



De essentie van prestatimaatstaven

Auteur(s):

Cools, K.

Praag, C.M. van.

Kees Cools is werkzaam bij The Boston Consulting Group in Baarn en aan de Rijksuniversiteit Groningen. Mirjam van Praag is werkzaam aan de Universiteit van Amsterdam.

Verschenen in:

ESB, 88e jaargang, nr. 4394, pagina 56, 7 februari 2003

Rubriek:**Trefwoord(en):**

belonen

Tussen een beloningsmaatstaf die zo zuiver mogelijk meet wat de inzet van een werknemer is, en een maatstaf die weergeeft in hoeverre die inzet het ondernemingsdoel dient, bestaat een afruil.

Dat verklaart sommige missers bij het meten en belonen van prestaties.

In opdracht van 'de baas' moet iedere werknemer, van de voorzitter van de Raad van Bestuur tot de magazijnbediende, bepaalde prestaties verrichten. Omdat de meeste werknemers niet uit zichzelf precies datgene doen wat van ze verwacht wordt en om arbeidsprestaties adequaat te kunnen belonen, moeten er prestatimaatstaven worden vastgesteld. Daarmee kan periodiek worden bepaald welke mate van inspanning de werknemer heeft verricht en hoe effectief deze inspanning is. Bovendien heeft een juiste, transparante en eerlijke prestatiemeting en -beloning een sterk positief effect op de motivatie en daarmee op de prestaties van werknemers¹. Daarom is de selectie van de juiste prestatimaatstaven een van de belangrijkste instrumenten voor de aansturing van een onderneming. Dit artikel bespreekt de essentie van prestatimaatstaven; wat zijn de belangrijkste criteria en afwegingen bij het kiezen van prestatimaatstaven. De theoretische inzichten worden ondersteund door de resultaten van eigen empirisch onderzoek.

Tekortkomingen van de klassieke theorie

In de klassieke 'agency theorie' over de relatie tussen de werkgever - principaal - en werknemer- agent - wordt de mate van prestatiebeloning bepaald door de afruil tussen de behoefte aan 'verzekering' van de risico-averse agent en de behoefte deze agent te prikkelen om te presteren teneinde de marktwaarde van de onderneming te maximaliseren. Daarbij wordt echter voorbij gegaan aan twee belangrijke fenomenen. In de eerste plaats zijn er ondernemingen waarvoor het niet zo vanzelfsprekend is dat, of hoe, de marktwaarde gemaximaliseerd moet worden: dit geldt voor ondernemingen die niet beursgenoteerd zijn, een onduidelijk of meervoudig doel hebben, of waarvan de marktwaarde sterk fluctueert (afgelopen jaar betrof dat de meeste beursondernemingen). Het is dan natuurlijk erg lastig om aan werknemers te vragen om zich bij hun dagelijks werk voortdurend te richten op het maximaliseren van de lange termijn marktwaarde van hun bedrijf. In de tweede plaats, en zeker zo belangrijk, is het effect van de inspanningen van de meeste individuele medewerkers op de marktwaarde of beurskoers zeer beperkt, zo niet verwaarloosbaar. Dat betekent dat de marktwaarde in de meeste gevallen niet gebruikt kan worden als prestatie- dan wel beloningsmaatstaf. Een caissière of magazijnbediende bij Albert Heijn zal niet erg gestimuleerd worden om sneller en beter te gaan scannen of sorteren als hij of zij voortaan wordt beloond met een prestatieloon op basis van de door Ahold behaalde beursresultaten. Het aandelenrendement wordt met name door factoren bepaald die buiten de invloedssfeer van de individuele medewerker liggen, zoals de inzet van de overige 260.000 Ahold-werknemers. Daarom moeten (van de aandelenkoers) afgeleide prestatimaatstaven worden ontwikkeld.

Ruisminimalisering

Een belangrijke overweging bij de selectie van een afgeleide maatstaf is dat het risico of de ruis door externe factoren (olieprijzen, conjunctuur, grondstofprijzen, activiteiten van collega's, enzovoort) of door meetproblemen (hoe meet je bijvoorbeeld kwaliteit?) van de prestatimaatstaf moet worden geminimaliseerd. De maatstaf moet maximaal beïnvloedbaar zijn door de werknemer. Maar welke additionele eisen moeten aan die afgeleide prestatimaatstaven worden gesteld? In onze visie is dat er één.

Doelcongruentie

Naast het minimaliseren van de ruis is het andere cruciale criterium voor een goede afgeleide prestatimaatstaf, dat het effect van de inspanningen van een werknemer op zijn of haar maatstaf zo sterk mogelijk gecorreleerd is met het effect van diezelfde inspanningen op de doelstelling van de onderneming, zoals maximale marktwaarde of aandelenkoers. Wanneer bijvoorbeeld een medewerker van Sony een goed functionerende tv produceert, levert dat een directe en vrijwel constante (maar in de aandelenkoers onzichtbare) bijdrage aan Sony's marktwaarde. Dus het aantal afgeleverde, goed functionerende tv's scoort als prestatimaatstaf voor productiemedewerkers van Sony hoog op dit criterium. Met andere woorden, er moet een maximale doelcongruentie ('alignment') zijn tussen het effect van iemands acties op zijn maatstaf en het effect van die acties op het ondernemingsdoel, in casu de aandelenkoers.

Afruil

De perfecte prestatie maatstaf heeft dus een minimale ruis (is maximaal beïnvloedbaar door de werknemer) en een maximale 'doelcongruentie' (elke stijging of daling van de maatstaf gaat gepaard met een toename respectievelijk afname van de marktwaarde). De wereld van de prestatie maatstaven zou eenvoudig zijn als er maatstaven zijn waarbij én de ruis minimaal is én de doelcongruentie maximaal. Maar helaas, vrijwel alle functies bestaan uit multidimensionale taken, waarbij bijvoorbeeld zowel de kwaliteit dan wel snelheid, technisch vernuft, veiligheid, klantvriendelijkheid, service, samenwerking, maximaal moeten zijn, als de kosten minimaal, en dat niet voor één product maar voor een hele reeks producten en enkele diensten. Onze stelling is dat er voor de prestatie maatstaven van alle niet-eendimensionale functies en taken een afruil bestaat tussen ruis en doelcongruentie. Het is de kunst om precies die maatstaf te vinden waarbij de ruis zo laag mogelijk is en de werknemer dus een zo groot mogelijke prikkel voelt om zijn inspanningen te (kunnen) maximaliseren en waarbij tegelijkertijd zijn acties het ondernemingsdoel zo veel mogelijk ten goede komen (dus de congruentie groot is).

Meer deeltaken maakt afruil belangrijker

Naarmate een functie meer deeltaken omvat en in die zin multidimensionaler is, zal deze afruil belangrijker zijn. Alleen voor heel eenvoudige, eendimensionale taken is de afruil tussen ruis en doelcongruentie beperkt. Maar veel taken die eenvoudig lijken zijn toch meerdimensionaal waardoor de 'perfecte' maatstaf niet bestaat. Bijvoorbeeld niet alleen het aantal klanten dat een caissière of ober bedient is van belang voor de waarde van de onderneming, maar ook de - moeilijk te meten - klantvriendelijkheid; niet alleen het aantal paspoorten dat een douanebeampte controleert is van belang, maar ook de mate van accuratesse, enzovoort. Bekende voorbeelden waarbij relatief te veel aandacht wordt besteed aan ruisminimalisatie en te weinig aan doelcongruentie zijn het belonen van onderwijzers op basis van de testresultaten van leerlingen of het belonen van typisten op basis van het aantal aanslagen per dag. Ook het belonen van programmeurs op basis van het aantal geschreven programmeregels (bij IBM in de jaren zestig) is een goed voorbeeld. Het gevolg van de introductie van deze maatstaven is dat onderwijzers geprikkeld worden alleen nog maar het oefenen van testvaardigheden in de klas aan de orde te stellen, dat typistes tijdens het telefoneren, koffiedrinken en pauzeren blind steeds dezelfde toets aanslaan en dat de softwareprogramma's bij IBM langer (en dus duurder) werden dan noodzakelijk. Zodoende ontbreekt voor verreweg de meeste taken de 'perfecte' maatstaf en moet er altijd wat water in beide wijnvaten worden gedaan: meer dan minimale ruis in het ene vat en minder dan maximale doelcongruentie in het andere. Deze afruil geldt voor de selectie van prestatie maatstaven van alle soorten organisaties, zowel beurs- als niet-beursgenoteerde ondernemingen als not-for-profit instellingen. Voor de laatste twee categorieën bestaat niet eens de mogelijkheid om het aandelenrendement als prestatie maatstaf te hanteren en is het ondernemingsdoel zelf waarschijnlijk niet eens meetbaar.

Directieniveau

In vergelijking tot de werkvloer is op directieniveau doelcongruentie niet belangrijker dan ruisminimalisatie, of omgekeerd. Wel is het zo dat hoger in een organisatie functies relatief complexer (dus meerdimensionaal) worden, waardoor de keuze van de maatstaf zowel belangrijker (want meer van invloed op de marktwaarde) alsook moeilijker wordt. Moeilijker, omdat een groot aantal dimensies van het werk 'op directieniveau' slecht meetbaar zijn (veel ruis) en ook de relatie tussen de inspanningen op al die dimensies en het ondernemingsdoel vaak erg moeilijk is vast te stellen. Een groot aantal problemen en keuzes bij de selectie van prestatie maatstaven in de praktijk kan goed worden geanalyseerd met behulp van de hier besproken elementaire afruil. We laten enkele daarvan, het verschijnsel 'gaming' en de meest gangbare afwegingen bij maatstafselectie (individueel versus team en subjectief versus objectief) de revue passeren.

Gaming

Het fenomeen 'gaming' (cijfermanipulatie) is aan de orde van de dag. Gaming betekent dat een werknemer op door de werkgever onbedoelde wijze de realisatie van de prestatie maatstaf kan verhogen; dit met een positief effect op zijn beloning (de gemeten prestatie neemt immers toe), maar met een heel klein of zelfs negatief effect op de doelstelling van de onderneming. De mogelijkheid voor gaming ontstaat wanneer een prestatie maatstaf onvoldoende doelcongruent is. Wanneer een werknemer in staat is de maatstaf te 'manipuleren' zonder relevante arbeid te verrichten betekent dit immers een lage correlatie tussen het effect van zijn inspanning op de maatstaf en op het ondernemingsdoel. Bekende voorbeelden zijn 'window dressing' door creatief boekhouden (bijvoorbeeld het activeren van lopende kosten of het vrij laten vallen van voorzieningen), het IBM-voorbeeld van te lange softwareprogramma's, het niet leveren van nazorg door medisch specialisten die op basis van aantal operaties worden betaald, geen service verlenen als omzetverantwoordelijke autodealer die steeds na twee jaar wordt overgeplaatst, het doorvoeren van absurde prijsverlagingen als je wordt afgerekend op het aantal verkochte wasmachines en zo verder. Bij al deze voorbeelden geldt steeds dat de gemeten prestatie verbetert, terwijl de onderneming er geen cent wijzer van wordt. Een daarmee samenhangend probleem is dat indicatoren die gevoelig zijn voor gaming een goed beeld kunnen geven van de geleverde prestaties met toegevoegde waarde totdat die indicator tot prestatie maatstaf wordt verheven en mensen zich gaan uitputten om de maatstaf te manipuleren. Vanaf dat moment is de doelcongruentie tussen de maatstaf en het ondernemingsdoel grotendeels verdwenen en zal waarschijnlijk ook de directe correlatie tussen maatstaf en ondernemingsdoel sterk afnemen. Een aardig voorbeeld daarvan is dat de correlatie tussen de populair geworden financiële maatstaf 'Economic Value Added' (een maatstaf voor overwinst; bedrijfsresultaat minus de kosten van vermogensbeslag) en aandelenrendement in de VS sterk is gaan afnemen in het midden van de jaren negentig, precies in de periode dat de toepassing ervan een grote vlucht nam. Dat wijst erop dat vanaf het moment dat de maatstaf werd toegepast deze deels werd gemaximaliseerd door boekhoudkundige en andere trucs en niet door acties die echte economische waarde creëren.

Gangbare keuzedimensies

Een belangrijke afweging bij de selectie van prestatie maatstaven in de praktijk is die tussen objectieve en subjectieve maatstaven. Ook die keuze heeft betrekking op de afruil tussen ruis en doelcongruentie. Een belangrijk voordeel van objectieve maatstaven is dat de (meet)ruis minimaal is; allerlei productieaantallen, het aantal aanslagen per minuut, de gewerkte tijd of het aantal storingen is vaak eenvoudig vast te stellen. Maar meestal is dat maar één aspect van iemands taak. Met name kwaliteitsaspecten, klantvriendelijkheid en coöperatief of collegiaal gedrag zijn vaak belangrijk maar veel lastiger objectief vast te stellen. Dus uit een oogpunt van congruentie zijn juist subjectieve maatstaven vaak te prefereren, daarvan is de ruis echter doorgaans groter. Een andere afweging in de praktijk, ook te vertalen naar de afruil tussen ruis en congruentie, is individuele versus groepsmaatstaven. Uit een oogpunt van minimale ruis zijn individuele maatstaven verkieslijk omdat die door individuele werknemers relatief makkelijk kunnen worden beïnvloed en geen liftersgedrag uitlokken. Maar in veel gevallen is het belangrijk om samen te werken om de productie te maximaliseren. En juist

groepsmaatstaven kunnen dicht tegen bedrijfsdoelstellingen aanliggen. Dus met het oog op maximale congruentie verdienen groepsmaatstaven de voorkeur. Het productieverlies door onvoldoende samenwerking moet dan worden afgewogen tegen de mindere prikkel (inclusief liftersgedrag) omdat groepsmaatstaven vanuit het perspectief van de individuele werknemer slechter te beïnvloeden zijn en dus meer ruis bevatten.

Empirie

Omdat de primaire afruil bij de selectie van prestatie maatstaven die we in dit artikel bespreken het resultaat is van zeer recent theoretisch onderzoek, bestaan er nog geen ondersteunende empirische analyses. Daarom hebben we zelf geprobeerd te toetsen of er inderdaad een afruil bestaat tussen ruis en doelcongruentie. Onze dataset bestaat uit jaarrekening- en koersgegevens van 1650 Amerikaanse niet-financiële beursgenoteerde ondernemingen over de periode 1994-1998 met een beurskapitalisatie van minimaal \$ 120 miljoen in 1998. Voor deze steekproef hebben we indicatoren ontwikkeld voor de ruis en de doelcongruentie van zeven ondernemingsprestatie maatstaven. Tevens hebben we ondernemingen ingedeeld in vijftien verschillende bedrijfstakken. De resultaten zijn weergegeven in [tabel 1](#). We zien dat alle correlatiecoëfficiënten het verwachte, op afruil duidende positieve teken hebben. Bovendien is het merendeel (zeer) significant verschillend van nul. Dus naarmate de ruis van een maatstaf groter is (ongunstig) zal de doelcongruentie ook groter zijn (gunstig). Deze resultaten wijzen er sterk op dat er inderdaad sprake is van een afruil tussen ruis en doelcongruentie: hoe beter een maatstaf beïnvloedbaar is door het management (lage ruis), des te lager de correlatie tussen het effect van zijn werkzaamheden op de maatstaf en het effect op het ondernemingsdoel (lage doelcongruentie), lees aandelenrendement.

Tabel 1. Correlatie tussen ruis en doelcongruentie per bedrijfstak.

Een grote correlatie wil zeggen dat wanneer de bedrijfstak bijvoorbeeld hoog scoort op ruis (wat slecht is), dan scoort ze ook hoog op doelcongruentie (wat goed is). Een hoge correlatie geeft dus aan dat je niet op beide aspecten goed kunt scoren met een beloningsmaatstaf

Bedrijfstak	Pearson correlatie ^a	Spearman rangcorrelatie ^a
amusement	0,91c	0,85c
bouw	0,57b	0,55b
zakelijke dienstverlening	0,52	0,63b
chemie	0,81c	0,71c
communicatie	0,00	0,03
duurz. cons. Producten	0,77c	0,82c
commerciële gezondheidszorg	0,58b	0,41
high tech	0,70c	0,65c
informatietechnologie	0,64c	0,86c
machinebouw	0,88c	0,92c
papier	0,40	0,47
olieindustrie	0,85c	0,74c
uitgeverijen	0,63b	0,69c
transport	0,87b	0,72b
nutsbedrijven	0,40	0,56b

a. Het aantal maatstaven waarmee is gerekend bedroeg 14.

b. significantieniveau 90%. c. significantieniveau 95%.

Conclusie

Een van de belangrijkste instrumenten voor het aansturen van elke soort organisatie en het motiveren van medewerkers is het selecteren van de juiste prestatie maatstaven. De twee belangrijkste selectiecriteria voor prestatie maatstaven zijn de ruis en de doelcongruentie van de maatstaf. Bovendien is er, zeker voor de minder eenvoudige meerdimensionale taken en functies, sprake van een afruil tussen ruis en congruentie. Naarmate iemand een maatstaf zelf minder goed kan beïnvloeden door de aanwezigheid van externe factoren (economisch klimaat, inspanningen van collega's, weersomstandigheden enzovoort) of meetproblemen (bijvoorbeeld kwaliteit, service en klantvriendelijkheid) is de ruis groter en is de maatstaf minder geschikt voor evaluatie, beloning en promotie. Met doelcongruentie duiden we op de (cor)relatie tussen het effect van iemands inspanning op zijn prestatie maatstaf en het effect van die inspanning op het ondernemingsdoel. De perfecte maatstaf heeft minimale ruis en maximale doelcongruentie. Maar helaas, niets is perfect, ook prestatie maatstaven niet. Met ons empirisch onderzoek hebben we laten zien dat er inderdaad sprake is van een afruil. Als de ruis (tijdreeksvariantie) van maatstaf a kleiner is dan die van maatstaf b dan zal doorgaans de doelcongruentie van a juist hoger zijn dan die van b. Er is dus sprake van een afruil tussen ruis (beïnvloedbaarheid van de maatstaf) en doelcongruentie van prestatie maatstaven.

Implicatie

Dit betekent bijvoorbeeld dat het geven van aandelenopties aan caissières als optimale prestatie prikkel niet effectief is. De waarde van die opties is weliswaar zeer sterk gecorreleerd met het aandelenrendement (hoge doelcongruentie) maar de ruis is enorm, omdat de optiewaarde van bijvoorbeeld Ahold-aandelen wordt beïnvloed door nog 260.000 andere medewerkers en een groot aantal externe factoren. Opties voor de voorzitter van de raad van bestuur zijn daarentegen veel zinvoller omdat hij wel sterke invloed kan uitoefenen op het aandelenrendement, dus op de optiewaarde. Of een ander voorbeeld, de heer Van Wijk zou als directievoorzitter van klm prima kunnen worden afgerekend op onder meer de punctualiteit van alle vertrekkende blauwe vogels. Die maatstaf is voor hem aan de ene kant beter beïnvloedbaar (lagere ruis) dan de aandelenkoers maar aan de andere kant zijn er natuurlijk nog meer factoren (bijvoorbeeld kosten, beladingsgraad, gemiste aansluitingen, klantvriendelijkheid, enzovoort) die de marktwaarde van klm bepalen. Voor een medewerker in de bagagekelder daarentegen bevat die gemiddelde klm-punctualiteit juist te veel ruis en is voor hem of haar het aantal

koffers dat hij of zij per uur op de bagagekarretjes zet een betere maatstaf omdat die maatstaf minder ruis heeft, dus beter beïnvloedbaar is. Ook die maatstaf is echter niet perfect voor de bagagemedewerker, want bijvoorbeeld ook de tijdigheid van de het plaatsen van die koffers en de wijze waarop ze op de karren worden geplaatst (met of zonder beschadigingen) is van belang voor klm. Deze tijdigheid en zeker de wijze waarop koffers worden geplaatst zijn echter slechter te meten, dus ontstaat meer ruis. Er is dus altijd sprake van een afruil: als de ene maatstaf beter beïnvloedbaar is (minder ruis) dan zal hij doorgaans slechter gerelateerd zijn aan het ondernemingsdoel. De kunst is het vinden van die maatstaven per functie of taak waarvoor de optelsom van beïnvloedbaarheid en doelcongruentie maximaal is. Deze fundamentele afruil ligt aan de basis van elk selectieproces van prestatie maatstaven. Hiermee kan onder meer het fenomeen van gaming en window dressing beter worden begrepen en kunnen ook keuzeprocessen in de praktijk tussen bijvoorbeeld individuele en groepsmaatstaven, of tussen subjectieve en objectieve maatstaven beter worden begrepen en begeleid. Tot welke negatieve gevolgen en schandalen het hanteren van prestatie maatstaven met onvoldoende doelcongruentie kan leiden, is de laatste tijd helaas maar al te vaak gebleken

Methode Pearson- en Spearman-rangcorrelaties

In navolging van anderen nemen we als indicator voor de ruis van elke maatstaf, de mediane waarde binnen een bedrijfstak van de varianties van elke maatstaf in individuele ondernemingen over vijf jaar. Naarmate deze 'tijdreeksvariantie' groter is, zal immers ook de invloed van externe factoren op schommelingen van het bedrijfstakresultaat groter zijn en dergelijke bedrijfstakfluctuaties - ten gevolge van valutakoersen, olieprijsen, bedrijfstakcycli, enzovoort - zijn doorgaans door individuele managers en bedrijven moeilijk te beïnvloeden. Hoe groter de variantie, des te groter de ruis van die betreffende maatstaf, en des te lager de beïnvloedbaarheid door een individuele manager. De correlatie tussen het effect van inspanning op de maatstaf en het effect op het ondernemingsdoel meten we door de mediane correlatie te bepalen tussen de prestatie maatstaf en het aandelenrendement van de onderneming in de loop van de vijf bestudeerde jaren. Naarmate die correlatie groter is, zal ook de doelcongruentie groter zijn. Om vast te stellen of er inderdaad sprake is van een afruil tussen ruis en doelcongruentie hebben we voor elk van de vijftien bedrijfstakken de Pearson- en Spearman-rangcorrelaties berekend tussen de tijdreeksvarianties en de tijdreekscorrelaties. Pearson-correlatie is de standaardmanier om correlaties te berekenen, maar in dit geval zijn de Spearman-rangcorrelatiecoëfficiënten het meest interessant. Bij Spearman-rangcorrelaties wordt alleen gekeken naar de volgorde van de waarnemingen en wordt vervolgens bepaald in welke mate de volgorde van de twee variabelen gecorreleerd is. Met andere woorden, of de afstand tussen twee opeenvolgende waarnemingen groot of klein is, is niet van invloed op de berekende Spearman-rang correlatie. Zodoende corrigeren Spearman-rangcorrelaties voor het feit dat de tijdreeksvarianties, per definitie, sterker variëren dan de tijdreekscorrelaties en geven daarom een beter beeld van de 'werkelijke' mate van correlatie tussen ruis (tijdreeksvarianties) en doelcongruentie (tijdreekscorrelaties).

¹ Zie R.S Kok, C. M. van Praag en K. Cools, Motiverend belonen loont, ESB, 29 november 2002, blz. 867-870.