



LITERATUUR

Ioannidis, J. (2005)

Contradicted and initially stronger effects in highly cited clinical research.

Journal of the American Medical Association, 294, 218-228.

Laband, D. & R. Tollison

(2003) Dry holes in economic research, *Kyklos*, 56, 161-173.

Oswald, A. (2006)

An examination of the reliability of prestigious scholarly journals: evidence and implications for decision makers.

Coventry: Warwick University, mimeo.

Tegenspraak zonder vooruitgang

Volgens de klassieke wetenschapsfilosofie van Karl Popper is tegenspraak de motor van de vooruitgang in de wetenschap. Onder wetenschapsfilosofen is dit inzicht omstreden, maar algemeen worden anomalieën toch gezien als het begin van nieuwe inzichten. Dit rooskleurige beeld van de wetenschap is in strijd met de werkelijkheid. Binnen de sociale wetenschappen – en dan in het bijzonder binnen de economische wetenschap – is tegenspraak meestal de uitkomst van gebrekkig onderzoek.

Een paar jaar geleden was ik tijdens een wetenschappelijk congres voorzitter van een sessie over het rendement van bedrijfsopleidingen. Het rendement van bedrijfsgerelateerde scholing is een belangrijke maatschappelijke vraag. Een hoog rendement kan betekenen dat er sprake is van onderinvestering en marktfalen. Een laag rendement geeft aan dat misschien alle aandacht voor scholing in bedrijven overdreven is en het geld beter anders besteed kan worden.

Op het congres werden vier papers gepresenteerd. Het eerste paper concludeerde dat het rendement van bedrijfsopleidingen meer dan honderd procent was, het tweede paper kwam tot de conclusie dat het rendement nihil was. De conclusies van het derde en vierde paper lagen tussen die van het eerste en het tweede in. Vier papers met volledig tegengestelde conclusies. Als het rendement nul procent is, is investeren in bedrijfsopleiding een slechte beslissing. Als het honderd procent is, is het een uitstekend idee om geld aan bedrijfsopleidingen te besteden. Neem je alle vier bij elkaar, dan zou je moeten concluderen dat het rendement van bedrijfsopleidingen ergens tussen nul en honderd procent ligt. Van een dergelijke conclusie worden we niet veel wijzer.

Tegenspraak komt vaker voor dan we denken. In een recent artikel onderzoekt John Ioannidis hoe vaak het voorkomt dat zeer veel geciteerde artikelen later worden weerlegd of effecten vinden die sterker of groter zijn dan in vergelijkbare latere studies (Ioannidis, 2005). Hij neemt daarvoor artikelen, gepubliceerd in medische toptijdschriften tussen 1990 en 2003, die meer dan duizend keer in de literatuur geciteerd zijn. Een van de meest opzienbarende uitkomsten van dit onderzoek is dat 83 procent van de studies op basis van *survey data* in de loop van de tijd door ander onderzoek weerlegd werden door andere studies. Van de studies op basis van een gecontroleerd experiment – een *randomized controlled trial* – werd daarentegen 'slechts' 32 procent weerlegd. Wat het onderzoek van Ioannidis duidelijk maakt is dat survey onderzoek vaker

weerlegd wordt dan de kwalitatief betere gecontroleerde experimenten. Survey onderzoek is kwalitatief minder omdat het vrijwel altijd onmogelijk is om alle mogelijke versturende factoren constant te houden. In een gecontroleerd experiment kan dat wel. Verder bleken vooral uitkomsten van studies met een geringe steekproefomvang vaak niet gerepliceerd te kunnen worden.

Ik vermoed dat het aantal contradicties binnen de economische wetenschap eerder hoger dan lager ligt dan de cijfers die Ioannidis noemt. Om te beginnen kijkt Ioannidis alleen naar studies die heel vaak geciteerd worden. Geen enkel economisch artikel haalt zo veel citaties als de medische studies waar Ioannidis naar kijkt. Zo vindt Andrew Oswald dat het meest geciteerde economische artikel gepubliceerd in 1981 in de jaren daarna slechts 401 keer geciteerd is (Oswald, 2006). Veel geciteerde studies zijn studies, die om de een of andere reden veel aandacht hebben getrokken. Vanwege deze aandacht worden deze studies vaker gerepliceerd. Replicatie is echter eerder uitzondering dan regel. Negeren komt vaker voor dan repliceren. Slecht onderzoek en onderzoek met onwaarschijnlijke uitkomsten wordt meestal genegeerd. Zo vinden David Laband en Robert Tollison dat 26 procent van de artikelen in economische tijdschriften in de vijf jaar na publicatie niet eenmaal wordt geciteerd (Laband & Tollison, 2003). Voor ons economen zijn er denk ik twee lessen te trekken uit het onderzoek van Ioannidis. De eerste is dat we meer tijd en aandacht moeten schenken aan replicatie van onderzoeksresultaten. Binnen de economie heerst het paradigma van de originaliteit: iedereen wil het net iets anders – en beter doen – dan de voorgangers. Replicatie van goed onderzoek is echter waardevol omdat het bijdraagt aan de zekerheid van onderzoeksuitkomsten.

De tweede les is dat we meer aandacht moeten schenken aan meta-analyses, dat wil zeggen onderzoek waarin de beschikbare onderzoeksresultaten systematisch geanalyseerd worden. Nu baseren we ons vaak op de uitkomsten van de studie die onder vakgenoten als 'kwalitatief de beste' wordt beschouwd. Het onderzoek van Ioannidis leert echter dat kwalitatief goed onderzoek ook vaak weerlegd wordt, zeker als het om survey onderzoek gaat. Als we niet weten waar de waarheid ligt – in de uitkomsten van de originele studie of in de tegengestelde conclusies uit de replicatie – kan een meta-analyse van al het beschikbare onderzoek uitkomst bieden. Wellicht dat we zo ooit nog eens te weten komen wat het rendement op bedrijfsgerelateerde scholing nu echt is.