

Mathiness of truthiness in de economische wetenschap

Harrie Verbon

Paul Romer, grondlegger van de endogene groeitheorie, verwijt economen wiskunde te gebruiken om er academische politiek mee te bedrijven in plaats van er mee de economische wetenschap te dienen. Hij noemt dit mathiness. Zijn opvatting over wat het doel van (economische) wetenschap is, namelijk het bereiken van op consensus gebaseerde waarheid, is echter niet te handhaven. In de economie bestaat 'de' waarheid niet; welk economisch model actueel is, is afhankelijk van plaats en tijd. Ook Romer's opvattingen over de link tussen theoretische variabelen en waarneembare verschijnselen is discutabel.

1 Inleiding

Romer (2015) verwijt economen het gebruik van wiskunde om er academische politiek mee te bedrijven. *Mathiness*, zoals hij dat noemt, wordt door academici gebruikt om er hun eigen academische positie mee te dienen, maar niet om tot eensluidende kennis over de waarheid te komen. Hij noemt als voorbeeld de groeitheorie, zijn eigen vakgebied, waarin volgens hem de laatste twee decennia geen “wetenschappelijke vooruitgang in de richting van een consensus” heeft plaats gevonden.

De stelling van Romer heeft tot veel debat geleid dat vooral plaatsvond op blogs. Uit de reacties bleek dat meerdere economen onvrede hebben over het gebruik van wiskunde. Noah Smith, bijvoorbeeld, merkte op dat *mathiness* niet een probleem is van dat handjevol door Romer met name genoemde economen, maar van de hele professie op het gebied van de macro-economie. Hij schrijft dat wiskundige vergelijkingen in de macro-economie meestal niet ondersteund worden door empirisch bewijs. De Euler-vergelijking bijvoorbeeld, die een verband aangeeft tussen de groei van consumptie en de rentevoet, blijkt empirisch nooit te werken als verondersteld door de theorie. Toch is de Euler-vergelijking de kern van bijna ieder macro-economisch model (Smith 2015).

Er zijn echter ook tegengeluiden. Stephen Williamson, de vicepresident van de Federal Reserve Bank van St. Louis, betoogt dat Romer de beoefenaars van de economische wetenschap wil dwingen in een keurslijf dat alleen Romer zelf goed past. Williamson betoogt dat alle modellen noodzakelijkerwijs in een bepaalde richting onjuist zijn. Modellen kunnen echter radicaal van elkaar verschillen en toch in overeenstemming met dezelfde empirische verschijnselen. Waarom zouden we deze diversiteit niet juist als een zegen zien? (Williamson 2015).

In dit artikel wordt de wetenschapsfilosofie die Romer impliciet veronderstelt bij zijn opvatting over *mathiness* aan een nader onderzoek onderworpen (hoofdstuk 2). Die filosofie blijkt contraproductief te zijn, omdat hij tegenspraak wil uitsluiten, terwijl de economische wetenschap juist vooruitgang boekt door de concurrentie tussen tegengestelde opvattingen over de verklaring van verschijnselen. Daarnaast is zijn aanname over de verhouding tussen theoretische grootheden, zoals kapitaal, en waarneembare verschijnselen, zoals de kapitaalgoederenvoorraad die uit de nationale rekeningen afgeleid kan worden, methodologisch onhoudbaar (hoofdstuk 3). Vanuit Romer's visie valt te begrijpen waarom hij de Engelse kant van de Cambridge-Cambridge-discussie over de meetbaarheid van kapitaal achteraf van *mathiness* beschuldigt (hoofdstuk 4), maar het is minder evident dat ook Piketty van *mathiness* wordt beschuldigd door Romer. Piketty (2014) heeft immers veel energie besteed aan het meetbaar maken van kapitaal door de eeuwen heen (hoofdstuk 5). In hoofdstuk 6 besteden we aandacht aan een meer moderne opvatting die economiebeoefening ziet als het opbouwen van een verzameling cases. Deze opvatting gaat impliciet terug tot de opvatting van J.M. Keynes, die het als taak van economen zag op het juiste moment en op de juiste plaats het meest geschikte model uit de aanwezige voorraad van modellen te kiezen. We eindigen met conclusies (hoofdstuk 7).

2 Wetenschapsfilosofie

Wat precies misbruik van wiskunde is, raakt aan de vraag hoe de verhouding tussen wiskunde en de werkelijkheid in de economische wetenschap is, of zou moeten zijn, en welke wetenschapsfilosofie onderzoekers er op na kunnen of moeten houden. Wat de wetenschapsfilosofie betreft, Romer begint zijn AER-artikel met de definitie van wetenschap als een proces dat tot consensus leidt over theoretische en empirische stellingen die waar zijn: "Science is a process that does lead to a broadly shared consensus. It is arguably the only social process that does. Consensus forms around theoretical and empirical statements that are true" (Romer, 2015, p.89). Hij vervolgt dat, in tegenstelling tot in de wetenschap, er in de politiek niet naar consensus wordt gestreefd. Daarom gebruikt men in de politieke arena suggestieve of dubbelzinnige woorden om die tegenstellingen tussen partijen te overdrijven. Die gewoonte begint nu, volgens Romer, ook academia binnen te kruipen. Romer haalt diverse auteurs van papers op het gebied van de groeitheorie aan als voorbeelden van het bedrijven van academische politiek met behulp van wiskunde "presumably because they were losing the scientific debate (Romer, 2015, p.90)".

In de bovenstaande quotes beschouwt Romer de convergentie naar consensus in de wetenschap als een feitelijke eigenschap ("does lead to..."), maar indirect ook als een gewenste eigenschap, want academische economen, die de tegenstelling zoeken, begeven zich volgens Romer op het pad van de 'academische politiek'. Dit oordeel velt Romer in zijn AER-paper over het werk van een aantal toonaangevende economen, waaronder Lucas, Prescott en Piketty.

De stelling dat het wetenschappelijke bedrijf tot consensus zou leiden over de kern van de wetenschappelijke kennis, is eerder door Kuhn (1960) beroemd gemaakt. Kuhn beperkte zich tot de exacte wetenschappen. Volgens hem bestaan er 'paradigma's' die door de onderzoekers worden gedeeld en die leidend zijn voor de onderzoekprogramma's. 'Normaal' onderzoek heeft voornamelijk tot doel de kennis over de gevolgen van het gedeelde paradigma te verdiepen of om de 'puzzels' op te lossen die het paradigma opwerpt. Dergelijk onderzoek kan echter ook onoplosbare paradoxen of tegenstrijdigheden genereren. Het zal echter niet gauw leiden tot de verwerping van het paradigma. Het onvermogen om een puzzel op te lossen, zal eerder aan de onderzoeker dan aan de theorie worden geweten.

Volgens Kuhn is wetenschappelijk onderzoek erop gericht de theorie te bevestigen, niet om die te verwerpen. Hiermee beschrijft Kuhn de praktijk van het wetenschappelijk onderzoek, zoals hij dat waarnam. Consensus is in zijn waarneming een doel dat wetenschappers zelf nastreven. Dat kan echter volgens Kuhn slechts een doel zijn zolang de paradigma's gedeeld worden. Een onwrikbaar 'geloof' in het paradigma zou op de duur contraproductief zijn, volgens Kuhn, omdat de paradoxen en tegenstrijdigheden binnen een gegeven paradigma zo'n grote omvang kunnen aannemen dat er ruimte ontstaat voor een nieuw paradigma dat dan, tot nader order, leidend zal zijn voor het wetenschappelijke onderzoek.

Wanneer is het dan onwetenschappelijk om van het paradigma af te wijken, of, in de terminologie van Romer, wanneer wordt een afwijking van de consensus 'academische politiek'? Neem als voorbeeld de 'breed gedeelde' consensus dat in ons zonnestelsel de zon het middelpunt is, waar omheen de andere planeten, inclusief de aarde, draaien. Men zou alle wetenschappelijke bewijzen hierover moeten negeren om toch, zoals de kerk lang gedaan heeft, een geocentrisch standpunt te kunnen blijven verdedigen. Anderzijds kunnen tegengestelde interpretaties van waarneembare verschijnselen eeuwenlang naast elkaar bestaan, zonder dat beide interpretaties in strijd zijn met methodologische beginselen. Een voorbeeld is de golf- versus de deeltjesinterpretatie van licht. Dit verschil heeft tot op heden slechts tot de 'consensus' geleid dat beide interpretaties waar kunnen zijn.

Terwijl Kuhn het wetenschappelijk bedrijf ziet als werken onder een gedeeld paradigma dat pas wordt los gelaten als dit onvermijdelijk wordt door de omvang van de onopgeloste puzzels rond het paradigma, beschouwt Popper (1934) de wetenschappelijke praktijk erop gericht zijn eigen theorieën te verwerpen. Een theorie moet zo geformuleerd worden dat de kans dat hij wordt verworpen zo groot mogelijk is. Een kwetsbare theorie die desondanks niet verworpen wordt, is een sterke theorie. Dit uitgangspunt van 'falsificatie' heeft onder economisch methodologen veel invloed gekregen. Volgens Blaug (1992) ging die invloed echter niet zo ver dat falsifieerbaarheid van theorieën à la Popper de overheersende onderzoeksfilosofie onder economen was: "The prevailing methodological mood is not only highly protective of received economic theory (...): almost any model will

do provided it is rigorously formulated, elegantly constructed, and promising of potential relevance of real-world situations” (Blaug 1992: 110/111).

Op het eerste gezicht lijkt het dus dat het economisch onderzoek eerder de lijn van Kuhn dan de lijn van Popper volgt. Die indruk wordt versterkt door het bestaan van een centraal paradigma in de economische theorie. Dat paradigma blijkt, zoals (Rosenberg 2009) schrijft, uit het feit dat de leerboeken in de economie allemaal op elkaar lijken en dat economische ‘wetten’ in wiskundige vorm worden geformuleerd, die op diverse domeinen kunnen worden toegepast. Maar, zoals Rosenberg eerder verklaarde, hoewel er na de Tweede Wereldoorlog een steeds grotere nadruk op (wiskundige) techniek is komen te liggen bij het formuleren van theorieën, en er daardoor een grotere verklarende kracht aan de theorie kan worden toegekend, blijft de voorspellende kracht ver achter. In de natuurkunde kan men hopen dat experimenten vroeg of laat de juistheid of de onjuistheid van een hypothese kunnen aantonen. In de economie zijn dergelijke beslissende experimenten ondenkbaar. De economische theorie blijft daarom steken in de fase van toegepaste wiskunde, waarvan de resultaten empirisch niet getoetst kunnen worden. (Rosenberg 1992).

Er is een opvallende analogie tussen het oordeel van Rosenberg uit 1992 over de status van de economische wetenschap en het *mathiness*-begrip zoals dat door Romer werd gebruikt. Volgens beide auteurs kan wiskunde leiden tot theorieontwikkeling zonder een grotere empirische voorspelkracht. Theorie is dan niet veel meer dan het oplossen van puzzels zonder enige relevantie voor de economische werkelijkheid. Het verschil tussen Rosenberg en Romer is echter dat terwijl Romer zijn oordeel van *mathiness* op sommige wetenschappers werkzaam in de macro-economische groeitheorie van toepassing verklaart, Rosenberg zijn oordeel van ‘toegepaste wiskunde zonder relevantie’ van toepassing verklaart op de hele professie. Volgens Rosenberg (1992) is er geen hoop voor de economie als wetenschap, volgens Romer echter geldt dit alleen voor (groei)theoretici die hun gelijk wensen te halen en daarmee de wetenschap niet dienen. In feite is Romer optimistischer dan Rosenberg over de mogelijke vooruitgang van de economische wetenschap.

Anderzijds is de wetenschapsfilosofische opvatting van Romer, in tegenstelling tot die van Kuhn, veel meer rigide. Volgens Kuhn is het op een gegeven moment onvermijdelijk dat een gangbaar paradigma vervangen wordt door een ander paradigma dat een grotere verklarende kracht heeft dan het voorheen fungerende paradigma. Het is niet duidelijk of Romer iedere afwijking van het paradigma verwerpt. In ieder geval geldt dit volgens hem als de afwijking de vorm aanneemt van *mathiness*, oftewel als de theorie geen duidelijke band heeft met waarneembare verschijnselen.

3 Waarneembare verschijnselen en de groeitheorie

Volgens Paul Romer moeten de gebruikte symbolen in een economische theorie een duidelijk verband hebben met waarneembare verschijnselen. Romer beschouwt het beroemde artikel van Robert Solow uit 1956 daar een goed voorbeeld van. Dit artikel van Solow dat door velen als de start van de neo-klassieke groeitheorie beschouwd wordt, geeft een modelmatige beschrijving van economische groei met een neo-klassieke productiefunctie. Solow maakte gebruik van de abstractie dat een economie goederen en diensten produceert met twee productiefactoren, namelijk arbeid en kapitaal. Wat kapitaal betreft: “Solow’s (1956) mathematical theory of growth mapped the word ‘capital’ onto a variable in his mathematical equations, and onto both data from national income accounts and objects like machines or structures that someone could observe directly” (Romer 2015: 89).

Het beginsel van waarneembaarheid is echter voor het ‘succes’ van een wetenschap niet een noodzakelijk vereiste. Veel door filosofen en wetenschappers geïntroduceerde fysische begrippen waren niet altijd direct waarneembaar. Het meest duidelijk is dat bij het begrip atoom. In het oude Griekenland speculeerde 400 jaar v. Chr. de filosoof Democritus al over het bestaan van atomen, de kleinst denkbare deeltjes van iedere vorm van materie. Het zou nog ruim 2000 jaar duren voor de wetenschap in staat was het bestaan van atomen daadwerkelijk aan te tonen. Toch waren er in de tussentijd, gebruikmakend van de hypothese van het bestaan van (vooralsnog niet waarneembare) atomen, theorieën geformuleerd die waarneembare verschijnselen konden verklaren.

Variabelen in de theorie hoeven dus niet waarneembaar te zijn om toch wetenschappelijk van nut te kunnen zijn. Het is daarom weinig zinvol om aan de hand van Romer’s waarneembaarheids criterium over het wetenschappelijke van economische theorie te oordelen. Het criterium van Rosenberg (1992) kan hier behulpzamer zijn. Volgens Rosenberg is een nieuwe theorie wetenschappelijk als die theorie waarneembare verschijnselen beter kan verklaren en voorspellen (“empirically progressive” in zijn woorden).¹ De macro-economische groeitheorie kan aan de hand van dit criterium op zijn ‘wetenschappelijkheid’ getoetst worden. De neo-klassieke groeitheorie, zoals uiteengezet door Solow (1956), bevat als belangrijkste constatering dat de economie, ceteris paribus, naar een steady state convergeert waarin geen groei van het nationaal product als gevolg van investeringen in kapitaal meer mogelijk is. De reden voor dit ‘non-resultaat’ is het bestaan van afnemende marginale productiviteit van kapitaal. Hierdoor zal het voorbij een gegeven hoeveelheid kapitaal niet meer lonen om meer te investeren dan noodzakelijk is voor het in standhouden van de kapitaalgoederenvoorraad per hoofd. Mocht er desondanks nog sprake zijn van economische groei, dan kan die alleen het gevolg zijn van een exogeen bepaalde toename van de productiviteit van de productiefactoren.

¹ Rosenbergs wetenschapsfilosofie is ontleend aan het werk van Lakatos dat dateert uit 1970. Voor een bespreking van Lakatos, zie Boumans en Davis (2016).

Het neo-klassieke groeiemodel heeft een tamelijk simpele en voor de hand liggende voorspelling, namelijk dat minder ontwikkelde economieën sneller groeien dan ontwikkelde economieën. Afgezien van de vraag hoe deze hypothese empirisch getoetst zou moeten worden (waarover hieronder meer), kan deze voorspelling alleen al op basis van oppervlakkige waarneming 'gefalsifieerd' worden. Veel laag ontwikkelde economieën groeien structureel ondermaats, al zijn er klassieke tegenvoorbeelden, zoals, onder meer, China, India en andere zuidoost-Aziatische landen.

De neo-klassieke groeitheorie kon het ontbreken van 'catching-up' in sommige laag ontwikkelde economieën dus niet verklaren. De endogene groeitheorie, aan de ontwikkeling waarvan Romer een belangrijke bijdrage heeft geleverd, voorzag in deze lacune. De belangrijkste verandering ten opzichte van het neo-klassieke groeiemodel is dat er niet langer wordt uitgegaan van een afnemende marginale productiviteit van kapitaal. Door de introductie van menselijk kapitaal in de productiefunctie en door aan te nemen dat menselijk kapitaal afhangt van de aanwezige omvang van het fysieke kapitaal, krijgt fysiek kapitaal op macroniveau een dubbele rol in de productie. Als gevolg daarvan dalen de meeropbrengsten van kapitaal niet, waardoor endogeen continue groei van het nationaal product mogelijk wordt.

Was deze introductie van 'menselijk kapitaal' als een extra variabele in de productiefunctie "empirically progressive" in de terminologie van Rosenberg en dus geen voorbeeld van *mathiness* à la Romer? Empirisch gezien hangt menselijk kapitaal af van de kwaliteit van het onderwijssysteem. Veel empirische studies hebben daarom proberen vast te stellen of de kwaliteit van het onderwijs van een land samenhangt met de economische groei. Een inmiddels klassieke studie in dit verband is Mankiw et al. (1992), die menselijk kapitaal aan de productiefunctie toevoegden en uit een cross-sectie van landen een verband tussen de kwaliteit van het onderwijs en de hoogte van het bruto binnenlands product afleidden. Zij concludeerden "that the Solow model is consistent with the international evidence if one acknowledges the importance of human as well as physical capital" (Mankiw et al. 1992: 433). Barro (2013), daarentegen, die recent dergelijke *cross-country*-regressies nog eens overdeed, concludeerde dat endogene groei modellen geschikt zijn om te begrijpen waarom ontwikkelde economieën kunnen blijven groeien, terwijl de neo-klassieke groeitheorie geschikt is om relatieve groeiverschillen tussen landen te begrijpen. Deze boodschap dat beide theorieën soms waar kunnen zijn, begint ook tot de leerboeken door te dringen. Gärtner (2013, p. 303), bijvoorbeeld, concludeert op basis van een cross-sectie van landen dat een hogere onderwijskwaliteit (gemeten door gemiddelde jaren van scholing per persoon) leidt tot een hoger nationaal inkomen per hoofd, maar toont ook aan dat landen met hogere kapitaalinvesteringen hogere groeivoeten kennen (Gärtner 2016, p. 307). De eerste conclusie bevestigt de neo-klassieke denkwijze, maar de tweede conclusie duidt op de relevantie van de endogene groeitheorie.

Het lijkt er dus op dat de macro-economische groeitheorie zich in eenzelfde fase bevindt als de deeltjes- versus de golveninterpretaties van licht die vanaf de 17e eeuw tot ver in de

20e eeuw tot tegenstrijdige voorspellingen leken te leiden. Een overtuigend empirisch bewijs op basis van waarnemingen van cross-secties van landen dat een van beide benaderingen ‘waar’ is, komt er hoogstwaarschijnlijk niet. Solow (1994) kon dit soort exercities in ieder geval niet waarderen. In zijn visie kan de groei van zeer verschillende landen niet verklaard worden door ervan uit te gaan dat de waarnemingen verschillende punten van eenzelfde universele productiefunctie beschrijven.

Dit raakt aan een meer algemeen probleem van macro-economische theorie, namelijk dat gecontroleerde empirische testen van de theorie ondenkbaar zijn en de resultaten altijd afhankelijk blijven van plaats en tijd, zodat universele geldigheid ontbreekt. Keynes constateerde dit al in de jaren 1930. Hij keurde de pogingen om economische relaties statistisch te toetsen sterk af, zoals hij in de beroemde discussie met Tinbergen uiteenzette.² Keynes zag de economische wetenschap als een vorm van logica. Zodra men feitelijke getallen gaat invullen voor variabele functies verliest een model volgens Keynes zijn algemeenheid. Het volgende fragment uit een brief aan Harrod (gedateerd 4 juli 1938), dat door veel methodologen is geciteerd (zie bijvoorbeeld Blaug 1992: 79), geeft aan hoe Keynes de rol van modellen versus de economische werkelijkheid zag: “Economics is a science of thinking in terms of models joined to the art of choosing models which are relevant to the contemporary World” (Johnson en Moggridge 1978: 296).

4 Het Cambridge-Cambridge kapitaaldebat

De wetenschapsopvatting van Keynes is ver verwijderd van het criterium van ‘empirical progressivity’ van Rosenberg, maar ook ver weg van het *mathiness*-begrip van Romer, want de vraag of een theorie zich baseert op waarneembare variabelen (zoals Romer eist) is in de opvatting van Keynes niet relevant. Empirische toetsing van economische theorieën aan de hand van waarneembare variabelen degradeert theorieën immers tot cases met een beperkte geldigheid.

Er is bovendien, wat Romer’s criterium van waarneembaarheid betreft, nog een tweede probleem, namelijk dat niet alle academische economen eensluidend zijn over de vraag of macro-economische variabelen objectief waarneembaar zijn, of dat de metingen het gevolg zijn van definities van deze variabelen die wellicht niet logisch consistent zijn. Dit laatste geldt met name voor de grootte van kapitaal. Romer gaat ervan uit dat de macro-economische variabele kapitaal zoals die in de theorie wordt gebruikt waarneembaar is, zowel direct (via materiële zaken als machines en gebouwen) als indirect via de nationale rekeningen.

² Zie Boumans en Davis (2016) voor een compacte uiteenzetting van dit debat. Zie ook Verbon (2013) waar de afkeer van Keynes voor regressieonderzoek gerelateerd wordt aan de meer ‘moderne’ afkeer van Taleb, 2007, voor het doortrekken van lijnen.

Kapitaal is in de theorie een abstractie op macroniveau die op een of andere manier verband houdt met materiële zaken die bij de productie gebruikt worden. Alle materiële zaken die in een land gebruikt worden bij de productie van goederen en diensten moeten derhalve opgeteld worden tot de grootte 'kapitaal' dat samen met 'arbeid' (al evenzeer een abstractie) de mogelijke en feitelijke omvang van de productie bepaalt. Sinds de opkomst van de endogene groeitheorie in de jaren 1980 lijkt het er op dat er een eensluidend oordeel is over het kapitaalbeprijp, zoals (Cohen en Harcourt 2003, p. 200) schrijven: "When theories of endogenous growth (...) took off in the 1980s using aggregate production functions, contributors usually wrote as if the controversies [on measurement of capital] had never occurred."

Cohen en Harcourt (2003)³ herinneren ons eraan dat, in tegenstelling tot wat Romer (2015) beweert, de meetbaarheid van kapitaal allermindst boven iedere twijfel verheven is. De Cambridge-Cambridge discussie duurde van midden jaren 50 tot midden jaren 70. De Engelse Cambridge economen, onder aanvoering van Joan Robinson, ontkenden de meetbaarheid van de geaggregeerde kapitaalvoorraad. In de productie worden heterogene kapitaalvoorraaden gebruikt die niet zomaar opgeteld kunnen worden. Men zou afzonderlijke kapitaalvoorraaden kunnen waarderen door het berekenen van de contante waarde van de toekomstige stroom van inkomsten. De aldus verkregen bedragen kunnen dan worden opgeteld om een meting van geaggregeerd kapitaal te verkrijgen. De verdiscontering betekent dat de rente gebruikt moet worden, maar die rente wordt in de neo-klassieke theorie bepaald door de omvang van de geaggregeerde kapitaalvoorraad. We stuiten dus op een cirkelredenering die niet alleen betrekking heeft op het meten van kapitaal maar, bijgevolg, ook op het empirisch schatten van het kapitaalaandeel in de productie.

Backhouse (2014) laat zien dat Solow de meetbaarheid van kapitaal ook niet boven iedere twijfel verheven vond, maar de geaggregeerde productiefunctie met kapitaal en arbeid als productiefactoren slechts als een heuristisch middel beschouwde om macrodata mee te analyseren. Het Solow model "abstracted from heterogeneity of capital in the same way that it abstracted from problems of unemployment, money and finance" (Backhouse 2014: 260). Solows opvatting lijkt op een directe falsificatie van de stelling van Romer dat een theorie moet gaan over waarneembare verschijnselen. Met theorie mag men pragmatisch omgaan, ook al leidt dat tot interpretatieproblemen.

Zoals met andere theoretisch-economische problemen die moeilijk zijn op te lossen (andere voorbeelden zijn de onmogelijkheidsstelling van Kenneth Arrow om individuele nutten tot sociale welvaartsfuncties te aggregeren, of de Allais-paradox in verwachte nutsmaximalisatie, zie Kahneman 2011, p. 213), is het probleem rond de waardering van

³ Cohen en Harcourt zijn niet onpartijdig in het kapitaaldebat. Harcourt gaf, bijvoorbeeld, in 1969 een overzicht van dit debat waarin hij 'partij trok' voor de Cambridge-Engeland zijde. Blaug uitte, daarentegen, in 1974 felle kritiek op de Cambridge-Engeland benadering, want, hoewel de voorgestelde alternatieve kapitaaltheorie logisch consistent was, vloeide er naar zijn inzicht geen enkele toetsbare voorspelling uit voort. Zie Mata (2004) voor een 'neutraler' overzicht van dit debat.

de macro-economische kapitaalgoederenvoorraad uit het zicht verdwenen. Er wordt aangenomen dat heterogene kapitaalgoederen kunnen worden opgeteld en dat verdiscontering van toekomstige opbrengsten niet tot een cirkelredenering leidt. De (onbedoelde) bijdrage van Romer (2015) aan deze oude controversie zou kunnen zijn dat deze weer in het bewustzijn van economen opduikt. Romer beschuldigt Joan Robinson immers postuum van *mathiness*: “Joan Robinson (1956) was engaged in academic politics when she waged her campaign against capital and the aggregate production function.” en: “Then too, a faction like Robinson’s that risked losing a battle might resort to mathiness as a last-ditch defense ...” en: “Where would we be now if Robert Solow’s math had been swamped by Joan Robinson’s mathiness” (Romer, p.93). De positie die Romer hier inneemt is kennelijk die van Blaug (2009), namelijk dat de Cambridge-Engeland benadering weliswaar logisch consistent, maar empirisch niet relevant is. Opvallend is dat ook de politiek geëngageerde econoom Thomas Piketty in deze discussie door Romer onder vuur wordt genomen.

5 *Mathiness* door Thomas Piketty

De Cambridge-Engeland kant van de kapitaaldiscussie had een overduidelijke ideologische motivering. Joan Robinson vond kapitaal vooral van belang omdat het eigendom ervan “confers on capitalists the legal right and economic authority to take a share of the surplus created by the production process” (Cohen en Harcourt, 2003, p. 208). Robinsons inzet bij het debat had dus een Marxistische ondertoon waarbij niet zozeer de omvang, maar eerder de functie van kapitaal centraal werd gesteld. Die functie was dat, in de Marxistische visie, ‘kapitalisten’ dankzij het bezit van ‘kapitaal’ arbeiders konden uitbuiten. Vergelijken we deze visie met de visie van Thomas Piketty, die met zijn boek uit 2014 de discussie over kapitaal weer in het centrum van de discussie heeft geplaatst. Piketty heeft het heel indirect ook over macht en zou, als hij 60 jaar geleden participant in de Cambridge-Cambridge discussie was geweest, wellicht aan de Engelse kant gestaan hebben met zijn (genuanceerd) uitgesproken voorkeur voor de Marxistische benadering van het kapitaaldebat. In tegenstelling tot Marx baseert Piketty zich echter op bestaande statistieken van de kapitaalgoederenvoorraad, waar de gebruikelijke conventies om kapitaal te meten zijn toegepast. Hij gaat dus uit van de aanname dat de stroom van diensten die een kapitaalgoed nu en in de toekomst zal leveren op een objectieve wijze bepaald kan worden. De cirkelredenering die de Engelse Cambridge-economen signaleerden, is volgens Piketty niet relevant en de neo-klassieke productiefunctie in combinatie met de veronderstelde beloning van productiefactoren volgens hun marginale productiviteit kan (uitzonderingen daargelaten) gehanteerd worden om voorspellingen over de inkomensverdeling tussen arbeid en kapitaal te genereren.

Daarmee past Piketty met zijn 2014-boek in de neo-klassieke traditie die meent dat, uitgaande van een neo-klassieke productiefunctie, de waarde van de geaggregeerde kapitaalgoederenvoorraad een zinvol begrip is waarmee macrodata geanalyseerd kunnen worden. De Cambridge-Cambridge discussie (waar hij nog geen drie bladzijden aan

besteedt in zijn boek) interpreteert hij als een debat over de substitueerbaarheid van arbeid en kapitaal, waarbij de Engelse kant van de discussie uitging van vaste technische coëfficiënten en de Amerikaanse kant van substitueerbaarheid. Deze discussie had volgens hem niet hoeven plaats te vinden als de strijdende partijen de juiste gegevens hadden gehad: “In my view, the virulence – and at times sterility – of the Cambridge capital controversy was due in part to the fact that participants on both sides lacked the historical data needed to clarify the terms of the debate” (Piketty 2014: 232). Als de economen in jaren 50-70 de data hadden gehad die Piketty nu wel heeft, hadden zij zich met belangrijkere zaken bezig kunnen houden. Impliciet miskent hij in dit citaat dat de Engelse kant van de discussie de data juist wantrouwt en zelfs verwerpt, wegens de boven aangehaalde cirkelredenering. Als een pragmatisch neo-klassiek econoom acht Piketty het gebruik van een macro-economische productiefunctie met de geaggregeerde (en meetbare) kapitaalgoederenvoorraad gerechtvaardigd. Immers, vanaf de jaren 1970 is het zo dat “Solow’s so-called neoclassical growth model definitively carried the day (Piketty 2014: 231).”

Romer en Piketty lijken op één lijn te zitten wat de meetbaarheid van kapitaal betreft. Daarom is het verwonderlijk dat Romer ook Piketty van *mathiness* beschuldigt. De verklaring daarvoor kan zijn dat Piketty in zijn boek de marxistische stelling van de voortdurend toenemende concentratie van kapitaalbezit en de dalende winstvoet interpreteert in termen van de neo-klassieke groeitheorie (pp 227-230). De continue stijging van de kapitaal-inkomensratio, zoals die door Piketty voor het laatste deel van de 20e en het begin van de 21e eeuw wordt beschreven, kan verklaard worden met een simpel neo-klassiek groeimodel met lage productiviteitsgroei. Door extreem lage productiviteitsgroei en een lage groei van de bevolking, kan het kapitaalaandeel in het nationaal inkomen naar ongeëvenaard hoge waarden gaan. In het uiterste geval kan de kapitaal-inkomensratio zo hoog worden dat de opbrengstvoet op kapitaal naar nul moet gaan, hetgeen (volgens Marx) het einde van het kapitalisme zou inluiden: het kapitalistische stelsel zal imploderen, terwijl het kapitaalaandeel juist explodeert. Piketty weet zo, aan de hand van (simpele) neo-klassieke theorie iets van de door Marx gesignaleerde tegenstrijdigheden in het kapitalisme in de theorie te incorporeren. De door hem verzamelde data over een lange periode suggereren inderdaad dat er imminent in het kapitalistische stelsel sprake is van een toenemende ongelijkheid tussen kapitaal (rijk) en arbeid (arm), zoals Marx die voorspelde. Het apocalyptische einde van het kapitalistische systeem doet zich alleen maar niet voor omdat alle economieën productiviteitsgroei kennen. Bovendien wordt toenemende ongelijkheid tussen kapitaal en arbeid tegengewerkt door crises als de Grote Depressie van de jaren 1920/30 en door catastrofale wereldoorlogen die een groot deel van de kapitaalgoederenvoorraad beschadigen of vernietigen.

Volgens de endogene groeitheorie, waar Romer de vader van is, zou een dergelijke extreme ontwikkeling van sterk divergerende beloningsverhoudingen tussen kapitaal en arbeid zich niet kunnen voordoen omdat het kapitaalaandeel in het nationaal inkomen constant is en er geen sprake is van afnemende marginale kapitaalproductiviteit. Het is

natuurlijk speculatief, maar het lijkt er op dat Romer de hypothetische toekomstige ontwikkeling van de vermogensongelijkheid die Piketty beschrijft binnen het huidige economische systeem verwerpt. Wellicht daarom plakt Romer het label *mathiness* op werk van Piketty.

6 *Mathiness en mathematisering in de economie revisited*

Rosenberg (1992) nam een verdere mathematisering van de economische wetenschap waar. Dat voldeed aan het aan Lakatos ontleende criterium dat een nieuwe theorie ‘*theoretically progressive*’ moet zijn, namelijk meer moet kunnen verklaren dan oudere theorieën. Deze ontwikkeling was echter niet ‘*empirically progressive*’, namelijk het leidde niet tot de ontdekking van nieuwe feiten. Volgens Romer is mathematisering van de economie echter geen probleem, zolang de theorie maar berust op waarneembare verschijnselen (en past binnen de ‘consensus’). Over empirische verificatie of falsificatie van de theorie zwijgt Romer, terwijl Rosenberg (1992) dat (met een nu gedateerd beeld) beschrijft als een cheque die op de post is gedaan, maar nooit aankomt.

Volgens Rosenberg (1992) blijft de economische wetenschap steken in toegepaste wiskunde door de gebrekkige empirische steun voor de theorie. Rosenberg beschouwde dit als een diskwalificatie voor de economie als wetenschap, maar herzag later dit oordeel wanneer hij de economie als een historische wetenschap typeert (Rosenberg, 2009). Deze interpretatie ten aanzien van het ‘wetenschappelijke van de economie’ past bij de in hoofdstuk 3 aangehaalde opvatting van Keynes. Keynes zag de economische wetenschap als een verzameling modellen waaruit voor iedere plaats en tijd het meest relevante model gekozen moet worden. De (al dan niet wiskundige) generalisaties van de economische wetenschap zijn daarom ook geen ‘wetmatigheden’, maar kunnen lokale economische verschijnselen verklaren die maar gedurende een beperkte tijd geldig kunnen zijn.

Dat er uit het verloop van de geschiedenis geen wetmatigheden kunnen worden afgeleid, is ook met kracht door Popper in *The poverty of historicism*⁴ betoogd. Het verloop van de geschiedenis is eenmalig en kan niet door experimenten worden herhaald. De geschiedenis herhaalt zich niet en het is dus ook onmogelijk om het verloop van gebeurtenissen te kunnen voorspellen. Als de economie een historische wetenschap is, zijn economische modellen door hun ‘historische karakter’ ongeschikt om te dienen als basis voor economische voorspellingen. De economische theorie vormt dan eerder een verzameling van case studies die soms van toepassing zijn, maar soms (misschien vaker) niet.

Een recente uitwerking van dit idee van de economische wetenschap als een verzameling case studies wordt gegeven door Gilboa et al. (2014). Zij maken een onderscheid tussen rule-based redeneren, waarbij het doel is te komen tot algemene theorieën, en case-based

⁴Het is vermeldenswaard dat de ‘*poverty*’ al in 1944-1945 in drie delen in *Economica* was verschenen. Popper beschouwde de economie overigens niet als een historische wetenschap.

redeneren, waarbij men gelijke situaties uit het verleden probeert te identificeren als basis voor voorspellingen. Hun stelling is dat de economische wetenschap grotendeels gebaseerd is op een groot reservoir aan (theoretische, empirische of experimentele) cases die relevant zijn als ze gekoppeld kunnen worden aan 'reële' cases. Nieuwe economische theorieën zijn toevoegingen aan dat reservoir. Waargenomen cases kunnen op zichzelf die theorieën niet tegenspreken, want er kunnen andere cases zijn waarmee de theorie wel in overeenstemming is. Omdat in deze opvatting theorieën geen algemene geldigheid bezitten, kunnen ze ook nooit onwaar zijn.

Economie gebruikt meer dan de andere sociale wetenschappen wiskunde bij het formuleren van theoretische resultaten. Economische theorie kan men bijvoorbeeld nauwelijks begrijpen zonder in staat te zijn wiskundig geformuleerde functies te differentiëren. Dit is een ondeugd, volgens Rosenberg (1992) omdat deze mathematisering niet gepaard is gegaan aan een verbetering van de verificatie van de theorie. In de opvatting van Gilboa et al. (2014) is de 'mathematisering' van de economie echter een deugd, omdat de generalisatie van bestaande theoretische resultaten die via wiskundige analyse worden verkregen, de potentiële relevantie van de theorie vergroot. Economische theorie moet uiteraard logisch consistent zijn en moet bij voorkeur zo intuïtief mogelijk zijn, zodat het duidelijk is dat de theorie gelijkenis vertoont met bestaande situaties (Gilboa et al., p. F521). Maar, volgens de case-based opvatting hoeft een theorie niet absoluut (empirisch) waar te zijn. Wat men van een theorie mag vragen is dat die in feitelijke situaties van toepassing kan zijn. Er volgt dan direct dat als de theorie via wiskundige methoden algemener wordt gemaakt, dit de kans vergroot dat de theorie gelinkt kan worden aan feitelijke situaties.

Zo komen we via Keynes, Popper, Rosenberg (2009) en Gilboa et al. tot de conclusie dat 'de' waarheid in de economische theorie niet bestaat. Resultaten in de economische wetenschap kunnen (en mogen!) elkaar volop tegenspreken, omdat ook cases in de werkelijkheid elkaar kunnen tegenspreken. Deze wetenschapsfilosofische opvatting over de economische wetenschap verhoudt zich slecht met de opvatting van Romer dat wetenschap een proces is dat naar een breed gedragen consensus leidt "around theoretical and empirical statements that are true". Bijgevolg past ook het door Romer geïntroduceerde begrip *mathiness* niet in deze wetenschapsopvatting. Iedere theoretisch (empirisch of experimenteel) resultaat dat aan bepaalde minimumvoorwaarden voldoet, zoals logische consistentie en intuïtieve zeggingskracht, heeft recht op een plaats in het reservoir aan cases dat economische wetenschap heet.

7 Slot: *mathiness of truthiness*

De wetenschapsfilosofie van Romer, dat de wetenschap moet streven naar consensus over de empirische en theoretische waarheid, is contraproductief. De waarheid kan niet vastgesteld worden in de economie door het historische karakter van economische verschijnselen. Elkaar tegensprekende modellen sluiten niet uit dat ze toch allemaal op verschillende plaatsen of in andere perioden relevant kunnen zijn.

De tweede eis van Romer namelijk dat variabelen in theoretische modellen waarneembaar moeten zijn, kan al evenzeer een contraproductief effect hebben. Een voorbeeld waaruit blijkt dat waarneembaarheid niet noodzakelijk is voor theorievorming betreft het atoom. Over het bestaan van atomen werd al voor het begin van de jaartelling gefilosofeerd. Pas tegen het einde van de 19e eeuw waren natuurkundigen in staat indirect via experimenten het bestaan van atomen te bewijzen. Dat bracht de ontwikkeling van de atoomtheorie in een stroomversnelling. Waarneembaarheid van variabelen is dus wel een voldoende, maar niet een noodzakelijke voorwaarde voor theorievorming. In de economische wetenschap geldt dat ook; de waarneembaarheid van kapitaal is daar een voorbeeld van. De kapitaalgoederenvoorraad kan gemeten worden en het resultaat daarvan vindt men in de nationale rekeningen. De manier waarop kapitaal gemeten wordt is echter gebaseerd op theoretische overwegingen die niet door iedereen worden gedeeld. Als economische theorie verder ontwikkeld had moeten worden, door uit te gaan van eenduidig waarneembare data, had dat tot stagnatie in de theorievorming kunnen leiden.

De stelling van Romer dat wiskunde in de economische wetenschap misbruikt wordt voor politieke of ideologische doeleinden is ook weinig vruchtbaar. De mathematisering van de economie kan, juist als het historische en vaak unieke karakter van economische verschijnselen wordt erkend, de economische wetenschap verder helpen. Mathematisering maakt een model algemener waardoor de kans dat het model past op een situatie uit de 'echte' economische wereld toeneemt (Gilboa et al. 2014). Door de mathematisering van de economische wetenschappen wordt de econoom geholpen om, in de woorden van Keynes, de kunst van het kiezen van modellen die relevant zijn voor de hedendaagse wereld beter te beheersen.

Het is niet onmogelijk dat in sommige economische papers wiskunde alleen maar wordt gebruikt ter wille van zichzelf en niet omdat het de band met de werkelijkheid zou verduidelijken. Evenzeer is het mogelijk dat mathematisering gebruikt wordt om, in de woorden van Romer, academische politiek te bedrijven. De eis dat proposities en de bewijsvoering intuïtief aannemelijk moeten zijn (Gilboa et al. 2014: F520-522) beperkt dit risico. Toch zal het moeilijk zijn om eenduidig aan te wijzen welk soort onderzoek alleen maar *l'art pour l'art* of ideologisch gemotiveerd werk is. Er kan zich immers altijd een situatie voordoen waar het model op van toepassing is. Beschuldigingen van *l'art pour l'art* of *mathiness* leiden bovendien al gauw tot willekeur, meer gebaseerd op de persoonlijke smaak van de criticus dan op objectief waarneembaar misbruik van wiskunde. Zo is *mathiness* volgens Romer het gebruik van wiskunde dat alleen maar in schijn tot inzicht leidt, maar in feite ideologische motieven verhuult. Het woord *mathiness* is afgeleid van het woord *truthiness* dat verwijst naar het verkondigen van 'waarheden' die niet gebaseerd zijn op objectief waarneembare feiten. Uit de manier waarop Romer het label *mathiness* op academische papers plakt, lijkt hij zelf te lijden aan *truthiness*. Het voorbeeld van de *mathiness*-verklaring van werk van Piketty suggereert dat hij zelf ideologisch gemotiveerd is.

Auteur

Harrie Verbon (e-mail: H.A.A.Verbon@uvt.nl) is hoogleraar aan de Tilburg School of Economics and Management Department of Economics

Literatuur.

- Backhouse, R.E., 2014, MIT and the other Cambridge, *History of Political Economy*, vol. 46: 252-271.
- Barro, R.J., 2013, Education and economic growth, *Annals of Economics and Finance*, vol. 14(2): 301-328.
- Blaug, M., 1992, *The Methodology of Economics. Or, How Economists Explain*, Cambridge University Press.
- Blaug, M., 2009, The trade-off between rigor and relevance: Sraffian economics as a case in point, *History of Political Economy*, vol. 41(2009): 219-41.
- Boumans, M. en J.B. Davis, 2016, *Economic methodology (Understanding economics as a science)*, Palgrave, 2nd ed.
- Cohen, A.J. en G.C. Harcourt, 2003, Whatever happened to the Cambridge capital theory controversies? *Journal of Economic Perspectives*, vol. 17(1): 199-214.
- Gärtner, M., 2016, *Macroeconomics*, Pearson.
- Johnson, E. en D. Moggridge (red.), 1978, *The Collected Writings of John Maynard Keynes, The General Theory and After*, vol XIV.
- Gilboa, I., Postlewaite, A, L. Samuelson, en D. Schmeidler, 2014, Economic models as analogies, *Economic Journal*, vol. 124: F513-F535.
- Kahnemann, D., 2011, *Thinking, fast and slow*, London: Penguin Books.
- Mata, 2004, Constructing identity: the post Keynesians and the capital controversies, *Journal of the History of Economic Thought*, vol. 26(1): 241-259.
- Piketty, T., 2014, *Capital in the 21st century*. Cambridge: The Belknap press of Harvard University Press.
- Popper, K., 1957, *The Poverty of historicism*, Routledge.
- Romer, P. M., 2015, Mathiness in the theory of economic growth, *American Economic Review, Papers & Proceedings*, vol.105(5): 89-93.
- Rosenberg, A., 1992, *Mathematical politics or science of diminishing returns?*, Chicago University Press.
- Rosenberg, A., 2009, *If economics is a science, what kind of a science is it?* The Oxford handbook of philosophy of economics.
- Smith, N., 2015, *How mathiness made me jaded about economics*. Artikel op www.bloombergvew.com.
- Solow, R.M., 1956, A contribution to the theory of economic growth, *Quarterly Journal of Economics*, vol. 70(1): 65-94.
- Solow, R.M., 1994, Perspectives on growth theory, *Journal of Economic Perspectives*, vol. 8: 45-54.
- Taleb, N.N., 2007 *The black swan: The impact of the highly improbable*. US: Random House.
- Verbon, H., 2013, De late erkenning van de zwarte zwaan door Paul Samuelson, *ESB*, vol. 98: 462-464.
- Williamson, S., 2015, *Don't get mathy with me, or I'll give you a good shunning*, op het blog New monetarist economics (21 mei 2015).