang gebaseerd zijn, sterk afhangen van de keuze van het jaar. Per jaargang worden overigens maar weinig variabelen gevonden, die een significante invloed op de opkomstkans hebben. Als we nogmaals als criterium nemen het minstens drie keer voorkomen van een (op tenminste 10%-niveau) significante coëfficiënt, dan kunnen slechts de volgende (voorzichtige) conclusies getrokken worden: (1) wekelijkse kerkgangers gaan eerder stemmen

dan mensen zonder godsdienst; (2) de opkomstkans stijgt bij het ouder worden; (3) in vakbonden of politieke partijen georganiseerde kiesgerechtigden gaan eerder naar de stembus dan zij die niet georganiseerd zijn.

Tabel 1 geeft slechts de relaties weer tussen de categorieën die opgenomen zijn in de constante term en iedere andere categorie afzonderlijk. Daaruit kunnen echter ook de verschillen tussen ieder paar categorieën van een variabele, en de significantie daarvan, afgeleid worden. De resultaten hiervan voor de afzonderlijke jaren zullen we hier niet vermelden, omdat er evenals de reeds gepresenteerde cijfers, weinig positieve conclusies uit getrokken kunnen worden. De enige duidelijke conclusie ondersteunt de uitzonderlijke positie van de wekelijkse kerkgangers⁶.

Hoewel er per jaargang weinig significante coëfficiënten voorkomen, is er een aantal coëfficiënten dat elk jaar hetzelfde teken heeft, zij het dat significantie niet (altijd) optreedt. Dit is een aanwijzing dat het aggregeren van de data tot duidelijkere resultaten kan leiden. Dit is dan ook uitgetoetst, waarbij specifieke (macro-economische of politieke) jaareffecten worden opgevangen door een dummy-variabele toe te voegen met betrekking tot het jaar waarop de waarneming betrekking heeft. Het jaar 1971 wordt opgenomen in de constante term.

De impliciete veronderstelling hierbij betreffende het constant zijn van de coëfficiënten is door Foster (1984) voor de Verenigde Staten met behulp van een lineair model getoetst, en niet verworpen. Simpele 'likelihood-ratio' toetsen voor de Nederlandse data wijzen uit dat de stabiliteit van alle coëfficiënten, inclusief de constante term verworpen wordt (overschrijdingskans $\pm 0,005$). De hypothese dat alle coëfficiënten behalve de constante term stabiel zijn, wordt echter niet verworpen (overschrijdingskans $\pm 0,15$). Dit rechtvaardigt de gekozen werkwijze, waarbij de constante term per jaar verschilt door de opgenomen jaar-dummies die de specifieke jaareffecten weergeven.

Een vluchtige blik op de laatste kolom van tabel 1 is al voldoende om te zien dat onze conclusies nu geheel anders worden. Bij elke gecategoriseerde variabele zijn er nu minstens twee categorieën met significante coëfficiënten. Ook hier hebben we de verschillen tussen de categorieën van elke variabele berekend. Alle verschillen die statistisch significant bleken bij een onbetrouwbaarheidsniveau van respectievelijk 10% en 5%, worden vermeld in tabel 2.

Deze resultaten duiden erop dat de lagere twee inkomenscategorieën minder snel naar de stembus gaan dan de hogere inkomensgroepen (er zij opgemerkt dat het verschil tussen de tweede en vierde categorie (10%-) significant van nul verschilt, als er eenzijdig getoetst wordt). Wat opleiding betreft is de belangrijkste conclusie dat de laagste categorie (alleen

Tabel 2: Significante verschillen tussen dummy-variabelen

	5%	10%
Inkomen	1/3, 1/4, 2/3	—
Opleiding	1/3, 1/4	1/2, 2/4
Subjectieve		
Sociale Categorie	1/3, 1/4, 2/3, 2/4	1/2, 3/4
Objectieve		
Sociale Categorie	1/2, 1/3, 2/4	3/4
Frequentie van	1/2, 1/5, 2/4, 2/5,	—
Kerkbezoek	3/5, 4/5	

De nummering van de categorieën is overeenkomstig het schema op p. 43. Per variabele staan in deze tabel de dummy-variabelen (paarsgewijs) weergegeven die coëfficiënten hebben die significant van elkaar verschillen. Bij de variabele opleiding bijvoorbeeld, verschilt de coëfficiënt behorende bij de groep 1 (alleen lagere school) op 5%-niveau significant van de groepen 3 en 4 (respectievelijk middelbaar en hoger onderwijs). Op 10%-niveau verschilt deze coëfficiënt significant van die behorende bij groep 2, zij die enige jaren voortgezet onderwijs genoten hebben. Op dit niveau verschillen ook de coëfficiënten behorende bij de groepen 2 en 4 significant. Een significant verschil in coëfficiënten wordt geïnterpreteerd als een significant verschil in opkomstkans tussen de betreffende groepen.

lager onderwijs) minder snel naar de stembus gaat dan de overige. Dit verschijnsel doet zich niet meer voor zo gauw er enkele jaren voortgezet onderwijs gevolgd zijn.

De variabele subjectieve sociaal-economische klasse vertoont significante verschillen voor elk mogelijke paar categorieën. Hoe lager de klasse waartoe men meent te behoren, des te kleiner is de kans dat men zal gaan stemmen. Bij de objectieve sociaal-economische klasse zien we dat ambtenaren eerder gaan stemmen dan mensen in de particuliere sector. De categorieën met betrekking tot mensen zonder beroep/huisvrouwen of -mannen en gepensioneerden en WAO'ers – waarvoor kenmerkend is dat zij mensen betreffen die buiten het (formele) productieproces staan – nemen tussenposities in. Terloops merken we hier op dat de numerieke betekenis van laatstgenoemde groep naar verwachting in de komende vijftig jaren aanzienlijk zal toenemen, gezien de verwachte vergrijzing van de bevolking.

Wat betreft de frequentie van kerkbezoek wordt bevestigd, dat wekelijkse kerkgangers de grootste opkomstkans hebben. Verder is het zo dat mensen die nooit naar de kerk gaan eerder gaan stemmen als ze niet

godsdienstig zijn, dan als ze dit wel zijn. Men zou kunnen veronderstellen, dat de eerste groep een meer wereldse houding aanneemt.

Verder bevestigen de geaggregeerde data de belangrijke rol van leeftijd en organisatorische status. Tot slot zijn alle jaren significant verschillend van elkaar wat betreft opkomst, behalve 1972/1982 en 1977/1981.

Er zij op gewezen, dat deze conclusies alleen dan gelden, wanneer de overige variabelen constant gehouden worden. In de volgende paragraaf zal nog wat nader worden ingegaan op de schattingen voor de geaggregeerde data.

3. Marginale effecten

Met behulp van de coëfficiënten in tabel 1 kunnen de marginale effecten onderzocht worden van veranderingen in de verklarende variabelen. Omdat de verklarende variabelen hier (op leeftijd na) dummies zijn, houdt een verandering een (denkbeeldige) overgang naar de betreffende groep, vanuit de basisgroep in. In een lineair model is het effect hiervan gelijk aan de geschatte coëfficiënt voor deze categorie. Dit geldt niet voor de logit-specificatie (vergelijking (1)), waar het effect afhankelijk is van de waarde van de overige variabelen (zie paragraaf 1). Wij hebben deze marginale effecten als volgt bepaald (zie bijvoorbeeld ook Wolfinger en Rosenstone (1980)): voor ieder individu in categorie *i* van een of andere variabele werd de geschatte opkomstkans bepaald. Deze kans werd opnieuw geschat onder de veronderstelling dat het individu tot de basiscategorie (aangegeven met een * in bovenstaand overzicht van de variabelen) in plaats van *i* behoorde, en het verschil tussen deze twee kansen werd bepaald. Het gemiddelde verschil over alle individuen in groep *i* wordt gepresenteerd als het marginale effect van de overgang van de basiscategorie naar *i*. Deze marginale effecten zijn zowel voor de gehele steekproef als voor iedere categorie van een andere variabele bepaald. Een aantal resultaten wordt vermeld in tabel 3, waarin de categorieën met een coëfficiënt die niet significant van 0 afwijkt, weggelaten zijn.

Het effect met betrekking tot opleiding en inkomen is van dezelfde orde van grootte. Bij opleiding doet het sterkste effect zich voor bij de categorie betreffende middelbaar onderwijs, bij inkomen is dit het geval bij degenen wier inkomen tussen modaal en 2 x modaal ligt. De hoogste twee categorieën verschillen bij geen van beide variabelen echter significant (zie tabel 2). De conclusie van Schmidt (1983) dat opleiding geen effect heeft op opkomstgedrag moet derhalve verworpen worden. Wel ondersteunen de uitkomsten zijn conclusie met betrekking tot inkomen (in tegenstelling tot

de conclusie van Wolfinger en Rosenstone). Het effect van opleiding neemt af bij stijging van inkomen en vice versa. Geconcludeerd kan worden dat het bereiken en handhaven van een hoge inkomenspositie gedeeltelijke compensatie biedt voor gebrek aan opleiding, en andersom. Hierop komen we nog terug.

Als we de marginale effecten van objectieve versus subjectieve sociale klasse (tabel 3b) bekijken, zien we dat deze effecten groter zijn in de particuliere sector (particuliere werknemers en zelfstandigen) dan in de publieke sector (ambtenaren, gepensioneerden en WAO'ers), met één uitzondering, doch dit betreft wel een cel met slechts 12 waarnemingen.

Tabel 3: Marginale effecten

Het getal linksboven geeft de gemiddelde verandering in opkomstkans weer als individuen van de basisgroep naar de betreffende groep zouden verhuizen (zie de hoofdtekst). Rechtsboven staat steeds de standaardafwijking van deze verschillen voor de betreffende cel en het derde getal is het aantal individuen in de betreffende cel. Er zijn bijvoorbeeld 284 individuen met middelbare opleiding en een inkomen tussen modaal en 2 x modaal. De gemiddelde afname in opkomstkans voor deze individuen, als ze verondersteld worden tot de laagste inkomensklasse te behoren, bedraagt 3,2-procentpunten, zodat de gemiddelde stijging bij een verandering van de laagste naar de derde inkomensgroep eveneens 3,2-procentpunten bedraagt voor deze groep. De standaardafwijking van deze kansveranderingen is in dit geval 1.9.

3a INKOMEN	OPLEIDING						INKOMEN			
	voort- gezet		middel- baar		hoger		modaal- 2 x mod.		meer dan 2 x mod.	
nul- minimum	2.7	1.4	4.9	2.6	5.3	2.5	4.5	2.2	4.3	2.1
	502		71		54	lager	262		31	
minimum- modaal	2.5	1.2	4.2	2.3	4.3	2.4	voort- gezet	4.3	2.2	4.0
	956		138		117	gezet	889		178	
modaal- 2x modaal	2.1	1.1	3.2	1.8	3.1	2.1	middel- baar	3.2	1.9	3.4
	889		284		290	baar	284		111	
meer dan 2 x modaal	1.6	1.0	2.8	1.5	2.3	1.5	hoger	2.2	1.5	1.9
	178		111		203		290		203	
totaal	2.3	1.2	3.6	2.1	3.2	2.2	totaal	3.8	2.2	3.1
	2525		604		664		1725		523	

SOCIALE KLASSE	SUBJECTIEF					
	Midden- klasse		hogere arbeider		lagere arbeider	
OBJECTIEF						
ambtenaar	-1.8	1.2	-3.9	2.4	-6.3	3.2
	387		110		61	
werknemer	-3.0	1.6	-6.8	3.2	-8.3	3.8
	692		521		316	
zelfstandige	-2.7	1.7	-6.6	3.4	-4.6	2.1
	211		29		12	
zonder beroep	-2.5	1.4	-5.7	2.8	-6.9	3.7
	844		473		453	
gepensioneerd/WAO	-1.7	1.1	-4.1	2.0	-5.9	3.3
	215		164		180	
Totaal	-2.5	1.5	-5.8	3.1	-7.1	3.7
	2349		1297		1022	

De positie van degenen zonder beroep zit er wederom tussenin. Het blijkt dat ambtenaren, gepensioneerden en WAO'ers minder gevoelig zijn voor hun subjectieve sociale positie dan particuliere werknemers en zelfstandigen. Wij wijzen er voor alle zekerheid nogmaals op, dat deze conclusies alleen gelden voor de geaggregeerde data. Overeenkomstige conclusies bleken voor afzonderlijke jaren nauwelijks te trekken (hetgeen de betrekkelijke waarde van de analyse van Rosenthal aangeeft).

Tot slot toont tabel 3c de marginale effecten betreffende ambtenaren en particuliere sector, onderverdeeld naar verschillende leeftijdscategorieën. De leeftijd is als volgt gecategoriseerd: (1) 18-29 (1971: 21-29); (2) 30-39; (3) 40-49; (4) 50-59; (5) 60-69; (6) ouder dan 70 jaar. Wij concluderen dat het effect sterk afneemt bij stijging van de leeftijd. Dit verminderde effect van de sociale categorie is toe te rekenen aan verhoogde levenservaring en/of aan generatie-effecten.

Natuurlijk zou nog een groot aantal van dergelijke tabellen met marginale effecten gepresenteerd kunnen worden. Wij beperken ons echter tot het vermelden van de belangrijkste conclusies.

Op de eerste plaats is er een duidelijke invloed van opleiding, inkomen

leeftijd	OBJECTIEVE SOCIALE KLASSE			
	werknemer		zelfstandige	
18-29	-4.8	2.0	-5.7	2.8
	585		46	
30-39	-3.6	2.0	-3.5	1.8
	527		84	
40-49	-2.9	1.7	-3.3	2.3
	326		83	
50-59	-2.3	1.4	-2.6	1.6
	204		58	
60-69	-2.1	1.3	-2.1	1.2
	55		39	
ouder dan 70	-1.5	1.7	-1.4	0.7
	4		6	
Totaal	-3.7	2.0	-3.4	2.3
	1701		316	

en leeftijd. Deze variabelen lijken een soort levenservaring (c.q. generatie-effect) weer te geven. Een hoge ervaring op grond van een dezer variabelen vermindert het effect van de andere variabelen, uitgezonderd het effect van de subjectieve sociaal-economische klasse, dat belangrijk blijkt te blijven.

Ook vinden wij een belangrijk verschil tussen particuliere sector en publieke sector.

Frequentie van kerkbezoek blijft eveneens een belangrijke variabele, waarvan de betekenis slechts weinig afneemt bij toenemende ervaring via een der bovengenoemde variabelen.

Overigens zij opgemerkt, dat de gevonden verschillen weliswaar significant, doch niet uitzonderlijk hoog zijn. Zelfs de groepen met de laagste opkomstchansen hebben nog een aanzienlijke waarschijnlijkheid om te

gaan stemmen. Dit in tegenstelling tot de Verenigde Staten, bijvoorbeeld, waar verschillen tot 50-procentpunten voorkomen (zie Wolfinger en Rosenstone (1980)). Aan de verklaring hiervan beoogt dit stuk geen bijdrage te leveren.

4. Conclusies

Belangrijkste doel van het onderhavige onderzoek is geweest, om met behulp van een statistisch geëigend model het effect van een aantal in de literatuur (afzonderlijk of in combinatie) gehanteerde categoriale variabelen op het opkomstgedrag van Nederlanders te onderzoeken. Met behulp van vijf achtereenvolgende nationale kiezersonderzoeken is een logit-specificatie van dit opkomstgedrag geschat. Dit schatten is zowel gebeurd voor de jaren afzonderlijk als voor de geaggregeerde data. Wij hebben onze conclusies vooral vergeleken met de resultaten van Schmidt (1983) en Rosenthal (1983), die dezelfde problematiek onderzocht hebben. Het onderhavige stuk kan gezien worden als een replicatie van de studie van Wolfinger en Rosenstone (1980), toegepast op Nederland. Aangetoond is echter dat het zich beperken tot één verkiezing, zoals Wolfinger en Rosenstone gedaan hebben, vertekenend kan werken.

Onze conclusies hangen sterk af van het al dan niet gebruiken van geaggregeerde data. Worden de verkiezingen afzonderlijk bekeken, dan vinden we nauwelijks variabelen die de opkomst helpen verklaren. Bij aggregatie daarentegen vinden we een significante invloed van opleiding, inkomen, leeftijd, sociaal-economische status (subjectief en objectief), frequentie van kerkbezoek en organisatorische status. Daarbij wijzen we er nogmaals op dat we aanwijzingen gevonden hebben dat er geen structurele veranderingen plaats hebben gevonden die het aggregeren van de data zoals wij dit gedaan hebben ongeldig zouden maken.

Bij een vergelijking van onze conclusies met die van Schmidt (1983) en Rosenthal (1983) dient opgemerkt te worden dat naar onze mening hier, in tegenstelling tot genoemde studies, gebruik is gemaakt van een statistisch geëigend, gedragstheoretisch gefundeerd model. De conclusies met betrekking tot de geaggregeerde data zullen we nogmaals tegen die van beide auteurs afzetten.

De conclusies van Schmidt aangaande kerkbezoek en leeftijd worden ondersteund, waarbij aangetekend zij dat de categoriale indeling gedetailleerdere conclusies mogelijk heeft gemaakt dan bij Schmidt te vinden zijn. Er is een sterkere invloed van inkomen gevonden dan bij Schmidt en de invloed van opleiding wordt hier, in tegenstelling tot in zijn studie, wel

gevonden.

De conclusie van Rosenthal dat de sociaal-economische positie invloed heeft op de opkomst bij verkiezingen, wordt ondersteund. In dit artikel is echter onderscheid gemaakt tussen subjectieve en objectieve sociale categorie. Beide blijken van invloed te zijn, doch alleen laatstgenoemde is vergelijkbaar met de indeling van Rosenthal. Ook deze vergelijking gaat echter niet geheel op, daar wij een andere indeling in groepen gekozen hebben. Met name de opdeling tussen publieke sector en private sector blijkt van belang te zijn.

Verder onderzoek is zeker gewenst. Op de eerste plaats zijn wij van mening dat veranderingen in variabelen een belangrijkere verklaring kunnen leveren dan niveaus ervan. Iemand zal bijvoorbeeld eerder beïnvloed worden door een hoog inkomen als hij/zij gewend is aan een laag inkomen dan als altijd al een hoog inkomen genoten is. En verder zijn we niet alleen geïnteresseerd in het al dan niet gaan stemmen, maar ook in de partijkeuze en het verklaren daarvan. Hieraan wordt gewerkt.

Noten

* Het hier gepresenteerde onderzoek is onderdeel van het onderzoeksprogramma 'Economische Politiek en Belangentegenstellingen' van de Universiteit van Amsterdam. Van financiële ondersteuning door de Nederlandse Organisatie voor Zuiver Wetenschappelijk Onderzoek (ZWO) is dankbaar gebruik gemaakt.

De gebruikte gegevens in dit rapport zijn oorspronkelijk verzameld voor de nationale kiezersonderzoeken 1971, 1972, 1977, 1981 en 1982 door een aantal onderzoekers in opdracht van de interuniversitaire werkgroepen nationale kiezersonderzoeken. De oorspronkelijke onderzoekers dragen geen verantwoordelijkheid voor de analyses of de interpretaties in dit rapport.

Het onderzoeksbestand is samengesteld en gedocumenteerd door de oorspronkelijke onderzoekers, met ondersteuning van het Steinmetz-Archief. De data worden gedistribueerd door het Steinmetz-Archief, Amsterdam.

1. Een andere reden waarom logit- en probit-modellen hier de voorkeur verdienen, is dat de OLS-schatters in geval van een lineair model wegens heteroscedasticiteit inefficiënt en derhalve niet optimaal zijn, hetgeen de resultaten minder betrouwbaar maakt.

2. Momenteel wordt ook de invloed van een aantal variabelen op partijkeuze onderzocht, waarbij ook opkomstgedrag onderscheiden naar partijvoorkeur aan bod komt.

3. De categorie werklozen heeft in alle jaren te weinig waarnemingen om statistisch betrouwbare conclusies mogelijk te maken. Om deze reden wordt deze groepering verder niet in de analyse betrokken.

4. De coëfficiënt behorende bij het middelbaar onderwijs heeft overigens een

onverwacht (negatief) teken: zij die middelbaar onderwijs genoten hebben, hebben een lagere opkomstkans in 1981 dan zij die alleen lager onderwijs gevolgd hebben.

5. Tenzij anders vermeld, is statistische significantie hier steeds met behulp van een tweezijdige toets getest.

6. Er dient nog iets gezegd te worden over de R^2 , die in tabel 1 vermeld staat. Het betreft hier de zogenaamde pseudo- R^2 :

$$\text{pseudo-}R^2 = (1 - (L_0/L_\Omega)^{2/n}) / (1 - L_0)^{2/n} \quad (2)$$

met

L_0 : het maximum van de aannemelijkheidsfunctie onder de restrictie dat alle coëfficiënten behalve de constante term gelijk zijn aan 0;

L_Ω : het maximum van de aannemelijkheidsfunctie met betrekking tot alle coëfficiënten;

n : de steefproefomvang.

Deze grootte verdient de voorkeur boven de gebruikelijke R^2 als men met een kwalitatieve afhankelijke variabele te maken heeft, omdat de gebruikelijke R^2 een maximale waarde van kleiner dan 1 kan hebben in een dergelijk model. Bovendien is de gebruikelijke interpretatie in de zin van de 'verklaarde variantie' niet van toepassing, zo gauw men van een lineair model afwijkt, of in een lineair model met een kwalitatieve variabele te maken heeft. Om deze redenen dienen de R^2 -waarden die Schmidt (1983) vermeldt dan ook met een kritisch oog bekeken te worden. Opgemerkt dient echter te worden dat ook de pseudo- R^2 (die wel een maximum van 1 heeft) slechts betrekkelijke waarde heeft, mede omdat er niet gecorrigeerd wordt voor het aantal vrijheidsgraden. De enige betekenis die deze grootte met zekerheid heeft, is dat het een vergelijkingsmaatstaf is voor andere modellen die met eenzelfde aantal vrijheidsgraden op dezelfde data betrekking hebben. Een verdere uiteenzetting over 'goodness-of-fit' maatstaven bij modellen met kwalitatieve variabelen kan gevonden worden in Amemiya (1981) of Maddala (1983). We merken nog op dat we naast de gemelde pseudo- R^2 (vergelijking (2)) ook gebruik hebben gemaakt van het Akaike Information Criterion, te vinden in Amemiya (1981). Verder hebben we, waar mogelijk, een 'likelihood ratio test' uitgevoerd. Geen van deze drie criteria heeft ons een wezenlijke verbetering van het model helpen vinden. De hypothese dat de opkomstkans voor ieder individu gelijk is, werd elk jaar duidelijk verworpen.

Geraadpleegde literatuur

- Amemiya, T. (1981), Qualitative response models: A survey; *Journal of Economic Literature* vol. XIX, p. 1483-1536.
- Bennet, J., en W. Orzechowski (1983), The voting behavior of bureaucrats: Some empirical evidence; *Public Choice* 41, p. 271-283.
- Foster, C. (1984), The performance of rational voter models in recent presidential elections; *The American Political Science Review*, vol. 78, p. 678-690.
- Frey, B. (1971), Why do high income people participate more in politics?; *Public*

Choice 11, p. 101-105.

Frey, B., en W. Pommerehne (1982), How powerful are public bureaucrats as voters?; *Public Choice* 38, p. 253-262.

Katosh, J., en M. Traugott (1981), The consequences of validated and self-reported voting measures; *Public Opinion Quarterly*, vol. 45, p. 519-535.

Maddala, G. (1983), *Limited-dependent and qualitative variables in econometrics*, Cambridge University Press.

Rosenthal, U. (1983), Ambtenaren, particuliere werknemers, werklozen en arbeidsongeschikten, thuisblijvers en kiezers in slechte tijden; *Acta Politica*, 83/2.

Schmidt, O. (1983), Kiezersopkomst van 1971 tot 1982; *Acta Politica*, 83/2.

Sigelman, L. (1982), The non-voting voter in voting research; *The American Journal of Political Science* 26, p. 47-56.

Winden, F. van (1983), *On the interaction between state and private sector*, North Holland, Amsterdam.

Wolfinger, R., en S. Rosenstone (1980), *Who votes?*, Yale University Press.