



Kwaliteitsinformatie middelbare scholen maakt verschil

Sinds 1997 publiceert het dagblad Trouw jaarlijks kwaliteitsinformatie over scholen in het voortgezet onderwijs. De publicatie door Trouw blijkt mede van invloed te zijn op de schoolkeuze van ouders en leerlingen, vooral bij vwo-scholen. Tegelijkertijd verbeteren de scholen na een negatief oordeel hun onderwijsprestaties.

Sinds enige jaren staat de transparantie van onderwijskwaliteit hoog op de politieke agenda (Dijkstra *et al.*, 2001; Dijkgraaf *et al.*, 2008; Koning en Noailly, 2009). Inmiddels is kwaliteitsinformatie over scholen in het voortgezet onderwijs gemeengoed geworden; vooral de Schoolprestaties van Trouw hebben hier een leidende rol in gehad. Sinds 1996 publiceert het dagblad iedere herfst voor iedere middelbare school de eindexamenprestaties, de doorstroom in de onderbouw en een integraal eindoordeel. De beoogde effecten van deze vorm van kwaliteitstransparantie spreken voor zich: beter geïnformeerde onderwijskeuzes en sterkere prestatiestimuli voor de scholen. De internationale literatuur lijkt dit ook te bevestigen. Zo vinden Hastings en Weinstein (2008) in een veldexperiment dat ouders zich bij de keuze

van lagere scholen laten leiden door kwaliteitsinformatie. Daarnaast zijn meerdere studies verschenen naar de effecten van de *No Child Left Behind Act* (NCLB), die in de Verenigde Staten zorgde voor meer openbaarheid en intensievere onderwijsinspecties (Dee en Jacob, 2009; Chiang, 2009). Het algemene beeld is dat de NCLB heeft geleid tot een duidelijke verbetering van de prestaties van scholen. Het is interessant te bezien of de publicaties in Trouw, als initiatief vanuit de markt en niet vanuit de overheid, soortgelijke effecten hebben gehad.

Tegenover de beoogde meerwaarde van kwaliteitsinformatie staan echter ook potentiële gevaren. Zo zou een tweedeling tussen hoog- en laagkwalitatief onderwijs kunnen optreden als alleen ouders uit de hogere sociaaleconomische groepen gebruikmaken van de informatie en de beste scholen kiezen.

Daarnaast is het de vraag of scholen niet vooral op kunstmatige wijze hun prestaties verbeteren, bijvoorbeeld door slecht presterende leerlingen geen examens te laten afleggen (Chiang, 2009). Bovendien zouden scholen in een neerwaartse spiraal terecht kunnen komen als slechte oordelen, dalingen in leerlingenaantallen en de uitholling van kwaliteit elkaar blijven opvolgen. Dergelijke ontwikke-

PIERRE KONING
Programmaleider bij het
Centraal Planbureau

KAREN VAN DER WIEL
Wetenschappelijk mede-
werker bij het Centraal
Planbureau

lingen zijn extra wrang als een slecht kwaliteitsoordeel gebaseerd is op tegenvallers buiten de invloedssfeer van schoolbesturen en docenten.

De vraag is of de publicatie van Trouw's Schoolprestaties tot dusver inderdaad tot deze bedoelde en onbedoelde effecten heeft geleid. Deze publicaties strekken zich inmiddels uit over een langere periode, zodat het mogelijk is te onderzoeken hoe scholen, leerlingen en ouders op deze kwaliteitsoordelen reageren, zowel op korte als middellange termijn. Gegevens over kwaliteitsindicatoren en Trouw-rapportages zijn nu voor handen voor de periode van 1997 tot en met 2008. Daarbij komt dat het sinds kort mogelijk is individuele gegevens van scholieren via het CBS in te zien. Deze informatie biedt niet alleen zicht op de gevoeligheid van schoolkeuzes voor informatie over de kwaliteit van scholen, maar plaatst de effecten ervan ook in breder perspectief. Met de scholierengegevens is namelijk het relatieve belang van kwaliteitsoordelen te bepalen, vergeleken met die van reistijd en andere factoren. Bovendien kan onderzocht worden of de reactie op kwaliteitsoordelen verschilt tussen scholieren, met name tussen de lagere en hogere sociaaleconomische groepen.

Het Trouw-oordeel

Sinds 1997 publiceert Trouw, op basis van een eigen bewerking van leerprestaties van de Onderwijsinspectie, jaarlijks een ranglijst van middelbare scholen die onderverdeeld is naar schooltypen. Het dagblad had deze primeur nadat een journalist zich succesvol beriep op de Wet openbaarheid van bestuur om kwaliteitsgegevens van de Onderwijsinspectie openbaar te krijgen. Later volgden zowel de Inspectie zelf als het weekblad Elsevier; gezien het langere tijdsbestek richt dit onderzoek zich echter op de Trouw-publicaties. Anders dan de Inspectie kent Trouw een eindoordeel toe aan ieder schooltype dat een school aanbiedt. Dit eindoordeel is een gewogen gemiddelde van drie kwaliteitsindicatoren van de Inspectie, waarna correctie plaatsvindt voor enkele kenmerken van leerlingen. De drie kwaliteitsindicatoren die Trouw gebruikt zijn: het onderbouwrendement, dit is het netto percentage leerlingen dat na de onderbouw onder of boven het initiële schooladvies uitkomt; het bovenbouwrendement, dit is het percentage leerlingen dat zonder vertraging een diploma haalt vanaf het derde leerjaar; en het gemiddelde cijfer voor het centraal schriftelijk eindexamen (CE). Hoewel de procedure ieder jaar enigszins is aangepast, heeft Trouw de totale score in alle jaren in ieder geval gecorrigeerd voor het percentage leerlingen behorend tot een culturele minderheid. Vervolgens verdeelt de krant alle scholen over vijf opeenvolgende categorieën, van meest positief tot meest negatief. Dit is ook het oordeel dat uiteindelijk de meeste aandacht krijgt in de publicatie en dat een aantal regionale dagbladen voor de voor hun regio relevante lijstjes overneemt. Een belangrijk kenmerk van de Trouw-publicatie is dat deze oor-

Tabel 1

Beschrijvende statistiek van de kwaliteitsgegevens op schooltype niveau per Trouw-oordeel.						
	Niet beschikbaar	Meest negatief	Negatief	Gemiddeld	Positief	Meest positief
Observaties	1.704	169	2.077	6.429	2.352	97
Percentage vwo	12	1	16	50	20	1
Percentage havo	15	2	17	47	18	1
Percentage vmbo-gt (mavo)	13	1	16	52	17	1
Totaal aantal scholieren	664	880	1.008	1.068	1.031	785
Aantal eerstejaars-scholieren	49	58	73	81	81	69
Rendement onderbouw	97%	87%	92%	97%	102%	108%
Percentage diploma zonder vertraging	71	49	59	72	77	82
Gemiddeld CE-cijfer	6,3	6,0	6,2	6,4	6,5	6,6

deelt over de prestaties van scholen van twee jaar eerder. De oorzaak hiervan ligt in vertragingen in de aanlevering van de gegevens door de Inspectie en de verwerking van die gegevens door de krant. Scholen zijn hierdoor wel direct op de hoogte van hun eigen absolute scores op indicatoren, maar kunnen moeilijk anticiperen welke rangorde Trouw hun toebedeelt. Deze is namelijk gebaseerd op relatieve prestaties, waarbij de exacte formule ook nog ieder jaar anders is. Gevolg is dat scholen en scholieren niet eerder dan na drie jaar effectief kunnen reageren op de Trouw-publicatie; hetzij in de schoolkeuze, hetzij met extra of juist minder inspanning om de onderwijskwaliteit op peil te houden (Koning en Van der Wiel, 2010a; 2010b).

Scholengegevens en scholierengegevens

Tabel 1 laat enkele beschrijvende statistieken van de scholengegevens zien. Duidelijk is dat maar weinig schooltypen de meest negatieve of de meest positieve score ontvangen. Opvallend is verder dat niet alle schooltypen een Trouw-oordeel hebben. Over het algemeen gaat het hierbij om kleine scholen, waarvan de gemiddelde kwaliteitsscores te onbetrouwbaar worden geacht. De gemiddelde kwaliteitsindicatoren van de scholen zonder Trouw-score liggen overigens dicht tegen die van de gemiddelde categorie.

Tabel 2 presenteert vervolgens enkele statistieken van de individuele gegevens van scholieren die in 2003 voor het eerst naar de middelbare school gingen. Dit bestand volgt uit koppeling van administratieve leerlinggegevens van het CBS, het Regionaal Inkomensonderzoek en de scholengegevens uit tabel 1. Het gaat hier in totaal om een steekproef van 23.923 leerlingen, circa twintig procent van het aantal brugklassers in het schooljaar 2003/2004. Van elke leerling is bekend welke school en schooltype is gekozen en waar hij of zij woonachtig is op zescijferig postcodeniveau. Per leerling kan aldus een keuzeset gecreëerd

Tabel 2

Beschrijvende statistiek van de scholierengegevens per schooltype (2003/2004).			
	Vwo	Havo	Vmbo-gt
Aantal scholieren	7.430	7.176	9.317
Aantal combinaties scholieren en scholen binnen twintig kilometer van huisadres	194.834	184.655	290.783
Huishoudinkomen onder het 25^{ste} percentiel	14,8%	18,5%	23,9%
Huishoudinkomen boven het 25^{ste} percentiel	42,0%	30,4%	21,8%
Aantal scholen in de keuzeset binnen 20 km	29	27	37
Minimumafstand naar een school	2,7	2,9	2,3
Percentage scholieren dat dichtstbijzijnde school kiest	36,4	41,9	40,1
Afstand naar de gekozen school in km	4,3	4,2	3,9

Tabel 3

Aantal leerlingen dat extra een school kiest door Trouw-oordeel vergeleken met neutraal oordeel¹.

Oordeel	Scholieren dataset	Scholen dataset	
Meest negatief	-15,834**	-2,279	-2,667
Negatief	-8,760**	-1,885**	-2,667**
Neutraal	Referentiecategorie		
Positief	7,128**	1,207*	1,610
Meest positief	5,658	7,729*	16,356**
Aantal observaties	670.272	9.064	2.702
Steekproef	Alle combinaties van scholieren en scholen	Alle scholen	Scholen in vwo
Schattingmethode	Conditional logit	Vaste effecten per schooltype	Vaste effecten per schooltype

* Significant op vijfprocent-niveau, ** significant op éénprocent-niveau.

¹ Additionele controlevariabelen zijn jaardummy's, de bevolkingsgrootte in de gemeente in de leeftijd tien tot twintig jaar, het aantal scholen in de gemeente, het totaal aantal scholieren, het aantal vestigingen dat onder dezelfde school valt, het aandeel scholieren in ieder schooltype en het aandeel leerlingen behorende tot een culturele minderheid.

worden van alle scholen binnen twintig kilometer van huis. Het scholierenbestand bevat onder andere ook informatie over de etniciteit van de leerling en over het inkomen en de belangrijkste inkomensbron van zijn of haar huishouden. Ongeveer veertig procent van de leerlingen kiest de dichtstbijzijnde school. De gemiddelde reisafstand naar de gekozen school is het langst voor vwo-leerlingen en het kortst voor vmbo-gemengde en theoretische leerweg leerlingen.

Effect op schoolkeuze

In beginsel is het mogelijk om met zowel de longitudinale scholengegevens als de scholierengegevens het effect van de kwaliteitsinformatie op schoolkeuze vast te stellen. Om met de scholierengegevens te beginnen: er kan een *Conditional logit*-model geschat worden van de keuze die iedere leerling maakt uit alle scholen binnen een straal van twintig kilometer van het woonadres. Met deze schattingen is het relatieve belang van de kwaliteitsgegevens ten opzichte van de reistijden te berekenen. De resultaten die volgen uit de schatting van dit model zijn weergegeven in de eerste kolom van tabel 3. De effecten van het Trouw-oordeel op individueel niveau zijn geaggregeerd tot het niveau van scholen, om zodoende de geïmpliceerde toe- of afname in de instroom van nieuwe leerlingen per school en schooltype vast te kunnen stellen.

Uit de schattingen blijkt dat het Trouw-oordeel wel degelijk effect heeft op de instroom van leerlingen. Een schooltype met een negatieve beoordeling ontvangt negen minder scholieren, terwijl een schooltype met een positieve score zeven scholieren meer verwelkomt. Dit komt overeen met zo'n twaalf en negen procent van de gemiddelde instroom van 76 leerlingen per schooltype. Aanvullende analyses leren dat deze gevoeligheid voor de Trouw-oordelen weinig verschilt tussen inkomenspercentielen of etnische groeperingen. Met andere woorden: huishoudens uit hogere inkomensklassen wegen de Trouw-publicatie niet zwaarder. Tot slot is reistijd zonder twijfel de meest bepalende factor voor de schoolkeuze: per extra kilometer verder van huis neemt de kans dat een leerling een school kiest met dertig procent af. Omgerekend betekent dit dat een leerling indifferent is tussen een school die neutraal scoort en 844 meter verder weg ligt, en een school die de meest negatieve beoordeling scoort. Andersom geldt dat een leerling indifferent is tussen een neutrale school en een school die goed wordt beoordeeld en 222 meter verder ligt.

Ook de overige kolommen in tabel 3 geven inzicht in de reactie van het aantal aankomende leerlingen op de kwaliteitsinformatie, maar dan volgens schattingen met de longitudinale scholengegevens. In tegenstelling tot de scholierengegevens, die voor maar een jaar beschikbaar zijn, hebben deze het voordeel dat correctie mogelijk is voor factoren die constant zijn over de tijd en mede de populariteit van scholen bepalen. Denk hierbij aan de reputatie van een school of de aantrekkelijkheid van het schoolgebouw. Dit betekent dat het beter mogelijk is het afzonderlijke effect van het Trouw-oordeel te schatten. Wederom blijkt in dat

geval dat Trouw-scores een significante rol spelen bij de schoolkeuze; dit geldt in de eerste kolom voor drie van de vier geschatte coëfficiënten. De effecten zijn wel geringer, kennelijk omdat nu inderdaad correctie plaatsvindt voor kwaliteitsfactoren die samenhangen met het Trouw-oordeel en ook van invloed zijn op de schoolkeuze. Anders gesteld: de schattingen met de scholierengegevens geven een indruk van het effect van kwaliteitsinformatie in brede zin, terwijl de longitudinale gegevens dat doen voor het partiële effect van de Trouw-publicatie.

Tot slot laat de derde kolom in tabel 3 resultaten zien voor alleen de vwo-afdelingen van scholen. De reactie van leerlingen op de Trouw-scores blijkt bij deze groep sterker te zijn dan gemiddeld. Zo neemt de groep vwo-brugklassers met zestien scholieren toe in het jaar nadat het meest positieve kwaliteitsoordeel is ontvangen. Blijkbaar vinden vwo-leerlingen kwaliteit belangrijker of zijn zij beter in het vinden van de relevante informatie. Uit de eerder genoemde resultaten met de scholierengegevens blijkt dat dit effect niet te verklaren is uit het feit dat gezinnen met hogere inkomens meer op de Trouw-scores letten. Vwo-leerlingen uit alle geleidingen van de samenleving zijn dus meer gespitst op schoolkwaliteit.

Effect op kwaliteit van scholen

Kwaliteitsinformatie, en de Trouw-oordelen in het bijzonder, beïnvloeden dus de instroom van nieuwe leerlingen. Om die reden kan scholen er het nodige aan gelegen zijn hun kwaliteit op peil te houden. Het is de vraag of scholen zich in de volgende jaren van negatieve scores herstellen en zij hun kwaliteit verbeteren om verdere terugloop van instroom te voorkomen. Andersom geldt voor de scholen met een positieve score de vraag of zij minder druk voelen om te presteren en hun kwaliteit verslechtert, of dat een opwaartse spiraal van nog betere prestaties in werking treedt.

Om de effecten van het Trouw-oordeel op kwaliteit vast te stellen, is het in de eerste plaats noodzakelijk om de uitkomsten zowel na de meest positieve als na de meeste negatieve beoordelingen in termen van afwijkingen van gemiddelden te meten.

Net als bij het effect op de leerlingeninstroom betekent dit dat correctie plaatsvindt voor vaste effecten per school. Zo zullen kwalitatief goede scholen sowieso een hogere kans op een positief Trouw-oordeel hebben. Het gaat er dus om hoe scholen reageren wanneer het Trouw-oordeel hoger of lager uitvalt dan voorheen het geval was. Dankzij de vertraging tussen de registratie van gegevens en de rapportage van het oordeel ontstaat een goed beeld van deze reactie. Zonder vertraging zal na een uitzonderlijk goed of slecht jaar als vanzelf een meer gematigd jaar volgen, bij wijze van een regressie naar het gemiddelde. Gevolg is een vertekening van effecten, in die zin dat een negatief oordeel automatisch tot verbetering lijkt te leiden. Dat is nu dus niet het geval.

Tabel 4 laat de geschatte effecten zien van het meest positieve en het meest negatieve Trouw-oordeel op de drie kwaliteitsindicatoren van scholen: het percentage diploma's zonder vertraging, het gemiddeld centraal schriftelijk eindcijfer en het rendement van de onderbouw. Hierbij zijn de effecten van de Trouw-publicatie op twee, elkaar aanvullende termijnen geschat: het effect een jaar na de publicatie en daarna. De keuze voor alleen de meest extreme oordelen van de krant als impulsvariabelen is niet zonder reden. De overgrote meerderheid van scholen begeeft zich namelijk afwisselend in een van de drie middencategorieën, met als gevolg dat de langetermijneffecten van een reeks van deze gematigde oordelen moeilijk te interpreteren zijn. Voor de meest positieve en de meest negatieve scores is dit echter niet het geval. In de onderzochte periode ontving namelijk maar circa vier en zes procent van alle scholen uiteindelijk het meest positieve respectievelijk meest negatieve oordeel. Tabel 4 geeft aan dat de effecten van het Trouw-oordeel op kwaliteitsindicatoren het sterkst zijn bij het percentage diploma's en bij het gemiddelde CE-cijfer. Voor beide indicatoren leidt de meest negatieve score tot verbeteringen op kwaliteit, terwijl voor de meest positieve score, zij het in iets mindere mate, geldt dat de kwaliteit naderhand achteruitgaat. Verder blijkt dat de hogere slagingspercentages niet zozeer het gevolg zijn van hogere scores op de schoolonderzoeken, maar van de CE-cijfers. Het lijkt er dus niet op dat scholen de zwaarte van hun schoolonderzoeken verlaagd hebben.

Met een toename van circa 0,10 punt oogt het langetermijneffect van het meest negatieve Trouw-oordeel op het gemiddelde CE-cijfer misschien gering, maar dit komt overeen met circa veertig procent van een standaarddeviatie van deze variabele. Relatief gezien doet dit effect zeker niet onder voor schattingen in de Amerikaanse literatuur over de effecten van NCLB en is ook in lijn met door Chorny en Webbink (2010) gevonden effecten van inspectieverzwaring in Amsterdam. In tegenstelling tot de door de Amerikaanse overheid geïnitieerde NCLB is de Trouw-publicatie echter een privaat initiatief, met alleen *naming and shaming* als disciplinerend mechanisme. Ten slotte valt op dat de geschatte effecten op de lange termijn meestal ongeveer even groot zijn als die op de korte termijn. Bij de meest negatieve score suggereert dit dat scholen gedurende een langere periode hun prestaties verbeteren, alsof dit hen bewust maakt van de noodzaak tot investeringen in beter onderwijs.

Conclusies

Al met al laten dus niet alleen leerlingen en ouders zich leiden door kwaliteitsinformatie in hun schoolkeuze, ook schoolbesturen reageren hierop. Het effect van de Trouw-informatie op de schoolkeuze is wel beperkt; de reisafstand naar een school is duidelijk belangrijker. Bovendien zullen ook andere informatiebronnen meetellen, met name die van de

Tabel 4

Geschatte korte- en langetermijneffect van Trouw-oordeel op kwaliteitsindicatoren ¹ .			
	Diploma's zonder vertraging	Eindexamen-cijfers	Rendement onderbouw
Kortetermijneffect van meest negatieve oordeel	4,238**	0,090**	1,036
Kortetermijneffect van meest positieve oordeel	-3,704**	-0,031	-0,946
Langetermijneffect van meest negatieve oordeel	6,732**	0,111**	2,006
Langetermijneffect van meest positieve oordeel	-3,360**	-0,064**	-1,636

* Significant op vijfprocent-niveau, ** significant op éénprocent-niveau.

¹ Additionele controlevariabelen zijn jaardummy's, de bevolkingsgrootte in de gemeente in de leeftijd tien tot twintig jaar, het aantal scholen in de gemeente, het totaal aantal scholieren, het aantal vestigingen dat onder dezelfde school valt, het aandeel scholieren in ieder schooltype en het aandeel leerlingen behorende tot een culturele minderheid.

Onderwijsinspectie. Tegelijkertijd zijn de mogelijke nadelen van de publicatie ogenschijnlijk beperkt: de analyse leert dat niet alleen hogere inkomens hun voordeel doen met de informatie van Trouw en dat er bij slecht presterende scholen geen sprake is van een neerwaartse spiraal die door de openbare kwaliteitscores in gang wordt gezet.

Duidelijk moge zijn dat het Trouw-oordeel effect heeft, maar dat wil nog niet zeggen dat het een accuraat kwaliteitsoordeel is (Dijkstra *et al.*, 2001). Zeker bij het meest negatieve en het meest positieve oordeel gaat het om zeer kleine groepen scholen die in de regel gedurende een jaar extreem scoren. Niet altijd zullen zulke uitzonderlijke jaren de scholen aan te rekenen zijn, terwijl het effect van vooral de meest negatieve score op lange termijn doorwerkt op de schoolkeuze van leerlingen. Dit euvel zou deels voorkomen kunnen worden door kwaliteitsindicatoren te middelen over jaren, of door scholen evenrediger over de vijf categorieën te verdelen.

LITERATUUR

- Chiang, H. (2009) How accountability pressure on failing schools affects student achievement. *Journal of Public Economics*, 93(9), 1045-1057.
- Chorny, V. en D. Webbink (2010) The effect of accountability policies on primary education in the Amsterdam. *CPB Discussion paper*, 144.
- Dee, T.S. en B. Jacob (2009) The impact of No Child Left Behind on student achievement. *NBER working paper*, 15531.
- Dijkgraaf, E., S. van der Geest en M. de Jong (2008) Concurrentie bevordert onderwijsrendement niet. *ESB*, 93(4548), 730-732.
- Dijkstra, A.B., S. Karsten, R. Veenstra en A.J. Visscher (red.) (2001) *Het oog der natie: scholen op rapport. Standaarden voor de publicatie van schoolprestaties*. Assen: Koninklijke van Gorcum.
- Hastings, J. en J. Weinstein (2008) Information, school choice and academic achievement: Evidence from two experiments. *The Quarterly Journal of Economics*, 123(4), 1373-1414.
- Koning, P. en J. Noailly (2009) Schoolkeuze, concurrentie en kwaliteit in het basisonderwijs. *ESB*, 94(4554), 118-120.
- Koning, P. en K. van der Wiel (2010a) Ranking the schools: how quality information affects school choice in the Netherlands. *CPB Discussion Paper*, 150.
- Koning, P. en K. van der Wiel (2010b) School responsiveness to quality rankings. An empirical analysis of secondary education in the Netherlands. *CPB Discussion Paper*, 149.