

Grote kwesties...

Wijlen André Gorz...

Eerder toevallig stootte ik begin vorig jaar in een Brusselse boekhandel op een pas uitgebracht boekje van André Gorz, *Lettre à D.* (Editions Galilée, Paris, 2006). De ondertitel van het boekje – de auteur zelf noemt het “un récit” – is *Histoire d'un amour*. Het bevat een erg persoonlijke getuigenis over de cruciale rol die de vrouw waarmee hij meer dan een halve eeuw samenleefde in zijn leven speelde. Verhaal van een liefde, maar ook een aangrijpende schuldbekentenis. “Pourquoi es-tu si peu présente dans ce que j'ai écrit alors que notre union a été ce qu'il y a de plus important dans ma vie? Pourquoi ai-je donné de toi dans ‘Le Traître’ une image fautive et qui te défigure? Ce livre devait montrer que mon engagement envers toi a été le tournant décisif qui m'a permis de vouloir vivre?” Op weg naar Oostende heb ik de pakweg zeventig bladzijden in een ruk uitgelezen, beseffend dat ik naar alle waarschijnlijkheid net het allerlaatste boek van André Gorz had gekocht. *Lettre à D.* brengt niet alleen het relaas van een gezamenlijk levensavontuur, maar kondigt tegelijk ook het einde ervan aan. De eerste regels geven kort en sober weer aan welk fysiek aftakelingsproces D. de voorbije twintig jaar ten prooi was en de laatste regels beschrijven hoe de auteur 's nachts dikwijls het silhouet voor zich ziet van een man die op een eenzame weg in een verlaten landschap achter een lijkwagen stapt. “Je suis cet homme. C'est toi que le corbillard emporte. Je ne veux pas assister à ta crémation; je ne veux pas recevoir un bocal avec tes cendres. J'entends la voix de Kathleen Ferrier qui chante 'Die Welt ist leer, Ich will nicht leben mehr' et je me réveille. Je guette ton souffle, ma main t'effleure. Nous aimerions chacun ne pas avoir à survivre à la mort de l'autre. Nous nous sommes souvent dit que si, par impossible, nous avions une seconde vie, nous voudrions la passer ensemble.” Toen ik een achttal maanden later vernam dat André Gorz – samen met Dorine –

op 24 september 2007 uit het leven was gestapt, verraste mij dat dus niet echt. *Lettre à D.* was de erg uitgebreide, maar finale versie van de opdracht die in meerdere van zijn recente boeken op de eerste bladzijde stond: “Dorine sans qui rien ne serait”.

Wie is in godsnaam André Gorz, hoor ik u al vragen? En waarom ons over zijn werk onderhouden? Ik begrijp deze reactie wel. Rekening houdend met de beperkte aandacht voor zijn overlijden in de Vlaamse pers (en de vrij onnauwkeurige situering van zijn werk als men er toch over rapporteerde) valt te vrezen dat hij nu niet enkel dood is, maar al veel langere tijd ook vergeten. Een paar jaar geleden had ik de kans om enkele keren een lezing te geven over zijn werk en telkens weer bleek het noodzakelijk om omstandig uit te leggen wie André Gorz eigenlijk was en waarom zijn werk nog altijd verdient om gelezen te worden. Maar als één publicist arbeidsmarktwaarnemers van mijn generatie gedurende de tweede helft van de vorige eeuw steeds weer met grote kwesties confronteerde dan was het wel hij. Vandaar.

Geboren in 1924 als Gerard Horst – al een verbastering van de oorspronkelijke familienaam door zijn joodse vader met het oog op een gemakkelijker integratie na zijn huwelijk met een katholieke vrouw – verliet hij op jonge leeftijd zijn geboortestad Wenen, kort na de aanhechting van Oostenrijk onder Hitler. In Zurich maakte hij kennis met het werk van Sartre en later met Sartre zelf. Na de beëindiging van de Tweede Wereldoorlog vestigde hij zich in Parijs. Hij werkte er als journalist voor *L'Express* en stichtte mee *Le Nouvel Observateur*, een blad waarvoor hij blijft werken tot hij in 1983 op vervroegd pensioen ging om beter te kunnen zorgen voor zijn vrouw bij wie men kanker had vastgesteld. In Parijs bewoog hij zich in de kring rond Sartre en De Beauvoir en fungeerde vanaf 1961 in opvolging van Maurice Merleau-Ponty als politiek

directeur van het tijdschrift *'Les Temps Modernes'*. Zijn belangrijkste boek uit die periode (en misschien wel van zijn hele oeuvre) is zonder twijfel *'Le Traître'*, een biografische roman die in 1958 verschijnt met een voorwoord van Sartre. Uit dezelfde periode stamt ook een lijvig manuscript over moraalfilosofie dat pas in 1977 voor het eerst werd uitgegeven. Sartre beschouwde *'Fondements pour une morale'* (Editions Galilée, Paris) als het boek over moraal dat hij zelf nooit schreef en als een volwaardig complement van zijn *Critique de la Raison Dialectique*.

Vanaf het begin van zijn carrière publiceerde Gerard Horst zijn filosofische en sociaalkritische werken onder de naam André Gorz. Zijn journalistiek werk ondertekende hij met Michel Bosquet, het pseudoniem dat hem door de hoofdredacteur van L'Express werd voorgesteld. Pas in 1978 bij het verschijnen van *'Ecologie et Politique'* (Editions du Seuil, Paris) – een bundeling van recent journalistiek werk en het eerder verschenen essay *'Ecologie et Liberté'* (1975) – werd voor het grote publiek duidelijk dat achter de bekende journalist en de sociaalkritische filosoof een en dezelfde persoon schuil ging. Op de titelpagina stond als auteur André Gorz/Michel Bosquet. De korte inleiding vermeldde dat het journalistieke werk en de dieper gravende essays qua stijl tot hiertoe moeilijk te verzoenen waren maar dat de ecologische problematiek uiteindelijk leidde tot een verstrengeling van beiden. Vandaar het dubbele auteurschap. Vanaf dan schreef hij enkel onder de naam André Gorz.

Gorz bekleedde nooit een academische positie. Niettemin beïnvloedden zijn denkbeelden en geschriften ook de wetenschappelijke wereld. Wie de moeite neemt om oude jaargangen van vaktijdschriften te doorbladeren – bijvoorbeeld jaargangen van eind de jaren zestig en begin de jaren zeventig van de prestigieuze *Journal of Human Resources* – zal zelfs daarin meermaals verwijzingen vinden naar zijn werk. Ik kan me geen andere auteur zonder academische geloofsbriefvinden indenkende die toch zo sterk woog op het wetenschappelijke discours over de arbeidsmarkt. Eerst over de problematiek van de arbeidsdeling, daarna over de rol van de middenklassen en de intellectuelen, later over het werkgelegenheidsbeleid. Elk nieuw boek van Gorz bevatte onvermijdelijk iets controversieels. Elk nieuw boek vormde onvermijdelijk een bron van controversie en een aanleiding tot discussie. Al

was het maar over de empirische basis van zijn stellingen – empirische basis die door velen als weinig betrouwbaar werd beschouwd. Ten tijde van Ignace Lindemans werd Gorz ooit door de christelijke arbeidersbeweging uitgenodigd – ik vermoed als spreker op een Sociale Week – maar dit belette niet dat GiDS in de Gids op Maatschappelijk Gebied tegelijk op zijn gekende manier brandhout maakte van de cijfertjes waarop Gorz zich beriep om zijn stellingen over de arbeidsduurverkortings te onderbouwen.

Ter illustratie van de verrassende wendingen, die het werk van Gorz kon aannemen, het volgende voorbeeld. In 1997 verschijnt *'Misères du Présent, Richesses du Possible'* (Editions Galilée, Paris). Tot verbazing van vele lezers neemt Gorz in dit boek erg expliciet afstand van de stelling die zijn werk het vorige decennium kenmerkte. "J'ai longtemps refusé l'idée d'un revenu social qui permette de 'vivre sans travailler'. [...] Dans de précédents ouvrages, je voulais que la garantie à chacun d'un plein revenu soit liée à l'accomplissement par chacun de la quantité de travail nécessaire à la production des richesses. Cette formule, que je préconisais à partir de 1983. [...] Je l'abandonne donc"

Sinds 1983 argumenteerde Gorz inderdaad dat enkel een combinatie van (drastische) arbeidsduurverkortings en (drastische) arbeidsherverdelingsdoelstellingen bood indien men een beleid wenste te voeren dat aan iedereen een baan in het formele circuit garandeerde. Dit argument steunde op twee uitgangspunten: a) de vaststelling dat aanhoudende productiviteitsstijgingen onvermijdelijk een vermindering van het aantal noodzakelijke arbeidsuren voortbrachten met als gevolg dat steeds meer mensen dreigden in de werkloosheid of in de inactiviteit te belanden en b) de stelling dat een baan in het formele circuit tegelijk ook fungeert als een 'ticket' dat toegang verleent tot '(mede)burgerschap' en een lidmaatschap van de maatschappij fundeert. In zekere zin kan men deze stelling van Gorz beschouwen als een emancipatorische versie van pleidooien voor volledige werkgelegenheid. Door een aantal progressieve lieden is zij in de debatten over een activerend arbeidsmarktbeleid ook als dusdanig gebruikt. In Vlaanderen heeft Gorz met zijn argumenten voor een (gedeeltelijke) ont koppeling van inkomen en het aantal gepresteerde arbeidsuren via een vorm van 'tweede cheque' zon-

der twijfel de nodige invloed gehad. Men kan hem zelfs beschouwen als de verre inspirator van zaken als het Derde Arbeidscircuit, het alom gekende 'rugzakje' dat je aan uitkeringsgerechtigden wil meegeven omwille van meer kansen op sociale integratie via 'employability' of van het 'tijdskrediet'. In de discussie over een basisinkomen gaven deze argumenten voedsel aan tegenstanders om een volledige ontkoppeling van jobs en inkomen af te wijzen. Maar, zegt Gorz in *'Misères du présent, Richesses du possible'*, deze analyse en de daaruit volgende stelling deel ik zelf niet langer. Niet omdat de vaststelling over de effecten van de productiviteitsstijgingen fout waren, maar omdat jobs niet langer de rol van 'toegangsticket tot burgerschap' (kunnen) vervullen en een bepaalde vorm van basisinkomen het enige redmiddel is uit de huidige impasse.

Welke evoluties Gorz ertoe brachten om van mening te veranderen is niet echt duidelijk en hier is niet de plaats om dit in detail te onderzoeken. Sinds het verschijnen van het boek heb ik trouwens weinig commentaren op deze omslag gelezen. Het lijkt wel alsof de vroegere volgelingen van Gorz dit boek niet hebben gelezen of zijn verandering van positie simpelweg negeren. Wel is duidelijk dat deze verrassende wending ook deze keer, net als in alle andere werken, het gevolg is van uitgebreide discussies met woordvoerders van maatschappelijke bewegingen. Deze werkwijze vergemakkelijkt de lectuur van Gorz' werk niet bepaald. Enig zicht op met wie hij in een bepaald werk de dialoog aangaat, is meestal nodig om de essentie van zijn argumenten te begrijpen.

De analyses die hij in de jaren zeventig maakte van de negatieve effecten van de sociale en technische arbeidsverdeling vloeiden voort uit uitgebreide contacten met de Duitse, Franse, Engelse en Italiaanse vakbeweging. De theoretische concepten waarvan hij gebruik maakte, dragen duidelijk de sporen van discussies met belangrijke denkers zoals Herbert Marcuse, Ivan Illich of Jürgen Habermas. Voor wat betreft de recente wijzigingen in zijn visie op de maatschappelijke rol van 'jobs' is de invloed merkbaar van een groep Franse en Italiaanse intellectuelen (Yann Moulier-Boutang, Antonella Corsani, Patrick Dieuaide, Carlo Vercellone, Maurizio Lazzarato en anderen). In de slipstream van de Italiaanse filosoof Toni Negri en zijn erg specifieke interpretatie van het marxistisch gedachtegoed

brenge deze auteurs een erg aparte analyse van de zogeheten immateriële productie, die zij beschouwen als een onderdeel van een nieuw – 'cognitief' – kapitalisme. De neerslag van de discussie met deze stroming schrijft Gorz neer in *'L'immateriel'* (Editions Galilée, Paris, 2003). Maar het boek vertoont evenzeer invloeden van het werk van Hans Jonas of van Peter Sloterdijk... en van theoretische reflecties op de praktijk van hackers en de ontwikkeling van 'open-source software'.

Vandaag ligt er in dezelfde (en andere zogeheten 'betere') boekhandel(s) opnieuw een boek van André Gorz. Postuum uitgegeven, maar nog wel door Gorz zelf samengesteld. Het bevat zeven eerder verschenen teksten en interviews, waaronder één die hij pas twee uur voor zijn overlijden aan de redactie van *Ecorev* doorstuurde.

Ecologica (Editions Galilée, Paris, 2008) is het politieke testament van Gorz. Het brengt een overzicht van de argumenten die hij zelf als de kern van zijn werk beschouwde: de ontwikkeling van een kritische reflectie op de consumptiemaatschappij tot een gedegen fundament voor een politieke ecologie. Politieke ecologie betekent voor Gorz in eerste instantie een 'ethiek van bevrijding'. Als een rode draad loopt door zijn werk een dubbele waarschuwing. Waarschuwing tegen de expertocratie. Waarschuwing tegen de zelfbeperking. Waarschuwing tegen twee soorten antwoorden op de ecologische problematiek, die ertoe leiden om de vrijheid van de individuen zeer sterk in te perken en die men daarom nauwelijks als emancipatorische ontwikkelingen kan beschouwen. "Le sens fondamental d'une politique écosociale [...] est de rétablir politiquement la corrélation entre moins de travail et moins de consommation d'une part, plus d'autonomie et plus de sécurité existentielles, d'autre part, pour chacun et chacune." (pp. 66-67) De cruciale zin is hier 'rétablir politiquement' – niet door een ingreep van experts en technocraten, ook niet door een beweging van zichzelf beperkende opleggende individuen.

Voor wie nog nooit iets van Gorz heeft gelezen – of erger nog: wie nog nooit van Gorz hoorde – biedt *Ecologica* de uitgelezen kans om kennis te maken met een controversieel, maar verrassende inzicht biedende denkwereld. En voor ons allemaal een goede reden om via het herlezen en blijven herle-

zen van zijn werk de confrontatie aan te gaan met vragen over 'grote kwesties'.

Voer voor liefhebbers van statistiek... en 'evidence-based policy-makers'

Statistische significantietoetsen vormen een onvermijdelijk en daarom een omzeggens als ritueel te beschouwen onderdeel van geschriften die empirische onderzoeksresultaten presenteren. Statistische significantietoetsen. U weet wel, die voor de meeste lezers moeilijk te begrijpen – en zelfs voor specialisten niet echt gemakkelijk uit te leggen – verwijzingen naar cijfers die het resultaat vormen van een procedure, die de lezer iets moet zeggen over het belang dat hij aan de gerapporteerde gegevens moet hechten. U kent ze wel, de toetsen: chi-kwadraat, gamma, t-toets, F-toets, Kolmogorov-Smirnoff of iets dergelijks. Gewoonlijk vindt men deze onderdelen van een onderzoeksrapport onderaan kruistabellen of in aparte kolommen van een tabel als het gaat om resultaten van meer ingewikkelde analyses.

Meestal maakt de auteur gebruik van twee soorten ingrediënten om aan te geven of iets al dan niet en meer of minder significant is: cijfertjes en sterretjes. Eén sterretje hoort dan bij iets dat conventioneel wordt aangegeven als < 0.05 , twee sterretjes bij < 0.01 en drie sterretjes bij < 0.001 . In mensentaal geven deze symbolen aan dat er minder dan één kans op twintig (respectievelijk honderd en duizend) is dat het geconstateerde resultaat toe te schrijven is aan toeval. Recent, zo hebt u wellicht ook al gemerkt, bestaat er een tendens om bij presentaties op seminaries of workshops gemakshalve en, naar ik vermoed, omwille van de leesbaarheid minimalistisch tewerk te gaan. De echte resultaten van de test – de cijfers zelf – laat men dan achterwege en op de vertoonde plaatjes vermeldt men nog enkel de sterretjes.

Gezien de opgang van wat men zo mooi 'evidence-based policy' noemt, verrast het niet echt dat men ook bij beleidsonderzoek en vanwege beleids mensen meer en meer kritische opmerkingen krijgt als bij tabellen de verwachte significantietoetsen ontbreken. Op zich een te waarden praktijk want (zeker) van een beleidsman mag men verwachten dat zij of hij wil weten of men wel degelijk staat kan

maken op de resultaten van empirisch wetenschappelijk onderzoek alvorens iets te ondernemen. Stel je voor dat men hervormingen zou doorvoeren omwille van resultaten die achteraf blijken te berusten op toeval. Daarom: lang leve de significantietoetsen!!!

En toch. Wie ooit een min of meer ernstige cursus statistiek onderging, hoort te weten dat het toetsen van significantie enkel betekenisvolle resultaten oplevert als de gegevens voldoen aan welbepaalde voorwaarden. De meest belangrijke is dat de gegevens die men analyseert een getrouwe weergave zijn van de populatie waarop de conclusies geacht worden van toepassing te zijn. Significantietoetsen zijn nuttig om te bepalen of een uitspraak op basis van een gedeelte van de populatie met recht mag veralgemeend worden naar het geheel van de populatie. (Of significantietoetsen zinvol zijn als men beschikt over populatiegegevens is trouwens onder statistici een topic waarover men het niet eens is.) Daarom is het ook zo belangrijk om te benadrukken: een belangrijke vereiste voor het zinvolle gebruik van significantietoetsen is dat de gegevens zijn verkregen op basis van een toevalssteekproef. Een simpele of enkelvoudige toevalssteekproef om precies te zijn. Want bij gebruik van een meer efficiënte manier van steekproeftrekking (zoals bij stratificatie waarbij men bij voorbaat de nodige voorzorg neemt opdat de steekproef een aantal karakteristieken van de populatie, bijvoorbeeld de verhouding mannen versus vrouwen, exact weergeeft) krijgt men sowieso 'betere' resultaten en zou men dus de resultaten van de toets strenger moeten beoordelen. Even belangrijk is dat significantie van een resultaat ook afhangt van de grootte van de steekproef. Enigszins overdrijvend – maar slechts enigszins – kan men stellen dat elk resultaat significant wordt als de steekproef maar groot genoeg is. De belangrijkste les van dit alles is dat gebruik maken van significantietoetsen liefst niet mag gezien worden als een procedure die standaard en zonder verder nadenken op alle onderzoeksresultaten mag en kan worden toegepast. En ernstige statistici waarschuwen daar dan ook steeds heel nadrukkelijk voor.

En toch. Een groot probleem is dat men significantietoetsen ondanks deze waarschuwingen niet altijd volgens de regels van de kunst toepast. 'Niet altijd' is zelfs een understatement. Jaren geleden

heeft D. McCloskey na onderzoek van een groot aantal publicaties in economische tijdschriften geconcludeerd dat de verleiding om significantietoetsen te laten verworden tot een ritueel in de slechte betekenis van dit woord echt wel groot is. Het gros van de artikels in economische tijdschriften gebruikt significantietoetsen op een oneigenlijke en nietszeggende manier. Nog erger, voegde McCloskey daar later aan toe, is dat economen de overtuiging zijn toegedaan dat bij de interpretatie en de voorstelling van hun resultaten enkel de statistische significantie van belang is. Ze vragen zich nog nauwelijks af of de verklarende variabelen ook wel degelijk maatschappelijk belang en vooral 'impact' hebben. Niet statistisch significante, maar variabelen met 'oomph' – da's wat je hoopt met je analyses te kunnen aanduiden, zegt McCloskey. (Wie zich in deze materie iets meer wil verdiepen, kan ik als aanzet verwijzen naar: McCloskey, D. 1992. The Bankruptcy of Statistical Significance. *Eastern Economic Journal*, 18: 395-361; Ziliak, S., McCloskey, D. 1996. The Standard Error of Regressions. *Journal of Economic Literature*, 34: 97-114; Ziliak, S., McCloskey, D. 2004. Size Matters: the standard error of regressions in the American Economic Review. *Journal of Socio-Economics*, 33: 527-556.

Hoe is het zover kunnen komen? Wat de economen betreft, wijst McCloskey (1996. *The Vices of Economists. The Virtues of the Bourgeoisie*. Amsterdam University Press) als de grote schuldige naar Rudolf Klein – Nobelprijswinnaar, bij economen en beleidsmakers vooral bekend omwille van zijn werk met erg grote econometrische modellen met veel vergelijkingen... en in onze contreien misschien ook nog wel omdat Jean-Pierre Van Rossem zich er indertijd op beriep om een tijdje bij deze man te hebben gestudeerd en diens methode te gebruiken bij zijn macro-economische voorspellingen. In de eerste gepubliceerde paper na zijn doctoraat gebruikte Klein in 1943, zegt McCloskey, de formule die sindsdien ritueel door economen is overgenomen: "The role of Y in the regression is not statistically significant. The ratio of the regression coefficient to its standard error is only 1.812. This low value of the ratio means that one cannot reject the hypothesis that the true value of the regression coefficient is zero." (R. Klein, geciteerd op p. 31)

Maar wat dan met de sociologen waar het rituele gebruik van de standaardprocedure – steeds de-

zelfde significantiedrempels zonder veel uitleg over het waarom, maar met de sterretjes – minstens zo wijd verbreid is als bij de economen? En ook in die discipline, zo wijst uitgebreid onderzoek uit, is het foutieve gebruik van significantietoetsen de norm.

Erin Leahey (2005. Alphas and Asterisks: The Development of Statistical Significance Testing Standards in Sociology. *Social Forces*, 48(1): 1-24) geeft ons enig inzicht in wanneer de huidige praktijk van het toetsen van significantie in zwang kwam en welke factoren hierbij eventueel een rol speelden. Volgt iedereen deze procedure uit rationele overwegingen omdat ze het meest of het best is aangepast aan de aard en de logica van het onderzoeksverloop? Maar waarom maken andere disciplines dan geen gebruik van dezelfde procedures? Of is er sprake van een vorm van besmetting of nabootsing? Volgen onderzoekers het voorbeeld van grootheden uit hun discipline en fungeren ze zo zelf ook weer als voorbeeld voor anderen? Of gaat het om een proces van disciplineren? Leggen sommige instellingen of organisaties regels op waaraan onderzoekers moeilijk kunnen weerstaan omdat ze anders hun carrière in gevaar brengen? Om op dit soort vragen een antwoord te vinden gaat Leahey uitermate empirisch tewerk. Hij nam een steekproef van (1 215) artikels met cijfergegevens uit de jaargangen van 1935 tot 1995 van de twee belangrijkste sociologische tijdschriften: de *American Sociological Review* en de *American Journal of Sociology* en noteerde zorgvuldig de karakteristieken van de manier waarop men de significantietoetsen gebruikte.

Tot welke vaststellingen leidt dit alles? Ongeveer de helft van deze artikels (613) test hypothesen en ongeveer 80% daarvan (496) maakt daarbij gebruik van significantietoetsen. Maar de toenemende trend doorheen de tijd is duidelijk. Voor 1940 vindt men nauwelijks echte testen. In de periode nadien neemt de praktijk continu toe, hoewel een echte doorbraak (met aanhoudend meer dan 90% van de artikels die gebruik maken van significantietoetsen) pas vanaf 1975 te noteren valt. Niet toevallig, zo merkt Leahey op, het tijdstip waarop de gekende statistische pakketten SPSS (1975) en SAS (1976) op de markt kwamen. Het gebruik van de twee rituele elementen – significantiedrempel (of alphaniveau) > 0.05 en de sterretjes – kent een min of meer gelijkaardig verloop doorheen de tijd, maar

op een veel lager niveau. De 0.05-drempel werd voor het eerst gebruikt in 1935 door de bekende statisticus R.A. Fisher. Dit voorbeeld vindt duidelijk enige navolging in het daarop volgende decennium; maar de echte doorbraak naar een breder verspreid gebruik stelt men pas vast vanaf het einde van de jaren zeventig. Sinds 1980 conformeert ongeveer 80% van de artikels zich aan deze standaard. De sterretjes verschijnen al een eerste keer rond 1955, maar verdwijnen dan weer tot begin de jaren zeventig. De echte doorbraak van dit element valt te situeren in de periode 1985 tot 1995. Dan stijgt het percentage artikels dat er gebruik van maakt van 10% tot iets boven de 40%. Alles samen leidt dit tot een eerste belangrijke vaststelling: het volledige pakket – te weten: statistische significantietoets, inclusief drempelwaarde(n) van 0.05 (en 0.01 en 0.001) en het gebruik van sterretjes – is pas in de laatste twee à drie decennia echt een wijd verbreide standaard geworden. En men mag een verdergaande verspreiding verwachten nu een ander statistisch pakket, STATA, de sterretjes in de routines opneemt.

Het duidelijke verband tussen de beschikbaarheid van technologie die het mogelijk maakt om met één druk op de knop en zonder veel verder nadenken alle mogelijke statistische toetsen uit de machine te laten rollen, is in zekere mate verontrustend. Op een jobinterview heb ik een pas afgestudeerd iemand ooit in alle ernst een standaardafwijking horen omschrijven als iets dat te berekenen valt door de juiste toets op een rekenmachine in te drukken. Wat op zich uiteraard correct is, maar geen inzicht in de betekenis ervan garandeert.

Een geruststellende vaststelling in het werk van Leahey is dan ook dat de aard van de gebruikte gegevens kennelijk toch van invloed is op hoe men de significantietoets hanteert. “[...] Authors using probability samples of data were more likely to use statistical significance tests than authors with non-probability samples” (p. 18) Maar daartegenover staat dat er geen verband kon worden vastgesteld tussen de grootte van de steekproef en de gebruikte drempelwaarden. Een andere ietwat verrassende vaststelling is dat bepaalde editors van tijdschriften ondanks het gebruik van een ‘blind’ refereeproces toch wel een invloed blijken te hebben op de aard van de artikels die onder hun hoede verschijnen. Bijvoorbeeld: in de periode dat een belangrijk tijd-

schrift geredigeerd wordt door een academicus, waarvan men weet dat hij als onderzoeker zelf meer nadruk legt op inhoudelijke dan statistische significantie blijkt de populariteit van de 0.05-drempel te dalen.

Leahey eindigt zijn analyse met de bedenking dat standaardisering van rapportering van onderzoeksresultaten uiteraard zijn voordelen heeft. Het maakt de lectuur van rapporten en de vergelijking van resultaten veel gemakkelijker. Maar standaardisering heeft ook zijn nadelen. “[...] If statistical significance testing has become habitual not because samples are increasingly collected through simple random sampling, but because researchers think that their research will be deemed unworthy if they do not use such tests, then perhaps the purpose of statistical significance testing should be critically evaluated.” (p. 19) Zo'n kritische evaluatie kan er eventueel toe leiden om met meer aandacht te onderzoeken welke alternatieve manieren er zijn om wetenschappelijke hypothesen te onderzoeken en besluiten te trekken uit de resultaten van empirisch onderzoek.

Deze conclusie klinkt als een echo van een boodschap, die McCloskey jaren geleden met de nodige nadruk meegaf aan haar vakgenoten: “Focusing on the calculation, if it causes us to forget this obvious job of making human judgments, is going to make us forget what we are doing. [...] What Klein and everyone in modern science are looking for is a mechanical, uncontroversial way of deciding whether some effect is large or small. No human judgment, please: we're scientists. [...] The practice grew, and especially in the 1970s when the computer chip came to maturity. Pretty soon everyone in economics thought that statistical significance was the same thing as scientific significance, that you could skip that last step of scientific work, the human assessment of largeness or smallness.” (*The Vices of Economists. The Virtues of the Bourgeoisie*, pp. 30-31).

Zeker voor beleidsonderzoek is het van uitermate groot belang deze laatste stap – “human judgment of largeness or smallness” – niet te verwaarlozen. En te beseffen dat de ‘menselijke’ beoordeling van de wetenschappelijke betekenis van onderzoeksresultaten niet samenvalt met de ‘politieke’ beoordeling van de maatschappelijke betekenis. ‘Oomph’ dus!

Retoriek en de 'gender wage gap'

Vrouwen verdienen gemiddeld nog altijd een stuk minder dan mannen. De vaststelling is al talloze keren gemaakt en we worden er keer op keer weer aan herinnerd. Recent nog ter gelegenheid van de jongste *Equal Pay Day*. Maar we weten ook dat deze vaststelling in haar algemeenheid niet opgaat. Als we ietwat meer in detail gaan of meer differentiaties tussen verschillende maatschappelijke groepen verschijnt een ietwat complexer beeld. Hoger op de loonladder blijkt het verschil minder groot dan lager. Voor bepaalde sectoren zijn de verschillen groot; voor andere nagenoeg onbestaande. Als men vrouwen en mannen vergelijkt die nooit trouwen, geen kinderen hebben en rond de leeftijd van 45 cirkelen dan doen vrouwen het zelfs beter. Vanuit beleidsoogpunt is dit meer complexe beeld wellicht het betere uitgangspunt. Want men kan vermoeden dat deze complexiteit het gevolg is van het bestaan van meerdere mechanismen die de ongelijkheid tussen mannen en vrouwen blijven veroorzaken en dus is het erg onwaarschijnlijk dat één algemene ingreep de kwaal kan doen verdwijnen. Omdat de onderzoeksliteratuur over dit onderwerp ondertussen oeverloos begint te worden en een selectieve lectuur van het materiaal daarvoor bijna onvermijdelijk is, zijn gedegen overzichten onvoorstelbaar nuttig. Wie door de bomen het bos niet meer ziet, heeft de keuze uit meerdere mogelijkheden. Een recent, erg uitgebreid overzicht in de vorm van een zogeheten meta-analyse is Doris Weichselbaumer en Rudolf Winter-Ebmer. 2005. A Meta-Analysis on the International Gender Wage Gap. *Journal of Economic Surveys*, 19(3): 479-511.

De discussie over de 'gender wage gap' zal dus nog wel een tijdje voortduren. Een belangrijk aspect van deze discussie is of men deze verschillen in beloning van mannen en vrouwen nu al dan niet als 'discriminatie' mag of moet beschouwen. Misschien verrast het u dat deze kwalificatie een punt van discussie blijft. Maar het moet voor iedereen duidelijk zijn dat 'beschrijven' en 'beoordelen' twee verschillende registers zijn. Constateren dat er (nog steeds al dan niet) sprake is van een loonverschil tussen mannen en vrouwen is één zaak. Een andere zaak is of men dit verschil moet toeschrijven aan regelrechte discriminatie of aan (eventueel moeilijk te meten) factoren die verschillen in de productiviteit van mannen en vrouwen veroorzaken. Anders:

tussen 'beschrijven' en 'beoordelen' zit er zoiets als 'verklaren' en het hangt zeer sterk af van de verklaring die men aanhangt of men zal spreken van discriminatie of niet, zelfs als dezelfde gegevens ten grondslag liggen aan beide interpretaties. Hoezeer beleidsmakers en leken dit ook betreuren, 'wetenschappelijke verklaringen' zijn omzeggens nooit helemaal eenduidig terug te brengen op wat men empirisch constateert. Ze vergen beoordeling en laten daarom onvermijdelijk ruimte voor waardeoordelen. Zo ook in het geval van de 'gender wage gap' – al was het maar omdat de terminologie waarmee men de geconstateerde en niet door productiviteit te verklaren loonverschillen tussen mannen en vrouwen typeert gemakkelijk aanleiding geeft tot waardeoordelen. 'Discriminatie' zeggen de enen, 'onverklaarde verschillen' zeggen de anderen.

In *Rhetoric in Economic Research: The Case of the Gender Wage Differentials*. 2006. *Industrial Relations*, 45(3): 416-436 gebruiken Doris Weichselbaumer en Rudolf Winter-Ebmer de literatuur die zij verzamelden voor hun meta-analyse om te onderzoeken welke factoren economen er eventueel toe brengen om (schijnbare) waardeoordelen in hun conclusies te verweven.

Deze literatuur over de gender wage gap is een excellente case voor een dergelijk onderzoek, zeggen Weichselbaumer en Winter-Ebmer. Al jaren bestaat er immers een door iedereen aanvaarde techniek – de zogeheten Binder-Oaxaca decompositietechniek – die men gebruikt om de geconstateerde loonverschillen toe te wijzen aan twee soorten factoren: een part dat wordt verklaard door productiviteitsverschillen en een part dat hierdoor niet te verklaren valt. Omdat zowat iedereen deze techniek gebruikt en de feiten dus niet behept zijn door voorkeuren omtrent hoe men de verschillende delen van de 'gender wage gap' berekent, kan men hier duidelijker dan voor andere onderzoekstopics een onderscheid maken tussen wat men constateert en hoe men het 'residu' – het tweede part van de geconstateerde loonverschillen – benoemt: 'discriminatie' of 'onverklaarde verschillen'. Maar wat bepaalt of auteurs de ene dan wel de andere omschrijving bezigen? Ideologie? Leeftijd? Gender? Academische status? Kwaliteit van het onderzoekswerk? Betrouwbaarheid van de gebruikte gegevens? Aard van de onderzoekspopulatie? De groot-

te van de geconstateerde loonverschillen? Of nog iets anders?

Om dit te onderzoeken verzamelden Weichselbaumer en Winter-Ebmer een kleine tweehonderd gepubliceerde artikels, die gebruik maken van de Binder-Oaxaca decompositietechniek. Een eerste, ietwat verrassende vaststelling is dat de term 'discriminatie' doorheen de tijd minder frequent gebruikt wordt. In de jaren tachtig benoemt ongeveer 60% van de artikels het residu als 'discriminatie'; op het einde van de jaren negentig gaat het om nog slechts 30%. Als men enkel op de titels afgaat, daalt het gebruik van de term 'discriminatie' van 43% in de jaren tachtig tot 21% in de jaren negentig. Maar om eventuele problemen uit de weg te gaan die samenhangen met de positieve of negatieve connotaties van de term discriminatie in de inleiding of het besluit van de artikels focussen Weichselbaumer en Winter-Ebmer op de terminologie bij het bespreken van het berekende effect.

Een algemene vaststelling is dat hoe groter het geconstateerde verschil des te groter is de kans dat men het resultaat omschrijft als 'discriminatie'. Niet totaal onverwacht zou je denken. Maar minder verwacht is dat vrouwelijke auteurs niet minder, maar ook niet meer dan hun mannelijke collega's geneigd blijken om de term 'discriminatie' (of 'onverklaarde verschillen') te gebruiken. Opvallend is ook dat bij coauteurschap van vrouwen en mannen de kans dat men wijst op discriminatie twintig procentpunten lager ligt. Coauteurs van hetzelfde geslacht hebben dit effect niet. "[...] Maybe women feel they have to prove to their male coauthors that their scientific stance is not affected by personal experience." (p. 424) Auteurs met een duidelijke interesse, gezien vroegere publicaties hierover, voor de 'gender wage gap' gedragen zich op het vlak van de terminologie niet anders dan wie hierover niet publiceerde. Met Amerikaanse universiteiten geaffilieerde auteurs gebruiken de term 'discriminatie' 25 procentpunten minder dan hun vakgenoten uit andere landen. Toptijdschriften vermijden het gebruik van de term. De kans ligt iets minder dan dertig procentpunten lager dat men er het woord 'discriminatie' gebruikt. Of dit te maken heeft met de (vrijwillige) keuze van de auteurs dan wel met beïnvloeding door de redactie van het tijdschrift is op basis van deze gegevens onmogelijk na te gaan. Wel is er een duidelijk verband tussen het

gebruik van de term en het aantal artikels dat een auteur voordien in toptijdschriften publiceerde. Een voorafgaand artikel in een toptijdschrift reduceert de kans op het gebruik van de term 'discriminatie' met om en bij de veertig procentpunten.

Weichselbaumer en Winter-Ebmer besteden aparte aandacht aan de beleidsaanbevelingen die men in de geselecteerde artikels terugvindt. Waarom pleiten auteurs voor beleidsinterventies: omdat ze bijzonder sterk overtuigd zijn van hun empirische resultaten of omwille van hun ideologische overtuiging? Voor deze analyse construeren ze drie indicatoren: een schaal die de *gestrengheid* van de voorgestelde maatregelen meet, een index voor de *vastberadenheid* waarmee de maatregelen worden bepleit en het *aantal woorden* dat men besteedt aan beleidsaanbevelingen. De analyse is eerder ontgoochelend. Geen enkele van de gebruikte variabelen (demografische karakteristieken, indicatoren van overtuigingen, kwaliteit en betrouwbaarheid van de resultaten) zijn in staat om de verscheidenheid in de drie gebruikte maatstaven te verklaren: "neither female authors, nor those with a longer history in gender-related research, nor those having used the term *discrimination* in the interpretation of the binder-Oaxaca decomposition require more policy changes" (p. 429).

In zekere zin, besluiten Weichselbaumer en Winter-Ebmer, is dit resultaat geruststellend. Als men de geldigheid van de gebruikte indicatoren niet betwist dan blijken geslacht, ideologie en waardeoordelen de beleidsaanbevelingen van economen niet te beïnvloeden. Zelfs auteurs die het geconstateerde loonverschil als 'discriminatie' benoemen, onderscheiden zich op het vlak van de beleidsaanbevelingen niet van hun vakgenoten die spreken over 'onverklaarde verschillen'. "This means that there are differences in the rhetoric of economists, but these do not explain the policy prescriptions of authors." (p. 429) Er zijn ook geen aanwijzingen dat auteurs bewust gebruik zouden maken van meer restrictieve datasets zodat enkel omwille van de aard van het gebruikte materiaal de kans toeneemt op een grotere (of kleinere) 'gender wage gap'. "Only 4% of authors worked with data restrictions that might increase the gap compared to the entire population [...] These results are reassuring again: Authors might differ in their assessment and their rhetoric describing a phenomenon, but this does

not influence important decisions in the course of the empirical work.” (p. 432)

Natuurlijk, zoals voor elk onderzoek kan men de nodige kanttekeningen maken bij de gebruikte indicatoren. Trouwens, men kan zich afvragen of de vaststelling dat ook de kwaliteit en de betrouwbaarheid van het onderzoek geen invloed heeft op de beleidsaanbevelingen wel zo geruststellend is. Want als zelfs de overtuiging dat men kwaliteitsvol werk levert niet leidt tot een verduidelijking van de maatschappelijke betekenis van deze resultaten, wat heeft dit effect dan wel?

En als toemaatjes: alweer iets uit het leven gegrepen

Eerlijk, het lag niet in mijn bedoeling om hiervan een vaste rubriek te maken. Maar voorbeelden van de manier waarop de kleine lettertjes van courante regelgeving slachtoffers maken, blijven verbazen én liggen kennelijk voor het grijpen. Vandaar.

Neem nu het geval van een jongedame die met succes een opleiding verpleging voltooit. De vorige jaren werkte ze in de zomervakantie om haar studies te betalen. Dus: waarom niet op dezelfde manier ook dit laatste jaar uit de schoolperiode afsluiten – een beetje als orgelpunt en afsluiting van een periode in haar levensloopbaan. Daarna begint ze ineens te werken in een echte baan mét een echt contract in een groot en respectabel ziekenhuis. De baan valt best mee. Verpleging is iets wat ze voor de rest van haar leven wel wil doen. Niet noodzakelijk in loondienst, maar na voldoende ervaring

eventueel als zelfstandige in de thuisverpleging. Edoch. Na enkele maanden tijd blijkt de jongedame allergisch aan de producten waarmee ze heel de tijd in aanraking komt. Occasioneel contact tijdens de opleiding leidde niet tot problemen. Permanent contact is een andere zaak. Voorlopig resulteert dit enkel in huiduitslag maar, zo luidt de diagnose, op termijn kan dit leiden tot ontwikkeling van astma en andere kwalijke zaken. Advies van de arbeidsgeneesheer: zo snel mogelijk deze baan opgeven. Baan opgeven? Weg carrière in de verpleging, zeker! En weg ook het plan om samen met haar vriend een eigen woonst te zoeken. Want wat blijkt: de jongedame schreef zich na het verlaten van de school niet in als werkzoekende bij de VDAB – want toch een baan met een echt contract, nietwaar. Maar ondertussen heeft ze niet voldoende gewerkte dagen verzameld om uitkeringsgerechtigd te zijn. Dus: weg niet enkel carrière, maar tot na de wachttijd voorlopig ook geen middelen van bestaan die een transitie naar een ander loopbaanperspectief – een D-cursus in verpleging, een opleiding in kinderverzorging, tot kleuterleidster of iets dergelijks – ook psychologisch iets gemakkelijker verteerbaar zouden maken.

Ik moet het eerlijk toegeven. Dit verhaal verraste mij ietwat. Maar als ik de zoon van vrienden mag geloven, komt een dergelijke ervaring bij jongeren wel meer voor. Dus: werk aan de winkel voor wie invloed heeft op de inhoud van de kleine lettertjes!

Walter Van Trier