

De voorspellingsmarkt: een nieuw instrument ter ondersteuning van strategische beslissingen

A. Borison is bestuurder bij NERA en biomassa-energiebedrijf Agni Corporation. Gregory Hamm is hoofdeconoom bij NERA en bestuursvoorzitter van het gemeentelijk elektriciteitsbedrijf van Alameda (Calif.) Daarnaast zijn de auteurs bijzonder docent aan de universiteit van Stanford.¹

Onzekerheid is troef bij beleid – zowel in het bedrijfsleven als bij de overheid. De bestaande methoden om greep te krijgen op die onzekerheid zijn nog steeds zeer gebrekkig. Een instrument dat sinds enkele jaren steeds vaker wordt toegepast, biedt wellicht perspectief op verbetering.

Veel strategische beslissingen gaan gepaard met onzekerheid, zowel in het bedrijfsleven als bij de overheid. Bedrijven hebben er op allerlei terreinen mee te maken: bij onderzoek en ontwikkeling, wanneer ze een nieuw product lanceren, een nieuwe markt aanboren, een project ontwikkelen of een bepaalde technologie kiezen. Een goed voorbeeld is de beslissing van Toyota om het groene licht te geven voor de Prius. Tijdens de ontwikkeling en zelfs in het begin van de marktintroductie kampte dat project met allerlei commerciële en technische problemen. Succes was allesbehalve verzekerd. Ondanks alle onzekerheid stond ceo Watanabe van Toyota echter vierkant achter het project. In een interview met *Forbes* in 2006 zei hij: 'Ik had niet gedacht dat het zo'n groot succes zou worden. Sommigen verwachtten dat het snel zou gaan lopen, anderen dachten dat het geleidelijker zou gaan. Ik behoorde tot de tweede categorie.' De leider van het project, Takehisa Yaegashai, zou ooit gezegd hebben dat Toyota de Prius in het begin slechts vijf procent kans op succes gaf.² Inmiddels geldt de Prius natuurlijk als een behoorlijk groot succes in de markt. Onzekerheid speelt ook bij veel beleid en regelgeving van de overheid. Daarbij gaat het vaak om maatregelen met vergaande consequenties voor de samenleving. Denk maar aan onderwerpen als de klimaatverandering, de gezondheidszorg of de nationale veiligheid. Het meest bekende en controversiële voorbeeld uit het recente verleden is misschien wel de invasie van Irak. Die werd vooral gerechtvaardigd door de bewering dat Irak hoogst-

waarschijnlijk beschikte over massavernietigingswapens. Daar leek ook vrij veel bewijs voor te zijn, al waren niet alle aanwijzingen even overtuigend. In 2003 zei woordvoerder Ari Fleischer van het Witte Huis: 'We zijn er tamelijk zeker van, dat zij massavernietigingswapens hebben.'³ Het lijkt vreemd, maar kwantitatieve inschattingen van dit vermoeden die van vóór de invasie dateren zijn moeilijk te vinden. Twee officieren hebben echter naderhand alsnog zo'n kwantitatieve analyse uitgevoerd, met een interessante uitkomst. Zij concludeerden dat de destijds beschikbare aanwijzingen overeenkwamen met een kans op massavernietigingswapens van 75 procent.⁴ Inmiddels weten we natuurlijk dat die wapens niet zijn gevonden. Niet alle strategische beslissingen zijn zo ingrijpend (of veelbesproken) als de invasie van Irak. Maar er worden wel geregeld strategische beslissingen genomen die grote consequenties hebben voor degene die ze neemt, de organisatie in kwestie en de belanghebbenden bij die organisatie. In het bedrijfsleven zijn die belanghebbenden behalve de medewerkers ook de klanten en leveranciers, de aandeel- en obligatiehouders en de regelgevende c.q. toezichthoudende instanties. Beslissingen in de publieke sector kunnen veel of zelfs alle burgers raken. Verantwoordelijke partijen doen hun best om de best denkbare informatie te vergaren om hun beslissing te onderbouwen. De betrokken belanghebbende partijen verlangen dat ook steeds vaker. Sommige strategische beslissingen zijn voor het grootste deel gebaseerd op gegevens. De onzeker-

1. NERA: *National Economic Research Associates* (nadere toelichting in het artikel). De auteurs zijn te bereiken via adam.borison@nera.com resp. gregory.hamm@nera.com.

2. Adrian J. Slywotzky, *The Upside: The 7 Strategies for Turning Big Threats into Growth Breakthroughs*, Crown Publishing Group, New York, 2007, p. 24.

3. 'We have high confidence that they have weapons of mass destruction.' Bron: Don Van Natta, Jr. en David Johnston, 'A Nation At War: Banned Weapons', *New York Times*, 14 april 2003.

heid waarmee men te maken heeft, kan dan worden ingeschat aan de hand van uitgebreide historische informatie. Het kost weliswaar tijd en moeite om die gegevens te verzamelen en te interpreteren, maar een bijzonder complexe opgave is dat niet. Denk bijvoorbeeld aan het weer en de grondstoffenprijzen. Voor allebei zijn er uitgebreide historische gegevens voorhanden. Als de weersomstandigheden op een bepaalde plek in een bepaald seizoen relevant zijn voor een belangrijke beslissing – bijvoorbeeld de locatie voor een nieuwe waterkrachtcentrale – kan die beslissing worden genomen aan de hand van grote hoeveelheden gegevens.

In andere gevallen zal een strategische beslissing moeten worden genomen op grond van een oordeel. Er zijn dan geen, of bijna geen, direct relevante historische gegevens aan de hand waarvan de betreffende onzekerheden kunnen worden ingeschat. In dat geval moeten de beleidsmakers er zelf over oordelen (of zich door anderen laten adviseren) en dit oordeel meewegen in hun besluitvorming. Dit speelt bijvoorbeeld bij technologische ontwikkeling en maatschappelijke verandering. Voor dergelijke beslissingen zijn misschien wel meer of minder vergelijkbare gevallen uit het verleden bekend, maar een databank van volstrekt vergelijkbare gevallen bestaat niet. Als een belangrijke beslissing moet worden genomen op basis van een nieuwe technologie, zoals het afvangen van koolstofdioxide bij de uitstoot van afvalgassen, dan zal de beslissing eerder berusten op een oordeel dan op een analyse van kwantitatieve gegevens.

In de praktijk is het verschil tussen een beslissing op basis van kwantitatieve gegevens en een beslissing op basis van een oordeel natuurlijk nooit zo zwart-wit als wij hier schetsen. Ook wanneer er veel gegevens voorhanden zijn, speelt het oordeel een rol.⁵ De gegevens over het weer zullen we interpreteren in het licht van onze mening over de klimaatverandering. Omgekeerd zal er ook naar gegevens worden gekeken wanneer het persoonlijk oordeel de doorslag geeft bij de besluitvorming. Bij een beslissing over het al dan niet afvangen van broeikasgassen kan bijvoorbeeld worden gekeken naar beschikbare informatie over eerdere technologische innovaties. Kortom: zowel gegevens als het persoonlijk oordeel zijn belangrijk. Daarom wordt niet alleen onderzocht hoe de beste gegevens dan wel het beste oordeel kunnen worden verkregen – maar ook hoe oordeelsvorming en kwantitatieve analyse met elkaar kunnen worden gecombineerd.⁶ Persoonlijk oordeel speelt, afzonderlijk of in combinatie met gegevens, een cruciale rol bij de meeste

belangrijke beslissingen die worden genomen, zowel in de publieke als de private sector. Het belang ervan lijkt zelfs toe te nemen als we kijken naar de belangrijkste beleidsthema's van tegenwoordig voor ondernemingen en overheid: de oorzaak en het tempo van de klimaatverandering alsook de gewenste reactie daarop, de invloed van de mondialisering, de kans op herstel na de economische en financiële crisis, de rollen van particuliere en publieke instellingen in de samenleving, de risico's op regionale conflicten, de reikwijdte van milieuwetgeving, of de technologische ontwikkeling.

Met behulp van een nieuw instrument – de voorspellingsmarkt – kan onzekerheid bij strategische besluitvorming beter worden ingeschat. Zo kunnen beslissingen die iedereen raken op kwalitatief betere gronden worden genomen.

De huidige aanpak

Hoe wordt onzekerheid tot nu toe verdisconteerd in de strategische besluitvorming? Soms informeel. De beleidsmaker vertrouwt dan op intuïtie of ervaring. Of men hakt de knoop gewoon op goed geluk door. In andere gevallen wordt er zorgvuldiger te werk gegaan, maar ook dan op een kwalitatieve manier. Denk aan scenarioplanning, contingentieplanning, stresstests of risicogevoeligheidsanalyse. Met dergelijke methoden worden verschillende mogelijke uitkomsten benoemd, maar zonder een onderling vergelijkbare inschatting van de kans dat ze zich zullen voordoen.

Dit soort informele en kwalitatieve verkenningen hebben zeker hun waarde. Veel beleidsmakers die erop vertrouwen hebben er ook succes mee. De vraag blijft echter: in hoeverre kun je dergelijke inschattingen vertrouwen – en verdedigen – als een beslissing niet alleen afhangt van de mogelijkheid dat iets zal gebeuren, maar ook van de kans dat het gebeurt? Uiteraard geldt dit uiteindelijk voor iedere beslissing waarbij sprake is van onzekerheid. Hoe weet je of het zinvol is om actief te worden in een bepaalde markt? Dat hangt meestal af van een succesvolle entree in die markt. Moet er wel of niet zwaar worden geïnvesteerd in de ontwikkeling van een nieuw geneesmiddel? Dat hangt af van de kans dat het nieuwe geneesmiddel veilig zal zijn en het beoogde effect zal hebben.

Bij het nemen van strategische beslissingen wordt steeds meer vertrouwd op formele, kwantitatieve

Met formele analyse kunnen veel valkuilen van een louter intuïtieve besluitvorming worden vermeden

4. Maj. Ernest Y. Wong en Lt. Colonel Rod Roederer, 'Should the US Have Attacked Iraq? Can Decision Theory Shed Light on Polarizing Debate?' *OR/MS Today*, december 2006.

5. Voor een voorbeeld van een combinatie van oordeelsvorming en gegevens ontleend aan een model, zie Robert C. Blattberg en David J. Hoch, 'Database Models and Managerial Intuition: 50% Model + 50% Manager', *Management Science*, 36/8, 1990, pp. 887-899.

6. Voor een bespreking van methoden om statistische gegevens en oordeelsvorming te combineren bij tijdreeksprognoses, zie M.J. Lawrence, R.H. Edmundson en M.J. O'Connor, 'The Accuracy of Combining Judgmental and Statistical Forecasts', *Management Science*, 32/12, december 1986, pp. 1521-1532; J. Scott Armstrong en Fred Collopy, 'Integration of Statistical Methods and Judgment for Time Series Forecasting: Principles from Empirical Research', *International Journal of Forecasting*, 21, 1998, pp. 69-293.

Een kwantitatieve analyse gaf een kans van 75 procent op massavernietingswapens in Irak

7. J. Edward Russo en Paul J.H. Schoemaker, *Decision Traps: The Ten Barriers to Brilliant Decisions and How to Overcome Them*, Simon & Schuster, New York, 1989.

8. Carl-Axel S. Stael von Holstein en James Matheson, *A Manual for Encoding Probability Distributions*, SRI International, Menlo Park (Calif.), 1978.

9. Een uitgebreidere bespreking van vooringenomenheid is te vinden bij A. Tversky en D. Kahneman, 'Judgment under Uncertainty: Heuristics and Biases', *Science*, 27 september 1974, pp. 1124-1131.

10. Zie voor nadere informatie over waarschijnlijkheidscodering: Daniel G. Brooks en Timothy J. O'Leary, 'A Comparison of Encoding Techniques', *OMEGA: The International Journal of Management Science*, 11/1, 1983, pp. 49-55; Carl S. Spetzler en Carl-Axel S. Stael von Holstein, 'Probability Encoding in Decision Analysis', *Management Science*, 22/3, november 1975, pp. 340-358; Thomas S. Wallsten en David V. Budescu, 'Encoding Subjective Probabilities: A Psychological and Psychometric Review', *Management Science*, 29/2, februari 1983, pp. 151-173.

inschattingen van onzekerheid. Dat komt onder meer doordat de technologie voor het vergaren en verwerken van informatie in de afgelopen jaren sterk is verbeterd en wijd verbreid is geraakt. De mogelijkheden voor kwantificering zijn daarmee toegenomen.

Bovendien is het besef gegroeid dat formele analyse, en in het bijzonder kwantificering, de besluitvorming verbetert. Kwantificering is wenselijker geworden. Russo en Schoemaker beschrijven de diverse 'valkuilen' van een louter intuïtieve besluitvorming die met een meer formele analyse vermeden kunnen worden.⁷

Hoe kan de onzekerheid rond een persoonlijk oordeel worden gekwantificeerd? De huidige inzichten daarover zijn het resultaat van ruim een halve eeuw aan onderzoek en theorievorming omtrent besluitvorming ('decision theory'). Dit onderzoeksterrein heeft in de afgelopen decennia twee inmiddels gangbare, gestructureerde methoden voortgebracht om de mening van één respectievelijk meer deskundigen te peilen. Zowel bedrijven als overheidsinstellingen hebben deze twee methoden voor uiteenlopende doelen ingezet. Ze worden vooral gebruikt wanneer een beslissing grote consequenties kan hebben voor een organisatie. Voor de eerste methode, kanscodering, is zelfs een formeel handboek ontwikkeld.⁸

Kanscodering

Hoe onzeker is een mogelijke toekomstige gebeurtenis of ontwikkeling? Om dat te kunnen bepalen, is er meer nodig dan een deskundige simpelweg te vragen hoe groot hij of zij 'de kans op X' inschat. Iemands persoonlijke mening is niet altijd eenvoudig te registreren. Bovendien is die mening vaak gekleurd, zowel door cognitieve (onbewuste) vooroordelen als door bewuste overwegingen. Een voorbeeld van cognitieve kleuring is dat veel mensen een recente ervaring die indruk op hen heeft gemaakt onbewust relatief zwaar laten wegen in hun oordeel. Iets wat vers in het geheugen ligt, dringt zich op ('availability bias'). Een voorbeeld van bewuste kleuring is de vertegenwoordiger die zijn omzetprognose bewust laag houdt omdat hij een bonus krijgt als hij boven verwachting presteert.⁹

Kanscodering ('probability encoding') is hoofdzakelijk ontwikkeld om juist dit soort kleuring te neutraliseren. De formele coderingsprocedure zorgt ervoor, dat de ondervraagde begrijpt waarom

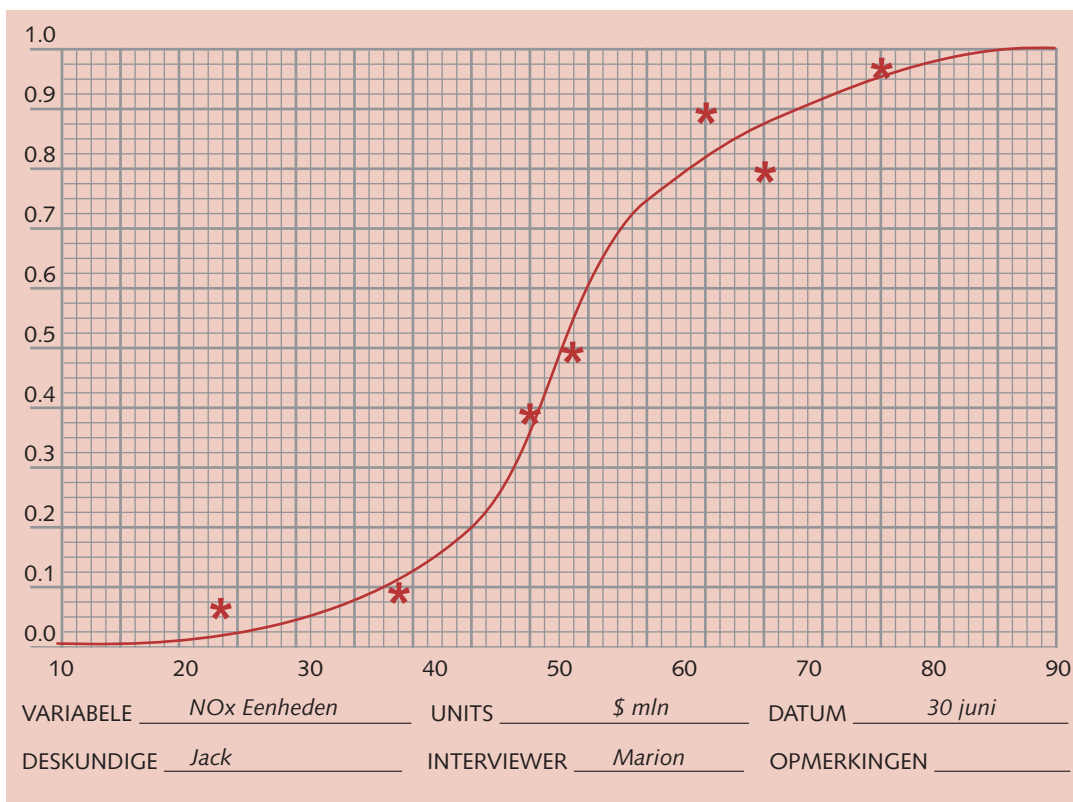
zijn inschatting nodig is en van goede kwaliteit moet zijn. Ook wordt datgene, waarvoor de kans wordt verkend, met deze methode scherp gedefinieerd. Indien nodig kan die gebeurtenis bovendien worden uitgesplitst in kleinere bestanddelen waar meer kennis over bestaat. Ten slotte zorgt de methode er ook voor, dat vooroordelen worden benoemd en geminimaliseerd en dat de ondervraagde expliciet zijn akkoord geeft aan de inschatting.¹⁰ Zoals Spetzler en Stael von Holstein laten zien, bestaat deze procedure uit vijf stappen:

- *Motivering*: contact leggen met de ondervraagde en onderzoeken of er sprake is van bewuste vooroordelen;
- *Structureren*: bepaling van de structuur van de in te schatten onzekerheid;
- *Conditionering*: de ondervraagde fundamenteel en zonder cognitief vooroordeel laten nadenken over zijn/haar oordeel;
- *Codering*: kwantificering van het oordeel van de ondervraagde in termen van kansen;
- *Verifiëring*: controle op consistentie van de door middel van de codering verkregen antwoorden.¹¹

Onderzoekers proberen deze procedure voor kanscodering verder te verfijnen en te versterken. Een van de aandachtspunten daarbij is de manier van ondervragen. Abbas e.a., bijvoorbeeld, hebben de uitkomsten van twee verschillende volgorden van vragen met elkaar vergeleken: bepalen van een waarde gevolgd door de benoeming van de kans resp. omgekeerd.¹² In ander onderzoek wordt er gewerkt aan scoringsregels gebaseerd op de feitelijke gebeurtenissen die eerlijke inschattingen zouden moeten bevorderen.¹³ Ondanks deze recentere ontwikkelingen is de procedure voor kanscodering in essentie al ruim twintig jaar hetzelfde.

Figuur 1 is gebaseerd op een toepassing in de praktijk van deze procedure voor een doorlopende variabele: de toekomstige kosten van een nieuwe technologie voor milieucontrole. Deze variabele kan zowel voor bedrijven als de overheid een belangrijke onbekende grootheid zijn om rekening mee te houden bij de besluitvorming. In figuur 1 vertegenwoordigt iedere asterisk (*) een antwoord op een specifieke vraag. Bijvoorbeeld: 'Waarvan is de kans groter – dat ik een zes gooi met een dobbelsteen of dat de nieuwe technologie minder kost dan twintig miljoen dollar?' Zijn alle antwoorden eenmaal verzameld, dan wordt een cumulatieve kansdistributie uitgetekend aan de hand van de

Figuur 1. Gecodeerde-kansdistributie voor de kosten van nieuwe technologie voor milieucontrol



gegevens; dit wordt vervolgens ter controle voorgelegd aan de ondervraagde. Is de codering afgerond, dan wordt de op het oordeel gebaseerde kansgrafiek ingezet bij de strategische besluitvorming (die verder plaatsvindt aan de hand van beslissingsanalyse, echte opties en dergelijke).

Met deze procedure worden behalve doorlopende ook eenmalige variabelen ingeschat. Denk bijvoorbeeld, om bij het thema van figuur 1 te blijven, aan de invoering van bepaalde milieuwetgeving.

Groepspeiling

Kanscodering volstaat wanneer men één enkele persoon wil raadplegen over de onzekerheid met betrekking tot een oordeel. Vaak is dat echter niet genoeg voor beleidsmakers en andere belanghebbenden. Door meer mensen te raadplegen, wordt beter duidelijk hoe er op dat moment naar een bepaalde kwestie wordt gekeken. Zo'n groepspeiling is ook beter te rechtvaardigen en verdedigen. In dit geval worden de inschattingen van een aantal personen tot één gezamenlijke inschatting gebundeld ('expert aggregation').

Een groepspeiling is in meer dan één opzicht lastiger dan kanscodering. Om te beginnen is de theorie ervoor analytisch complex. Bovendien moeten

de interacties tussen een aantal mensen worden gemanaged, wat ook niet eenvoudig is. Daarom zijn er voor dit doel verschillende benaderingen ontwikkeld, afhankelijk van de context. Het onderzoek hiernaar loopt sinds de jaren zestig.¹⁴

Zoals Clemen en Winkler schetsen, kunnen er in essentie twee benaderingen worden onderscheiden: wiskundig en op basis van overleg ('behavioral').¹⁵ In het eerste geval worden de kansinschattingen van de afzonderlijke ondervraagden verzameld en vervolgens aan de hand van een aantal regels of parameters wiskundig in één overkoepelende uitkomst vertaald. In het tweede geval overleggen de deskundigen met elkaar volgens een vooraf bepaalde procedure. Zij wisselen informatie uit, en passen hun inschatting van de kans dienovereenkomstig aan. Soms wordt er daarbij een consensus bereikt. In dat geval levert het proces een kansinschatting op waar de hele groep het mee eens is. Wordt er geen consensus bereikt, dan is er alsnog een aanvullende wiskundige bewerking nodig.

De wiskundige aanpak is licht in het voordeel ten opzichte van de overlegvorm, zeggen Clemen en Winkler. Van de verschillende wiskundige methoden lijkt een simpel gemiddelde van de afzon-

11. Spetzler and Stael von Holstein (1975), zie noot 10.

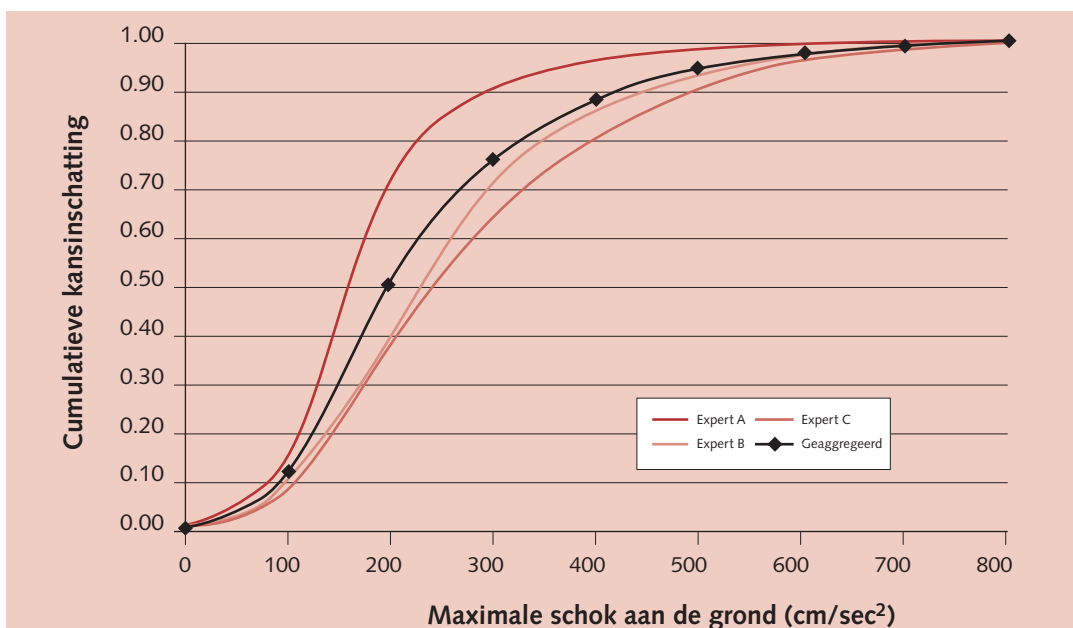
12. Ali E. Abbas, David V. Budescu, Hsiu-Ting Yu en Ryan Haggerty, 'A Comparison of Two Probability Encoding Methods: Fixed Probability vs. Fixed Variable Values', *Decision Analysis*, 5/4, 2008, pp. 190-202.

13. Robert L. Winkler, 'Scoring Rules and the Evaluation of Probabilities', *TEST*, 5/1, juni 1996, pp. 1-60.

14. Zie voor nadere informatie over de aggregatie van expertise: Robert T. Clemen en Robert L. Winkler, 'Combining Probability Distributions from Experts in Risk Analysis', *Risk Analysis*, 19/2, april 1999, pp. 187-203; Daniel Osherson en Moshe Y. Vardi, 'Aggregating Disparate Estimates of Chance', *Games and Economic Behavior*, 56/1, juli 2006, pp. 148-173; Thomas S. Wallsten, David V. Budescu, Ido Erev en Adele Diederich, 'Evaluating and Combining Subjective Probability Estimates', *Journal of Behavioral Decision Making*, 10/3, 1997, pp. 243-268.

15. Robert T. Clemen en Robert L. Winkler, 'Aggregation of Expert Probabilities', in: *Advances: Aggregate Probabilities (concept)*, 2006, hoofdstuk 9.

Figuur 2. Individuele en collectieve kansinschatting



Bron: Robert T. Clemen en Robert L. Winkler, 'Aggregation of Expert Probabilities', in: *Advances: Aggregate Probabilities (concept)*, 2006, Hoofdstuk 9.

16. Tversky and Kahneman (1974), zie noot 9.

17. Elizabeth F. Loftus, 'Eyewitness Testimony: Psychological Research and Legal Thought', *Crime and Justice*, 3, 1981, pp. 105-151.

18. Colin F. Camerer en Eric J. Johnson, 'The Process-Performance Paradox in Expert Judgment: How Can Experts Know So Much and Predict So Badly?', in: K.A. Ericsson en J. Smith (red.), *Towards a General Theory of Expertise: Prospects and Limits*, Cambridge Press, New York, 1991, pp. 195-217.

19. P.E. Tetlock, *Expert Political Judgment: How Good Is It? How Can We Know?*, Princeton University Press, Princeton (New Jersey), 2003.

derlijke verzamelde inschattingen de beste of bijna beste resultaten te geven. De resultaten met de overlegbenadering zijn minder eenduidig. Met de bekende Delphi-methode lijken even goede resultaten te worden geboekt als met de overige overlegmethoden, maar geen betere.

Figuur 2 geeft een voorbeeld uit de praktijk: de inschattingen van een aantal deskundigen en de aggregatie daarvan, in dit geval simpelweg een gemiddelde. (Het gaat hier om een andere variabele dan in figuur 1, te weten de maximale beweging van de bodem bij een aardbeving.) In dit voorbeeld worden de afzonderlijke inschattingen van drie deskundigen getoond. De uitkomst is de groepsinschatting, afgeleid van de verzamelde afzonderlijke meningen. Evenals afzonderlijke kansdistributies kan deze geaggregeerde distributie direct worden gebruikt bij de strategische besluitvorming.

Beperkingen

De besluitvormingstheorie ('decision theory') is in wezen het beste instrument tot nu toe om de onzekerheid bij een strategische beslissing kwantitatief in te schatten. Deze benadering heeft echter enkele belangrijke beperkingen.

Een eerste beperking is dat bij een methode die vertrouwt op individueel oordeel altijd rekening moet worden gehouden met de persoonlijke 'kleuring' van dat oordeel. Die vertekening is uitgebreid

gedocumenteerd; zoals gezegd hebben Tversky en Kahneman al jaren geleden in hun baanbrekende onderzoek de verschillende vormen ervan benoemd en gekarakteriseerd.¹⁶ Meer recentelijk hebben andere auteurs beschreven hoe het individuele oordeel in uiteenlopende omstandigheden te kort kan schieten. Loftus onderzocht de problemen waar mensen tegenaan lopen wanneer ze zich zelfs maar de eenvoudigste dingen proberen te herinneren.¹⁷ Camerer en Johnson¹⁸ en Tetlock¹⁹ laten zien dat het zelfs zogeheten 'deskundigen' niet veel beter af gaat in een medische respectievelijk politieke context. Formele kanscodering is bedoeld om af te rekenen met veel van de vooringenomenheid die het individuele oordeel kan kleuren. Er is echter geen degelijke theorievorming voor het kwaliteitsniveau van de inschattingen die deze aanpak oplevert – en vooral in hoeverre die overeenstemmen met wat er uiteindelijk in de praktijk gebeurt. Het ontbreekt aan vaste principes die aangeven in hoeverre de vertekening met een bepaalde coderingsmethode kan worden verkleind; of in hoeverre procedures moeten worden toegepast om een bepaald kwaliteitsniveau te bereiken. Vooraanstaande onderzoekers en gebruikers vinden het gebrek aan een degelijke onderbouwing een belangrijke tekortkoming van de huidige aanpak – en dus een belangrijk verbeteringspunt. In de woorden van Ferrell: 'Subjectieve kansinschatting en combinaties daarvan kunnen

alleen met vertrouwen worden gebruikt als er een betere theoretische onderbouwing komt.²⁰

Een tweede beperking geldt het empirisch bewijs, los van de theoretische onderbouwing, voor de kwaliteit van de kansinschattingen. Dat is nogal schaars en wisselend van aard. Er zijn zeker aanwijzingen dat met formele methoden betere resultaten worden geboekt dan met de informele. Browne e.a. melden bijvoorbeeld dat formele kanscodering aantoonbaar tot betere resultaten leidt.²¹ Er is echter weinig bewijs voor de kwaliteit van de inschattingen die formele kanscodering oplevert (iets anders dan de geconstateerde verbetering). Bovendien gaat het bij een groot deel van het bewijs om zeer specifieke toepassingen en instrumenten. Daardoor zijn algemene conclusies moeilijk te trekken. Een van de omvangrijkere evaluaties van kwantitatieve kansinschattingen is van Cooke; dit betreft vooral wetenschappelijk onderbouwde risicoanalyses met kansinschatting ('probabilistic risk analysis', PRA).²² Zijn bevindingen zijn veelzeggend. Hij bericht van een evaluatie van 84 empirische studies waarvan er '47 slechte resultaten meldden, en 37 goede resultaten'. Verder merkt hij op dat slechte resultaten in literatuurverwijzingen veruit in de meerderheid zijn (het aantal slechte citaten is 'een factor 6 groter'). Al met al komt Cooke tot de conclusie dat 'de meningen van deskundigen gepeild bij risicoanalyse met kansinschatting bijzonder breed gespreid zijn, enige clustering vertonen en matig reproduceerbare en kalibreerbare uitkomsten opleveren.' Desondanks blijft hij de methode steunen, vanuit de gedachte dat 'het glas half vol is'.

Een derde beperking betreft de geloofwaardigheid van deze methode in een organisatie en de verdedigbaarheid ervan tegenover de buitenwacht. Die staan ter discussie, wellicht als gevolg van de twee bovengenoemde beperkingen. Intern kunnen beleidsmakers niet makkelijk rechtvaardigen dat zij zich mede laten leiden door een volgens de beschreven methode gekwalificeerd oordeel. Dit is vaak vooral een probleem bij organisaties waar beslissingen doorgaans worden onderbouwd met (meer geloofwaardige) reeksen aan gegevens uit het verleden. Ook extern blijkt er vaak nauwelijks tot geen draagvlak voor te zijn onder belanghebbenden. Zo stelt Bier dat de trage invoering van risicoanalyse met kansinschatting (PRA) ten behoeve van de regelgeving voor kernenergie 'grotendeels te wijten is aan de subjectiviteit van PRA-uitkomsten'.²³ Onder juristen woedt er al minstens een eeuw een debat over de rol van getuige-

deskundigen, en het inschatten van kansen in het bijzonder, bij zowel civiele zaken als strafzaken. Eind negentiende eeuw sprak het Hooggerechtshof van Mississippi onomwonden zijn twijfel uit over subjectieve kansinschatting:

De gerechten ... hebben het leven, de vrijheid en de eigendommen van de burger nu wel genoeg blootgesteld aan de louter speculatieve meningen van heren die zich deskundig noemen in wetenschappelijke zaken en wier eigendunk in dezen vaak gelijk opgaat met hun onwetendheid. Wij zijn niet bereid, de rechtspraak over te geven aan hypothesen, gegis en de wereld van waarschijnlijkheden dan wel onwaarschijnlijkheden, en iets als bewijs toe te laten wat in de essentie niet anders dan volstrekt onzeker en onbevredigend kan zijn.²⁴

De rol van kans en kansberekening in rechtszaken is recentelijker tamelijk uitvoerig onderzocht en bediscussieerd, ook bij strafzaken die sterk in de publieke belangstelling stonden.²⁵ Naarmate beleid (publiek én privaat) meer blootstaat aan de openbaarheid, wordt het belangrijker om de gehanteerde methode voor kwantitatieve onzekerheidsinschatting te kunnen rechtvaardigen en verdedigen. Deze beperkingen worden op de koop toe genomen als men zegt dat de huidige aanpak weliswaar niet perfect is, maar wel de beste die we hebben. Dit is duidelijk de invalshoek die Morgan en Henrion hanteren: 'Het enige wat we kunnen doen, is voorzichtig te zijn en ons voortdurend te blijven realiseren dat een deskundige er ver naast kan zitten – maar ook dat we vaak niets anders hebben om op af te gaan.'²⁶ Dit brengt veel beleidsmakers en belanghebbenden in een lastig parket. Zij zouden de onzekerheid rond hun strategische beslissingen graag preciezer willen benoemen. Maar zij aarzelen om dat te doen met een methode die mogelijk ernstige beperkingen heeft. Die terughoudendheid is des te groter wanneer een beslissing ten overstaan van de samenleving moet worden genomen.

Voorspellingsmarkt

Ondanks alle beperkingen boden de genoemde instrumenten van de besluitvormingstheorie – kanscodering en groepspeiling – tot voor kort in de regel de beste manier waarop beleidsmakers en belanghebbenden de onzekerheid rond een oordeel konden benoemen. In de afgelopen jaren is er echter een nieuwe aanpak bij gekomen, die wellicht een betere grondslag biedt voor besluitvor-

20. William R. Ferrell, 'Discrete Subjective Probabilities and Decision Analysis: Elicitation, Calibration, and Combination', in: G. Wright en P. Ayton (red.), *Subjective Probability*, John Wiley and Sons, New York, 1994, pp. 411-451, op p. 448.
21. Glenn J. Browne, Shawn P. Curley en P. George Benson, 'Evoking Information in Probability Assessment: Knowledge Maps and Reasoning-Based Direct Questions', *Management Science*, 43/1, januari 1997, pp. 1-14.
22. Roger Cooke, *Experts in Uncertainty: Opinion and Subjective Probability in Science*, Oxford University Press, Oxford (VK), 1991.
23. Vicki M. Bier, 'Challenges to the Acceptance of Probabilistic Risk Analysis', *Risk Analysis*, 19/4, augustus 1999, pp. 703-710.
24. Henry Wade Rogers, *The Law of Expert Testimony*, Central Law Journal Company, St. Louis (Missouri), 1891.

Op de Hollywood Stock Exchange zetten deelnemers speelgoedgeld in op de 'top' of 'flop' van speelfilms

25. D.H. Kaye, 'The Laws of Probability', in: Peter Murphy (red.), *Evidence. Proof, and Facts: A Book of Sources*, Oxford University Press, Oxford (VK), 2003, pp. 529-546.

26. Granger Morgan en Max Henrion, *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*, Cambridge University Press, Cambridge (VK), 1992, p. 137.

27. Een uitstekende beschouwing van deze benadering is te vinden bij Justin Wolfers en Eric Zitzewitz, 'Prediction Markets', *Journal of Economic Perspectives*, 18/2, lente 2004, pp. 107-126.

28. Zie voor een bespreking van de ontwerpmogelijkheden en het effect ervan op de uitkomsten van het vergaren van marktinformatie: Charles Plott, 'Markets as Information Gathering Tools', *Southern Economic Journal*, 67/1, juli 2000, pp. 1-15.

29. James Surowiecki, *The Wisdom of Crowds: Why the Many are Smarter than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations*, Doubleday, New York, 2004.

ming op basis van een oordeel waar onzekerheid aan kleeft. Deze aanpak is vooral bekend geworden onder de naam 'voorspellingsmarkt' ('prediction market'), hoewel namen als voorspellings-, informatie- en besluitvormingsmarkt ook worden gebruikt.²⁷

Een voorspellingsmarkt is in essentie een georganiseerde plek waar mensen kunnen inzetten op een bepaalde uitkomst of een bepaald belang, bijvoorbeeld een politieke of maatschappelijke gebeurtenis. De markt is gesloten wanneer die gebeurtenis zich voordoet; vervolgens wordt de uitkomst bekend, en zijn er winnaars en verliezers. De inzetten die worden gedaan, bepalen de prijs voor een contract verbonden aan de uitkomst. Van die prijs kan direct of indirect worden afgeleid hoe groot de deelnemers de kans op de betreffende gebeurtenis inschatten: dat is hun voorspelling. In een voorspellingsmarkt zou bijvoorbeeld kunnen worden ingezet op de kans dat een Democraat de Amerikaanse presidentsverkiezingen van 2012 wint. Iemand die daarop wil inzetten, koopt dan een contract dat één dollar oplevert als een Democraat die verkiezingen wint. Totdat de uitslag van de presidentsverkiezingen van 2012 bekend zijn, zal de koers van dat contract iets onder één dollar liggen of hooguit één dollar bedragen; het contract levert immers hooguit één dollar op. Onder geschikte omstandigheden blijkt uit de evenwichtsprijs in deze markt, laten we zeggen zestig dollarcent, hoe groot de deelnemers de kans achten dat een Democraat wint in 2012. In dit voorbeeld is dat zestig procent (dit percentage is hier overigens volstrekt willekeurig gekozen).

De geschiedenis van de voorspellingsmarkt begint iets meer dan twintig jaar geleden: in 1988 lanceerde de universiteit van Iowa een elektronische markt, de Iowa Electronic Market (IEM), waar mensen konden inzetten op de presidentsverkiezing van dat jaar. Deze IEM is sindsdien op bescheiden voet in bedrijf gebleven en uitgebreid.

Sinds de lancering van de IEM is de wetenschappelijke en commerciële belangstelling voor de voorspellingsmarkt als instrument gegroeid. Het aantal van dergelijke markten, en het bereik ervan, is drastisch toegenomen. Er bestaan inmiddels tal van variaties op dit thema. Soms wordt er echt geld ingezet, soms virtueel of speelgoedgeld. De ene keer gaat het om een eenvoudig contract, zoals in

het zojuist gegeven voorbeeld, in andere gevallen om meer complexe contracten.²⁸ Een zeer belangrijk kenmerk is dat sommige markten algemeen en andere slechts beperkt toegankelijk zijn.

Er bestaat op dit moment een handjevol algemeen toegankelijke (openbare) voorspellingsmarkten. Iedereen kan er zijn gokje wagen. Deze markten benutten op deze manier ten volle de 'kennis van de massa' ('wisdom of crowds'); dat wil zeggen dat dit type markt optimaal functioneert wanneer er grote aantallen mensen aan deelnemen.²⁹ Een openbare markt is doorgaans gericht op politieke, maatschappelijke en economische gebeurtenissen die in de belangstelling staan van het grote publiek. Algemeen toegankelijke markten waar echt geld wordt ingezet, staan onder serieus toezicht van de regelgever – in het bijzonder in de Verenigde Staten. Dit toezicht heeft hun ontwikkeling belemmerd en heeft ook voor enige onrust gezorgd in deze nieuwe branche. Partijen die zo'n openbare markt verzorgen doen dat veelal omwille van de informatie die ze daarmee verkrijgen. Die kunnen ze vervolgens gebruiken voor hun professionele of commerciële doeleinden. Anderen is het vooral te doen om het geld dat de markt zelf oplevert. Hoewel er niets met zekerheid over te zeggen valt, lijken de deelnemers aan een openbare markt vooral gemotiveerd te worden door een combinatie van nieuwsgierigheid, imago en geld.

Treffende voorbeelden van een openbare markt buiten de IEM zijn Intrade, de Hollywood Stock Exchange en de CFO Prediction Market. Intrade, gevestigd in Ierland, is momenteel de koploper onder de openbare voorspellingsmarkten. Zozeer zelfs dat deze instelling zichzelf 'de voorspellingsmarkt' noemt. Op Intrade wordt echt geld ingezet in allerlei verschillende categorieën, waaronder kunst, bedrijfsleven, actualiteiten, amusement, politiek, wetenschap en het weer. Je kunt er bijvoorbeeld wedden op de kans dat er een wet komt op de beperking van de uitstoot van broeikasgassen en de handel in uitstootrechten. Of je kunt inzetten op de kans dat iemand aantoonbaar dat koude kernfusie inderdaad mogelijk is. Net als de meeste beurzen en commissarissen verdient Intrade een kleine provisie voor elke transactie. Om nog even bij de Amerikaanse presidentsverkiezingen van 2012 te blijven: figuur 3 laat zien dat de handel op Intrade op 25 oktober 2009 de kans op een Democratische overwinning inschatte op 65 procent. De Hollywood Stock Exchange ('beurs van Hollywood') heeft een veel kleiner aandachtsgebied. Daar zetten de deelnemers speelgoedgeld in op de 'top of flop' van speelfilms; hun opbrengst kunnen ze inwisselen voor prij-

Figuur 3. Koersontwikkeling op Intrade van propositie Democraat wint Amerikaanse presidentsverkiezingen van 2012 (1 nov 2008 – 25 okt 2009)



zen. De eigenaren van de Hollywood Stock Exchange verkopen de informatie aan de amusementswereld. De CFO Prediction Market, ten slotte, wordt verzorgd door de website CFO.com van het blad *CFO magazine*. Hier kan de financiële manager inzetten op variabelen die bij uitstek interessant zijn voor deze beroepsgroep, zoals de economische groei. Het tijdschrift wil hiermee vooral publiciteit genereren. De winnende deelnemers krijgen een bescheiden prijs, maar mensen doen vooral mee omdat ze hopen in hun beroepsgroep als de allerbeste uit de bus te komen.

Besloten voorspellingsmarkten zijn er veel meer. Die zijn uitsluitend toegankelijk voor de leden van een bepaalde groep. Veelal gaat het dan om het personeel van een onderneming, of een deel daarvan met bepaalde kwalificaties. Er wordt dan ‘gehandeld’ in gebeurtenissen waarin die groep sterk geïnteresseerd is, zoals de voltooiing van een belangrijk project of de lancering van een nieuw product door een concurrent. De partij die zo’n markt aanbiedt, is uitsluitend geïnteresseerd in de informatie die het oplevert. Die informatie wordt direct doorgegeven aan de relevante beleidsmakers in de organisatie. Ook nu weer is er niet veel met zekerheid over te zeggen, maar de belangrijkste motieven lijken de loyaliteit jegens de groep, imago en nieuwsgierigheid te zijn. Deelnemers wedden met namaakgeld en kunnen prijzen verdienen. Verscheidene organisaties vertellen openlijk dat zij intern met dit instrument werken. Het lijkt vooral populair te zijn bij technologiebedrijven in Silicon Valley. Google heeft er honderden georganiseerd, aldus Bo Cowgill, een productmanager die er verantwoordelijk voor is.³⁰ De belangrijkste onder-

werpen zijn het prestatieniveau van de onderneming, hoe nieuwe producten het doen en trends in de bedrijfstak. Google biedt dit soort markten ook aan als vorm van amusement, bijvoorbeeld over sport en andere evenementen; die dienen vooral ter ontspanning en als sociaal bindmiddel. BestBuy, General Electric, HP, Yahoo en allerlei andere vooraanstaande ondernemingen hebben ook openlijk verteld dat zij werken met een voorspellingsmarkt. De beroemdste (of beruchtste) voorspellingsmarkt van de overheid was een wedplek die het ministerie van Defensie wilde verzorgen over politieke gebeurtenissen. Dit initiatief ging niet door omdat de publieke opinie en het parlement erop tegen waren dat mensen zouden gokken op terroristische aanslagen, moord en dergelijke.

Toepassing bij strategievorming

Er is veel enthousiasme over de bijdrage die een voorspellingsmarkt kan leveren aan de strategische besluitvorming.³¹ Ook is er al een kleine maar groeiende onderzoeksliteratuur met een conceptueel of experimenteel karakter, over de inrichting en uitvoering van een voorspellingsmarkt ter ondersteuning van strategische beslissingen.³² Zoals al eerder opgemerkt, is de wetenschappelijke en commerciële belangstelling voor voorspellingsmarkten sterk gegroeid sinds de jaren tachtig. Er zijn inmiddels behoorlijk wat studies en experimenten over uitgevoerd. Daarbij is gekeken naar de markten zelf en naar de voorwaarden waaronder ze in verschillende contexten al dan niet succesvol zijn. Uit dit onderzoek kunnen drie criteria worden afgeleid die bij uitstek van belang lijken te zijn bij strategische besluitvorming. Het eerste criterium betreft het ontwerp van de

³⁰ Renee Dye, ‘The Promise of Prediction Markets: A Roundtable’, *McKinsey Quarterly*, 2, 2008, pp. 84-93.

Onder andere Google, BestBuy, General Electric, HP en Yahoo gebruiken de voorspellingsmarkt

markt zelf en de contracten die erin verhandeld worden. Er zijn veel denkbare 'smaken' voor een voorspellingsmarkt. Er zijn verschillende methoden om te bepalen hoe en hoeveel er kan worden ingezet; hoe vraag en aanbod bij elkaar kunnen worden gebracht; en wat voor beloning de winnaar moet krijgen. Met het oog op strategische besluitvorming is het heel belangrijk dat de contractprijzen zonder meer kunnen worden geïnterpreteerd als een collectieve inschatting van de kans op een gebeurtenis of ontwikkeling. Bovendien moet die inschatting zo accuraat en geloofwaardig mogelijk zijn. Een eenvoudige aanpak die goed werkt bij een reeds bestaande markt is in een andere context misschien ongeschikt.

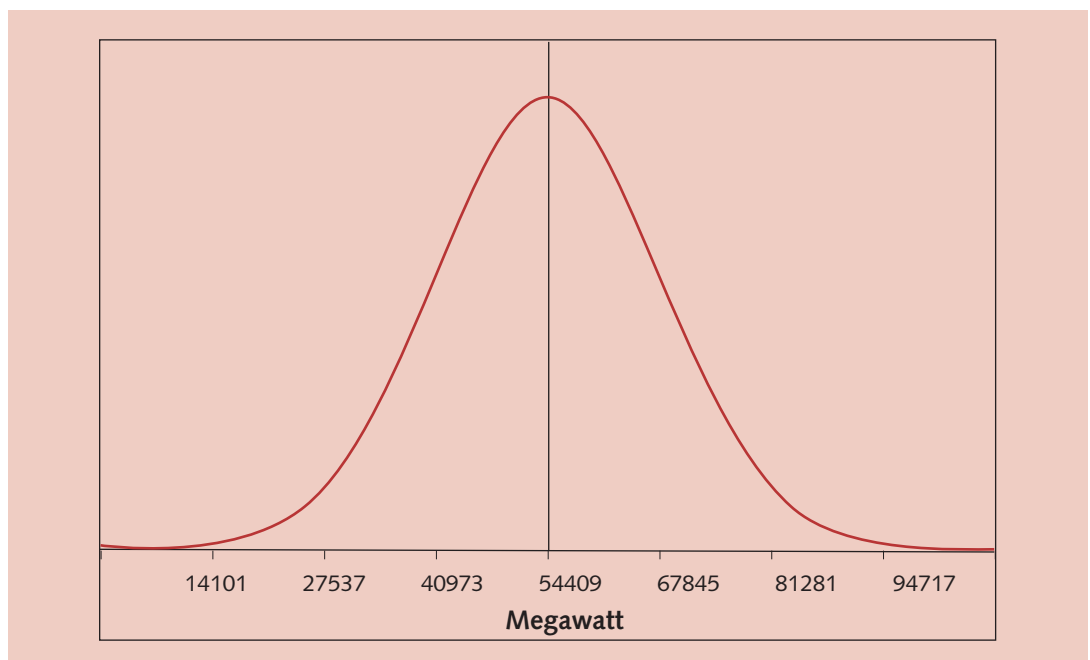
Het tweede criterium betreft de variabelen waarin kan worden 'gehandeld'. Die zullen zodanig moeten worden gedefinieerd dat ze voldoen aan twee mogelijk tegenstrijdige doelen. Ze dienen helder en relevant te zijn voor de deelnemers, want anders zijn de uitkomsten kwalitatief onvoldoende. En ze dienen direct aan te sluiten bij de besluitvorming waar het allemaal om begonnen is, anders hebben de uitkomsten te weinig nut. Bij de bestaande openbare voorspellingsmarkten geeft uitsluitend de eerste overweging de doorslag. Bij bestaande besloten voorspellingsmarkten

heeft het tweede criterium de boventoon gevoerd, wat soms ten koste is gegaan van de belangstelling voor de markt en het activiteitsniveau. Een goed voorbeeld van dat laatste is handel in ramingen voor verder in de toekomst gelegen gebeurtenissen. Een organisatie die voor een strategische keuze staat, zal misschien graag willen weten hoe een nieuwe technologie of markt zich op de lange duur ontwikkelt. Maar de deelnemers aan de voorspellingsmarkt zullen daar misschien niet op inzetten. Het duurt immers nog jaren voordat ze de uitslag te horen krijgen. Dit kan worden ondervangen met een zorgvuldig samengestelde portfolio van variabelen op korte respectievelijk lange termijn in combinatie met een doordachte inrichting van de markt.

Het derde criterium betreft de keuze van de deelnemers aan de markt. Veel bestaande markten, zowel openbare als besloten, hebben grote aantallen (honderden of duizenden) deelnemers nodig voor de benodigde liquiditeit en diversiteit. Voor de strategische besluitvorming zijn die grote aantallen echter misschien ondoenlijk of onwenselijk. Dan is het zaak, een betrekkelijk kleine groep deelnemers te vinden. Soms is een twintigtal mensen misschien al voldoende – mits ze voldoende kennis hebben en tegelijkertijd voldoende divers zijn. In de praktijk betekent dit, dat een voorspellingsmarkt ten behoeve van de strategische besluitvorming niet

31. Het enthousiasme over de mogelijkheden van voorspellingsmarkten voor de strategische besluitvorming blijkt onder meer uit recente artikelen van Teck-Hua Ho en Kay-Yut Chen in de *California Management Review* (Ho en Chen, 2007), *The Economist* (2009), de *McKinsey Quarterly* (Dye, 2008) en *Risk* (Patel, 2006). Aanvullende onderbouwing wordt verschaft in Teck-Hua Ho en Kay-Yut Chen, 'New Product Blockbusters: The Magic and Science of Prediction Markets', *California Management Review*, 50/1, herfst 2007, pp. 144-158; 'An Uncertain Future: Prediction Markets', *The Economist*, 28 februari 2009, p. 68; Dye (2008), zie noot 30; Navroz Patel, 'Deciding on a Future', *Risk*, 19/2, februari 2006, pp. 33-35; Peter A. Coles, Karim R. Lakhani en Andrew McAfee, 'Prediction Markets at Google', Harvard Business School Case #607088-PDF-ENG, 30 mei 2007.

Figuur 4. Indeling (semi-)publieke sector



Bron: Voorspellingsmarkt NERA

openbaar kan zijn. Maar evenmin dient deze markt beperkt te blijven tot één enkele organisatie. Men zal een combinatie moeten vinden van gekwalificeerde individuen met uiteenlopende achtergronden, ervaringen, invalshoeken en loyaliteiten.

Onze onderneming, National Economic Research Associates (NERA), heeft onlangs een voorspellingsmarkt ingericht die wij de Clean Energy Exchange (beurs voor schone energie) noemen. Deze markt draait op software van Crowdcast, een toonaangevende leverancier op dit gebied. Er nemen zowel mensen van NERA zelf als van buiten NERA aan deel. Deze markt is bedoeld om informatie te genereren die de deelnemers kunnen gebruiken bij risicomanagement en strategische planning. Tot nu toe gaat de aandacht uit naar onderwerpen waar de energiesector in geïnteresseerd is, zoals kern- en zonne-energie, het broeikasbeleid en grondstoffenprijzen. De markt is ingericht en wordt beheerd conform de hierboven beproven criteria; we werken aan oplossingen voor de uitdagingen die we daarbij tegenkomen. Figuur 4 toont de status op 25 oktober 2009 voor een marktvariabele van onze markt: de totale opwekkingscapaciteit aan het eind van 2011 van kerncentrales in aanbouw. Deze voorspellingsmarkt is opgezet met de uitdrukkelijke bedoeling om informatie te genereren in de vorm van een kansdistributie. In dit specifieke voorbeeld blijkt er een consensusprognose te bestaan onder de deelnemers van iets meer dan 54.000 (megawatt) MW. De laagste schatting is 14.000 MW, de hoogste 95.000 MW; beide hebben een kans van ongeveer één procent.

Voordelen

Met een doordachte inrichting en uitvoering langs de bovengeschetste lijnen krijgt een voorspellingsmarkt drie voordelen die een antwoord lijken te bieden op de eerder genoemde beperkingen van de huidige aanpak op basis van de besluitvormingstheorie. Deze voordelen hebben betrekking op de theoretische onderbouwing, de empirische bewijslast en de geloofwaardigheid/verdedigbaarheid.

Theoretische onderbouwing. In theoretisch opzicht onderscheidt de voorspellingsmarkt zich door drie belangrijke kenmerken van de bestaande aanpak. 1) Er worden werkelijke in plaats van denkbeeldige keuzen gemaakt. Bij de tot nog toe heersende aanpak zeggen mensen wat zij zouden kiezen of waar zij een zekere waarde aan zouden toekennen als zij de keuze hadden. De deelnemer aan een voorspellingsmarkt maakt die keuze ook echt en kent er daadwerkelijk waarde aan toe: geen woorden maar daden.

In principe is dit een betere graadmeter voor wat de deelnemer werkelijk denkt.

2) De relatieve stelligheid of overtuiging van iedere deelnemer wordt direct geregistreerd en expliciet afgemeten aan de hoeveelheid geld die hij/zij inzet. Bij de huidige benadering wordt die overtuiging helemaal niet gemeten of enkel door de betrokkene zelf aangegeven. In principe biedt de voorspellingsmarkt hiermee een betere graadmeter voor de relatieve kennis van de afzonderlijke deelnemer. 3) De voorspellingsmarkt kan putten uit de kracht van 'rationele verwachtingen'. Dit betekent dat de marktresultaten niet stelselmatig worden vertekend door vooroordelen (want anders zou er een kans zijn op een risicovrij rendement). De tot nog toe gangbare manier van werken is zeker niet vrij van vertekening. In principe leidt dit op den duur tot kwalitatief betere inschattingen.

Empirische bewijslast. In empirisch opzicht is er een vrij klein maar groeiend corpus aan bewijs dat voorspellingsmarkten veel betere resultaten kunnen opleveren dan andere prognosetechnieken. Er is bij herhaling gemeld dat de IEM trefzekerder was dan de meeste opiniepeilingen.³³ Dit is bemoedigend, maar het betekent niet dat een uitkomst van de IEM grotere voorspellende waarde biedt dan een groep politieke deskundigen die door middel van formele kanscodering worden geraadpleegd. Ook is gerapporteerd dat besloten voorspellingsmarkten het bij ondernemingen als Google en HP hebben gewonnen van deskundigen.³⁴ Dit is ook bemoedigend, maar we weten niet hoe die deskundigen waren geraadpleegd. Het statistisch bewijs, in mindere mate voorhanden dan anekdotes, is eveneens bemoedigend. Pennock e.a. rapporteerden bijvoorbeeld in een ingezonden brief aan het tijdschrift *Science* over de uitkomsten van onderzoek naar openbare voorspellingsmarkten: zij stelden dat de inschattingen ontleend aan deze markten goed 'gekalibreerd' waren, oftewel in hoge mate overeen kwamen met de feitelijke gebeurtenissen.³⁵

Geloofwaardigheid en verdedigbaarheid. De voorspellingsmarkt heeft belangrijke kenmerken die dit instrument geloofwaardiger en beter verdedigbaar kunnen maken dan de tot nog toe gehanteerde aanpak. De huidige benadering leunt zwaar op discussies één op één met een individu of een groep. Rechtbanken, regelgevers en het publiek twijfelen aan de inschattingen die dergelijke discussies opleveren. Het argument dat we 'het ermee zullen moeten doen' klinkt niet erg overtuigend. Tegelijkertijd lijkt er ondanks de financiële crisis nog steeds een vrij groot vertrouwen te zijn in het

32. Zie voor een vergelijking van voorspellingsmarkten en de Delphi-methode: Kesten C. Green, J. Scott Armstrong en Andreas Graefe, 'Methods to Elicit Forecasts from Groups: Delphi and Prediction Markets Compared', *Foresight: The International Journal of Applied Forecasting*, 8, herfst 2007, pp. 17-20.

marktmechanisme en in marktprijzen als methode voor het opstellen van prognoses. Het is nog vroeg dag, maar wellicht duidt dit erop dat kansinschattingen met behulp van voorspellingsmarkten ooit een groter draagvlak zullen krijgen dan de huidige aanpak – zowel in organisaties als daarbuiten.

Beperkingen

Natuurlijk heeft de voorspellingsmarkt als instrument ook beperkingen. Het instrument wordt pas sinds kort breder toegepast. Voor diverse conceptuele en praktische aspecten zijn daarom nog geen volledig antwoorden gevonden.

Een eerste beperking is simpelweg de ‘andere kant’ van het muntje dat ‘markt’ heet. Men kan allerlei overtuigende empirische en theoretische redenen hebben om de markt een groot voorspellend vermogen toe te dichten. Maar er is ook alle reden om markten met enige of zelfs veel argwaan te bekijken. Steeds meer gedragseconomisch onderzoek wijst op dat laatste voor conventionele markten. Thaler benoemt vele van de belangrijkste onderwerpen op dit gebied.³⁶ Ook is er vrij veel gepubliceerd over foutieve prognoses in andere relevante contexten, zoals totalisatorwedden bij paardenraces.³⁷ Al met al lijkt voorzichtigheid geboden: de markt levert vaak nuttige informatie op, maar ze is niet perfect en kan er zelfs volstrekt ‘naast’ zitten.

Een tweede beperking betreft de mogelijkheid van vertekening – niet zozeer door onbewuste vooringenomenheid maar door doelbewuste manipulatie. Ook hiervoor is zowel empirisch als theoretisch bewijs te geven. In een traditionele markt bestaat er een sterke directe prikkel om de markt te manipuleren: er valt veel geld mee te verdienen. Deze financiële prikkel ontbreekt of is bijzonder klein bij nagenoeg iedere voorspellingsmarkt. Daar wordt immers met zeer kleine bedragen (of zelfs helemaal geen echt geld) gewerkt. Er kunnen echter ook andere motieven dan geld zijn om een voorspellingsmarkt te manipuleren. Als de informatie uit een (openbare dan wel besloten) voorspellingsmarkt wordt gebruikt bij het bepalen van commercieel of publiek beleid, kan het voor de betrokken organisaties interessant zijn om die informatie te manipuleren. Bij een besloten voorspellingsmarkt (van één enkele onderneming) schatten individuele of groepen medewerkers de lanceringsdatum of verwachte omzet van een bepaald product mis-

schien mede in op basis van hun eigen belang. Bij een openbare voorspellingsmarkt levert een deelnemende onderneming misschien valse informatie omdat ze vermoedt dat een concurrent die informatie gebruikt bij zijn besluitvorming. De prikkel om een voorspellingsmarkt te manipuleren is misschien minder groot dan bij een conventionele markt. Daar staat echter tegenover dat het makkelijker is om zo’n markt te manipuleren. Opnieuw geldt: voorzichtigheid is geboden.

Een derde beperking, ten slotte, betreft de inzet van het instrument voor verkenningen van de verdere toekomst. Juist die zijn belangrijk bij strategievorming. De voorspellingsmarkt leent zich echter het beste voor een gebeurtenis die zich vrij snel (binnen ongeveer een jaar) zal voordoen en die duidelijk kan worden afgebakend (gebeurt iets wel of niet?). De korte termijn maakt de deelnemers alert en gemotiveerd; zij kunnen zich uitspreken over een concrete gebeurtenis (in plaats van de richting van de aanloop ernaartoe). Duidelijke afbakening betekent dat de kans op de gebeurtenis in kwestie kan worden afgelezen aan één enkele inzet. De marktkoers van het contract is de consensus omtrent de kans. Een markt voor langetermijnvragen (bijvoorbeeld: hoeveel kost technologie voor kernenergie in 2020?) is lastiger in te richten dan voor een vraag op korte termijn (hoeveel kosten zonnepanelen over een jaar?). Ook is een markt voor het inschatten van een doorlopende variabele als de prijs voor CO₂-emissierechten moeilijk te ontwerpen dan een markt voor de vraag of er wetgeving zal komen die de uitstoot van CO₂ reguleert. Dit is een probleem voor de voorspellingsmarkt als instrument ter ondersteuning van de strategische besluitvorming. Wellicht kunnen onderzoek en experimenten hier oplossingen voor aandragen.

Conclusie

Een ‘nieuwe muizenval’ is altijd spannend, vooral als die een groot positief effect kan hebben. Daarvan lijkt sprake te zijn bij de voorspellingsmarkt en de toepassing ervan bij de strategische besluitvorming. Natuurlijk mogen we ons niet van de wijs laten brengen door een nieuw speeltje, en moeten we oog houden voor de beperkingen van dit nieuwe instrument. Toch is onze inschatting dat de voorspellingsmarkt de kwaliteit van strategische beslissingen in de toekomst sterk zal helpen verbeteren.

33. Joyce E. Berg, Forrest D. Nelson en Thomas A. Rietz, ‘Prediction Market Accuracy in the Long Run’, *International Journal of Forecasting*, 24/2, april-juni 2008, pp. 285-300.

34. Kay-Yut Chen en Charles Plott, ‘Information Aggregation Mechanisms: Concept, Design, and Field Implementation for a Sales Forecasting Problem’, *California Institute of Technology Social Science Working Paper 1131*, maart 2002.

35. David M. Pennock, Steve Lawrence, C. Lee Giles en Finn Arup Nielsen, ‘The Real Power of Artificial Markets’, *Science*, 291, 9 februari 2001, pp. 987-988 (ingezonden brieven).

36. R.H. Thaler (red.), *Advances in Behavioral Finance*, Russell Sage Foundation, New York, 1993.

37. R.H. Thaler en W.T. Ziemba, ‘Anomalies: Parimutuel Betting Markets: Racetracks and Lotteries’, *Journal of Economic Perspectives*, 2/2, lente, 1988, pp. 161-174.

Translated and reprinted by permission of California Management Review. This article was originally published under the English title ‘Prediction Markets: A New Tool for Strategic Decision Making’, by A. Borison & G. Hamm, in the Summer 2010 issue of the California Management Review, Vol. 52, No. 4, pp. 125-141. Copyright © 2010 by The Regents of the University of California; all rights reserved. Vertaling: Raymond Gijzen.