

Locatie en loon in de VS

In de eerste maanden van 2003 zat ik met de handen in het haar. Ik was bezig met het laatste hoofdstuk van mijn proefschrift over economische geografie, zo op het oog een aardig boekje over de rol van ruimte en afstand in economische relaties. Voor mij, afgestudeerd econometrist, moest het een makkelijk hoofdstukje worden: een empirische analyse van de handel tussen Amerikaanse staten gezien in het licht van de eerder ontwikkelde theorie. Er was echter een klein probleem: dit was inmiddels het derde jaar dat ik probeerde dit hoofdstuk af te ronden. Het hele proefschrift had eigenlijk al in 2000 klaar moeten zijn. Na ettelijke pogingen had ik nog steeds geen flauw idee waar te beginnen.

Thijs Knaap

is universitair docent macroeconomie aan de Utrecht School of Economics. Naast economische geografie doet hij onderzoek naar de effecten van vergrijzing binnen het onderzoeksnetwerk Netspar. De website eco.nomie.nl is mede van zijn hand.

Aan data geen gebrek. Tijdens een verblijf op Brown University had ik de hand weten te leggen op een uitgebreid overzicht van de ruimtelijke organisatie van de Amerikaanse economie: de productie van de verschillende staten, de loop van de interne handel, de lokale loonniveaus, onderlinge afstanden. Volgens de Nieuwe Economische Geografie (NEG, de oude is van voor mijn tijd) zijn al deze zaken met elkaar in verband te brengen.

NEG is een prachtig verhaal dat ruimtelijke organisatie verklaart in algemeen evenwicht: bedrijven zitten daar waar ze de meeste winst behalen, werknemers vinden dat ze in de beste regio wonen, kortom iedereen is tevreden. Voor je gevoel betekent dat waarschijnlijk dat NEG-modellen met het soort lauwe oplossingen komen waarin iedere regio even groot is en de bedrijven keurig over het land verspreid zitten. Maar niets is minder waar: in de NEG-wereld is het volstrekt logisch dat miljoenen mensen en bedrijven met z'n allen in een grote stad gaan zitten, inclusief hoge huren en vastzittend verkeer, terwijl een eindje verderop geen klap te doen is. De theorie in verband brengen met de Amerikaanse economie, waarin zoals bekend mega-steden worden afgewisseld met uitgestrekte prairie, moest dus een fluitje van een cent zijn.

Alleen was het niet zo. Laat ik proberen uit te leggen hoe het zou moeten werken. De ruimtelijke variatie komt tot stand door de aanwezigheid van transportkosten: wat van ver komt is duurder. Dat pleit ervoor om productie zo dicht mogelijk bij de klant te laten plaatsvinden. Echter, er werkt ook een kracht de andere

kant op: schaalvoordelen. In dit geval lopen die via intermediaire producten, de goederen en diensten die een bedrijf nodig heeft om zelf te kunnen produceren. Een clustering van bedrijven op één plek zorgt voor lage kosten voor die intermediaire producten, en dus voor lage prijzen. Het evenwicht tussen transportkosten en schaalvoordelen levert een ruimtelijk patroon; mijn taak was om de parameterisatie van het model te vinden die een patroon oplevert dat zo dicht mogelijk op de werkelijkheid ligt. Dat is een bekend karwei waarvoor je een slimme optimalisatie kunt gebruiken, of (in deze tijd van brute rekenkracht) alle mogelijke combinaties van parameters af kunt lopen om uiteindelijk diegene te kiezen met de beste *fit*.

Maar er is iets bijzonders aan de hand met NEG modellen, of beter gezegd, twee bijzondere dingen. Ten eerste kunnen de modellen niet analytisch opgelost worden. Wie de oplossing wil weten moet aan de numerieke simulatie. Ten tweede voorspellen de modellen dan wel dat bijzondere patronen (agglomeraties) voor kunnen komen, maar ze zeggen niets over het *waar* van deze agglomeraties. Vaak heeft een model meerdere oplossingen en is op voorhand niet duidelijk welke van die oplossingen gevonden wordt. Dit verschijnsel staat bekend als hysteresis of pad-afhankelijkheid: het toeval bepaalt waar de grote stad terecht komt. Er blijkt een interessante empirische wetmatigheid te bestaan: hysteresis + numerieke simulaties = hoofdpijn. Want zo valt er natuurlijk weinig te optimaliseren: iedere verzameling parameters leidt tot een onbekend aantal oplossingen, waarvan er maar één geobserveerd wordt. En vaak komt er, bij herhaalde simulatie, steeds iets anders uit.

Nee, het moest allemaal op een andere manier, dat was duidelijk. Ik moest af van het idee dat een consistente modeloplossing dicht bij de data gebracht moest worden, en me concentreren op de omgekeerde vraag: bestaat er een om te

controleren in hoeverre de data consistent zijn met het model? Een minder ambitieus, maar meer haalbaar project. Ik besloot de te stoppen met simuleren en nog eens in de literatuur te duiken, om te zien of er methodes bestonden om kenmerken van een NEG-oplossing terug te vinden in de werkelijkheid. Die beslissing werd mijn redding.

Letterlijk, trouwens. Ik vond een paper geschreven door de Engelse econoom Stephen Redding, samen met een van de grondleggers van de NEG-theorie, Anthony Venables. Zij waren erin geslaagd om die theorie te testen op een dataset over wereldhandel en de lonen in verschillende landen. Dat was het natuurlijk: een belangrijk kenmerk uit het model is het feit dat elke producent overal beconcurrereerd wordt vanuit alle andere regio's, en dus niet te veel ruimte heeft om hoge prijzen te vragen. Tegelijkertijd zijn er transportkosten, zowel voor de import van intermediaire goederen als voor de export van finale goederen. Als de verkoopprijs dus min of meer vastligt, en voor gebieden buitenaf de inputs duur zijn en het eindtransport ook nog eens geld kost, dan schiet er weinig over voor de lokale factoren. Oftewel, de lonen zijn lager in gebieden die ver buiten de economische kern liggen. Redding en Venables tonen aan dat dit inderdaad het geval is op wereldschaal, maar de methode die zij daarvoor gebruiken kon natuurlijk net zo goed worden toegepast op mijn, Amerikaanse, dataset. Het was alsof ik een schriftelijke route naar mijn promotie opgestuurd had gekregen.

Een extra voordeel van het paper van de Engelsen was de grote schoonheid van de methode. Het vage verhaal over lonen en prijzen hierboven is natuurlijk precies te maken; de relatie tussen lonen, productie en transportkosten (de vraagfunctie) is

$$\sum_{r=1}^N E_r G_r^{\sigma-1} T_{sr}^{1-\sigma} = \bar{Y} \left[\frac{\sigma}{\sigma-1} W_s^\alpha G_s^{1-\alpha} \right]^\sigma$$

Er zijn N staten, elk met vraag E en prijsniveau G . De transport-*markup* tussen twee staten is T (een typische waarde: 1.2). De prijs p wordt vastgesteld als een gewogen geometrisch gemiddelde van het lokale loon w en het lokale prijsniveau (van de inputs) G . Een dergelijke vraagfunctie bestaat voor iedere staat s , en na even turen is te zien wat er met de lonen gebeurt. In de som links variëren de transportkosten T ; voor staten die ver uit de kern liggen zijn die kosten hoog, en dus is de linkerkant van de vergelijking voor die staten klein (σ is groter dan 1). Voor diezelfde staten is de term G aan de rechterkant hoog: ook vanwege de transportkosten is het duur om intermediaire goederen te kopen. Van twee kanten wordt het loonniveau w afgeknepen.

Allemaal goed en wel, maar hoe kom je aan praktische waarden voor de termen ter linker- en rechterzijde? Wat ik hierboven beschrijf is in

feite niet meer dan de potentiaal-theorie, die zegt dat gebieden met veel activiteit hoge lonen hebben. En daar schuilt de schoonheid van het artikel: Redding en Venables slagen erin om uitdrukkingen te vinden die, als het model klopt, exact overeen komen met de twee termen, en wel op een onverwachte plek: in de data over de handel tussen staten.

Laten we er nog één vergelijking tegenaan gooien. Als alles werkt zoals de theorie voorziet, is de handelsstroom tussen twee staten, X , gelijk aan

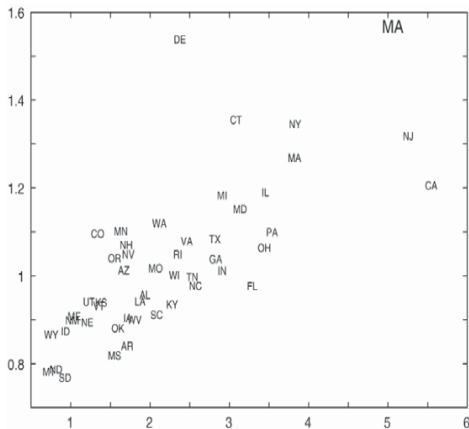
$$X_{rs} = n_s p_s^{1-\sigma} E_r G_r^{\sigma-1}$$

Hier heeft de versturende staat s een aantal van n bedrijven en prijs p . De transportkosten staan in het midden, en de bestedingen en prijsindex van de ontvangende staat aan de rechterkant. Het is een intuïtieve vergelijking: de waarde van de handel is proportioneel met de "grootte" van de staten en neemt af met de transportkosten. Eigenlijk is dit gewoon een opgevoerde versie van de aloude zwaartekrachtvergelijking, nog uitgevonden door Jan Tinbergen. Een prettige vergelijking, omdat hij bijzonder goed te schatten valt.

Wat is nu de truuk? Van de vijf termen aan de rechterkant horen er twee alleen bij de versturende staat en twee alleen bij de ontvangende staat. Je kunt die twee stellen opvatten als *fixed effect* van een staat en de vergelijking schatten als een panel. Dat levert je een schatting van het effect van transportkosten en een empirische waarde voor de twee linker- en de twee rechtertermen. Het mooie is dat de twee rechtertermen exact overeenkomen met de elementen van de som in de eerste formule. Door dus de handel te schatten kom je aan getallen die precies passen in de uitdrukking voor het loonniveau. Extra prettig is dat de twee linkertermen om te vormen zijn tot een uitdrukking voor G , zodat de hele eerste vergelijking ingevuld is.

Het is allemaal niet onlogisch: de theorie zegt dat lonen hoog zijn, daar waar de meeste activiteit geconcentreerd is. Bovendien vindt de meeste handel plaats tussen staten met een hoge concentratie van activiteit. Door naar de handel te kijken vind je de concentraties, en met die kennis kijk je naar het loonniveau.

Econometrisch is er natuurlijk nog genoeg om je zorgen over te maken. De formule met het loon maakt gebruik van de schattingsresultaten uit de handelsrelatie. Verstoringen in die data werken op de meest vreselijke manieren door in de tweede schatting, waardoor de standaardfouten meteen het raam uit kunnen. Maar ik had natuurlijk niet voor niets een snelle computer en een boek over *bootstrap* methoden op mijn bureau. Na een dag of wat klooiën had ik de eerste formule geschat.



Het zag er fraai uit: de gegevens over concentratie, omgevormd tot *market access* (MA) staan op de horizontale as en de log-lonen op verticale as. Lezers met goede ogen zien in elk datapunt de afkorting van een staat. Wie de ogen samenknijpt ziet een fraaie correlatie.

Graag was ik op dit punt gestopt. Tijdens een lange periode vol onzekerheid had ik dapper doorgerekend en de weerbarstige data eindelijk in een nette puntenwolk gedwongen. Het was tijd om de ramen open te zetten en het rokkostuum te laten stomen, de Amerikaanse staten achter me te laten. In een film begint op dat moment de aftiteling te rollen.

Helaas was dit zo'n film waarin op het laatste moment nog iemand binnenstapt met de woorden "*not so fast.*" De correlatie uit de figuur hierboven lijkt erop te duiden dat de NEG-theorie de lonen verklaart, maar er kan ook iets anders aan de hand zijn. Stel dat de productiviteit in bepaalde staten extra hoog is, bijvoorbeeld door het vlakke land, het mooie weer, de gunstig gelegen kust. Dan gebeuren er twee dingen: de lonen zijn extra hoog, als afspiegeling van de productiviteit, en de omvang van de economie in die staten is extra groot (productiviteit, opnieuw, en de aanzuigende werking van de hoge lonen). Zo'n toevallige schok werkt in op exogenen en de te verklaren variabele, en kan een verklaring zijn voor de mooie correlatie hierboven.

Dat, in het kort, was het commentaar dat ik van mijn Amerikaanse collega uit Brown kreeg. Om het probleem, waarvan we niet zeker wisten dat het bestond, te ondervangen kun je drie dingen doen. Ten eerste kun je de eigen staat uit de *market access* variabele halen. Je verklaart dan de lonen door de concentratie in de omliggende staten. Dat blijkt slecht te werken, en niet geheel onverwacht: de hoge lonen in Californië worden nou eenmaal niet veroorzaakt door factoren buiten die staat. De tweede methode, instrumentele variabelen, laat weinig heel van de oorspronkelijke correlatie. Dat kan komen door de extra onnauwkeurigheid van de methode, of

doordat er weinig correlatie te vinden valt. De derde truuk is het vangen van productiviteitsschokken in extra exogenen, zoals indicatoren van het klimaat. Het toevoegen daarvan is opnieuw dodelijk voor mijn resultaat: de mooie puntenwolk wordt uit elkaar geblazen.

Nou ja, het kan niet elke dag feest zijn voor de economisch geograaf. Het hoofdstuk gaf dan wel geen keihard bewijs voor de theorie waar ik jaren van mijn leven aan gegeven had, maar het lag er. En de conclusie dat er nog wel wat extra onderzoek nodig is, is met het oog op de toekomst helemaal niet zo slecht.

Referenties

Knaap, T. (2006). *Trade, location, and wages in the United States*. Regional Science and Urban Economics, forthcoming.