

Beleidsrapport STORE-22-019

Methodologie voor het bepalen van geografische groeiclusters in Vlaanderen a.d.h.v. een complexiteitsmaatstaf

Dieter Van Esbroeck ^{a,b,1}, Jakob Vanschoonbeek ^{a,b,2} en Astrid Volckaert ^{a,b,3}

^a *Steunpunt Economie en Ondernemen*

^b *Vlaams Instituut voor Economie en Samenleving (VIVES)*

Faculteit Economie en Bedrijfswetenschappen, KU Leuven

¹dieter.vanesbroeck@kuleuven.be

²jakob.vanschoonbeek@kuleuven.be

³astrid.volckaert@kuleuven.be

December 2022

**STEUNPUNT
ECONOMIE &
ONDERNEMEN**



Vlaanderen
is economie, wetenschap
& innovatie

© *STORE* en *KU Leuven* (2022). Dit rapport is gebaseerd op de Master Thesis van Siebe Schillebeeckx ([2022](#)) die werd uitgevoerd onder de supervisie van Jo Reynaerts, Jakob Vanschoonbeek en Astrid Volckaert. Wij wensen Aaron Putseys te bedanken voor opmerkingen en suggesties. De resultaten in dit rapport geven de mening van de auteurs weer en niet die van de Vlaamse overheid. De Vlaamse Gemeenschap/het Vlaams Gewest is niet aansprakelijk voor het gebruik dat kan worden gemaakt van de in deze mededeling of bekendmaking opgenomen gegevens.

Inhoudstafel

Samenvatting	3
1. Inleiding	4
2. Complexiteitstheorie	5
2.1 Wat is de complexiteitstheorie	5
2.2 Literatuur overzicht	6
2.3 Enkele bestaande voorbeelden	9
2.4 Economische complexiteitsindex (ECI)	10
2.5 Kennis complexiteitsindex (KCI)	11
3. Methodologie	12
3.1 De inputmatrix	12
3.2 De reflectiemethode	13
3.3 ECI op basis van lineaire algebra	15
3.4 Grafische weergave voor de Vlaamse economie	16
3.5 Een fictief voorbeeld	19
4. De geografie van bestaande clusterinitiatieven	26
4.1 Speerpuntclusters en cluster lidmaatschap	26
4.2 Maatschappelijke zetel versus vestiging	26
4.3 Complementariteit met geografische clusters	27
4.4 Mogelijke bijkomende clusterinitiatieven	28
5. Verdere toepassingen	28
Referenties	29
Appendix	31

Samenvatting

Het ontstaan en de groei van geografische clusters en hun impact op de regio is reeds geruime tijd gedocumenteerd in de wetenschappelijke literatuur. Meer recent wordt ook onderzoek gedaan naar de impact van ‘clusterinitiatieven’, waarbij het beleid de vorming van clusters ondersteunt om zo kennisoverdracht en innovatie te genereren, die niet organisch zou ontstaan.

Dit rapport beschrijft een methodologie om (i) de geografische groeiclusters in Vlaanderen te identificeren, (ii) in kaart te brengen waar deze groeiclusters geografisch overlappen met de bestaande clusterinitiatieven en (iii) na te gaan in welk mate geografische clusters en bestaande clusterinitiatieven complementair zijn.

Daartoe wordt gebruik gemaakt van de zogenaamde Economische Complexiteitstheorie, om deze complexe netwerken in kaart te brengen aan de hand van twee verschillende indexen: de Economische Complexiteits Index (ECI), berekend aan de hand van tewerkstellingsdata op het niveau van de individuele vestigingen van bedrijven; en de Kennis Complexiteits Index (KCI) op basis van patent gegevens (op bedrijfsniveau).

De voorgestelde methodologie laat ook toe om de evolutie in complexiteit doorheen de tijd te analyseren. Ten slotte kan deze methodologie ook benut worden om te bekijken welke bijkomende clusterinitiatieven in de toekomst mogelijk opportuun zijn; waar en op welke manier de beschikbaarheid van bedrijfsterreinen de uitbouw van (nieuwe) geografische groeiclusters kan ondersteunen; en wat de mogelijke verschillen zijn tussen KMO's en grote ondernemingen.

1. Inleiding

Dit rapport beschrijft de methodologie voor het identificeren van geografische groeiclusters in Vlaanderen aan de hand van een complexiteitsmaatstaf. Dit methodologische rapport dient als basis voor de empirische studie die in 2023 geografische groeiclusters in kaart zal brengen in Vlaanderen, in opdracht van EWI.

In wat volgt gaat hoofdstuk 2 dieper in op de economische complexiteitstheorie. Sectie 2.1 belicht de ontstaansgeschiedenis en interpretatie van de complexiteitindex. Sectie 2.2 biedt een uitgebreid literatuuroverzicht en overloopt de link tussen economische groei en complexiteit, evenals het pionierswerk in de afgelopen decennia om tot een dergelijke index te komen. Sectie 2.3. overloopt enkele voorbeelden uit bestaande studies voor het buitenland. De economische complexiteitsindex (ECI) en de kennis complexiteitsindex (KCI) worden vervolgens besproken in secties 2.4 en 2.5.

Hoofdstuk 3 geeft een gedetailleerd overzicht van de methodologie om de index te berekenen. Sectie 3.1-3.3 bespreken de verschillende stappen en sectie 3.4 toont een grafische weergave. Sectie 3.5 werkt een fictief motiverend voorbeeld uit om de centrale concepten van de methode toe te lichten.

Hoofdstuk 4 legt de link met de speerpuntclusters: het clusterlidmaatschap wordt besproken, alsook het belang van vestigingsdata en de complementariteit tussen economisch complexe regio's en de aanwezigheid van clusterleden. Hoofdstuk 5 overloopt een aantal mogelijke toepassingen van de methodologie voor toekomstige werkzaamheden binnen ECOOM-STORE.

2. Complexiteitstheorie

2.1 Wat is de complexiteitstheorie

Volgens de complexiteitstheorie hangt de materiële welvaart van een economie af van de economische draagkracht of het 'productieve vermogen' dat zij bezit. Hoe complexer een regio, hoe meer diverse economische activiteiten ze kan ontplooiën en hoe groter de opportuniteiten om economisch te groeien. Het verband tussen economische groei en complexiteit is ook empirisch aangetoond ([Hidalgo, 2021](#)). Het is daarom met name voor beleidsmakers interessant om een inzicht te krijgen in de complexiteit van een bepaalde regio en hoe deze zich tot andere regio's verhoudt.

Het meten van de complexiteit van een regio is niet evident aangezien dit van heel veel factoren kan afhangen. De complexiteitstheorie, voor het eerst voorgesteld door Hidalgo en Hausmann ([2009](#)), stelt dat het mogelijk is om een index van complexiteit op te stellen aan de hand van een beperkt aantal variabelen die gerelateerd zijn aan economische activiteiten. Enkele voorbeelden zijn de export van bepaalde producten, het aantal werknemers of het aantal patenten in bepaalde sectoren in bepaalde regio's.

De kracht van de complexiteitstheorie schuilt erin dat we op basis van eenvoudige variabelen een rangschikking van willekeurige regio's naar hun productieve vermogen kunnen opstellen. De complexiteit wordt berekend aan de hand van twee elementen: de diversiteit en de ubiquiteit. De *diversiteit* kijkt naar hoe divers een bepaalde regio is (hoeveel sectoren komen erin voor, hoeveel verschillende producten worden geëxporteerd, ...), de *ubiquiteit* gaat na hoe courant deze sectoren of producten zijn over het geheel van alle regio's (in hoeveel gemeenten komen deze sectoren voor, hoeveel landen exporteren hetzelfde product, ...). De complexiteit van een regio stemt overeen met de diversiteit én de ubiquiteit van de economische activiteiten die er plaatsvinden, waarbij meer verschillende en meer zeldzame activiteiten (met lage ubiquiteit) sterker bijdragen tot de gemeten complexiteit. Daarbij toont de bestaande empirische evidentie aan dat sterk gediversifieerde economieën vaker zeldzame economische activiteiten ontplooiën, zodat de meest gediversifieerde gebieden doorgaans ook de hoogste mate van complexiteit kunnen aanwenden.

Om de complexiteitsindex concreet te berekenen wordt er op een iteratieve manier gekeken naar de diversiteit en de ubiquiteit. Dit resulteert in een kwantitatieve maatstaf die regio's rangschikt naar hun complexiteit naarmate ze meer diverse én meer zeldzame activiteiten ontplooiën. Locaties die een groot aantal economische activiteiten ontplooiën die bovendien amper voorkomen op andere locaties zullen als de meest complexe regio's worden gerangschikt, terwijl de complexiteit van locaties die maar een

beperkt aantal vaak voorkomende activiteiten ontplooiën het laagst zal worden ingeschat. Een fictief voorbeeld wordt in detail uitgewerkt in sectie 3.5.

2.2 Literatuur overzicht

In deze sectie wordt een overzicht gegeven van de literatuur die aanleiding gaf tot de huidige berekeningen van de complexiteitsindex en worden enkele toepassingen besproken.

Traditionele empirische studies over economische groei, vaak gebaseerd op neoklassieke groeitheorieën, bestaan uit het definiëren van individuele factoren die zowel de grootte en de geografische verdeling van economische groei en welvaart proberen te verklaren. Bij deze methode moeten a priori assumpties gemaakt worden over welke factoren juist bepalend zijn ([Hidalgo, 2021](#)), wat niet evident is. In de wetenschappelijke literatuur komen niet minder dan 140 determinanten aan bod, van handelsopenheid tot onderwijsniveau, die een significante relatie hebben met economische groei. Deze overvloed aan variabelen leidt tot een fundamenteel probleem in empirische onderzoek naar welvaart en groei door het ontbreken van een consensus over welke factoren nu mee opgenomen moeten worden ([Durlauf et al., 2005](#)). Daarbovenop stelt zich de vraag hoe een determinant juist gemeten moet worden ([Durlauf et al., 2005](#)). Durlauf en Quah ([1999](#)) wijzen er verder op dat onderzoekers vaak een aantal variabelen willekeurig selecteren totdat er een set gevonden is die voldoende verklaring lijken te bieden voor economische groei, zonder stil te staan bij mogelijke alternatieve verklaringen. Als gevolg hiervan zijn moderne modellen om economische groei te verklaren en voorspellen aan de hand van exogene factoren enorm ingewikkeld geworden ([Tacchella et al., 2012](#)). Dit laat ruimte voor andere benaderingen die economische ontwikkeling proberen te verklaren.

Een alternatieve benadering om naar economische groei te kijken focust op het begrip van 'economische draagkracht'. Deze theorie stelt dat de welvaart van een economie afhangt van het 'productieve vermogen' binnen een economie ([Hartmann et al., 2017](#)). Dit vermogen wordt bepaald door de competenties, (productieve) kennis en de gehele sociale organisatorische structuur die een economie bezit en waar ze gebruik van maakt om haar set aan producten (en diensten) te realiseren ([Cristelli et al., 2013](#); [Hausmann et al., 2014](#)). Concreet betreft het alle vaardigheden en kennis die een regio bezit evenals zijn organisaties en instellingen die het mogelijk maken voor industrieën om zich in bepaalde regio's te vestigen. Een economie zal kunnen groeien in de mate dat zij dit productief vermogen kan vergroten en toepassen ([Hausmann et al., 2014](#)). Het productief vermogen beïnvloedt de economische ontwikkeling via twee wegen: enerzijds door de mate waarin ze *regionaal aanwezig* is en anderzijds door hoe *gespecialiseerd* deze is. Sommige regio's zullen een groot economisch prestatievermogen bezitten, wat dus betekent dat ze de juiste "ingrediënten" bevatten om een grote hoeveelheid aan ondernemingen en sectoren te faciliteren. Afhankelijk van hoe gespecialiseerd het

prestatievermogen is, zal deze grote hoeveelheid aan “vermogen” een bron zijn waaruit ze competitieve voordelen kunnen halen om welvaart en groei te creëren.

De voornaamste uitdaging van deze benadering is dat de economische draagkracht van een land zelden direct te observeren is ([Fritz & Manduca, 2021](#)). Ze bestaat meestal uit immateriële factoren (bijvoorbeeld instellingen, intellectuele eigendommen, handhaving, onderwijs, investeringen, technologieën en infrastructuur) die moeilijk concreet te meten zijn ([Cristelli et al., 2013](#)). Hidalgo en Hausmann ([2009](#)) introduceerden in 2009 een reflectiemethode om het productieve vermogen van landen te meten aan de hand van een economische activiteit, zoals de export van goederen. Het achterliggende idee is om informatie uit de data van geëxporteerde producten te gebruiken om iets te leren over de productieve structuur van een land. Cruciaal is de idee dat producten verschillende hoeveelheden kennis en competenties vergen om geproduceerd te worden ([Balland et al., 2022](#)). Hidalgo noemt producten ‘gekrystalliseerde verbeelding’ om de kennis en know-how aan te geven die vervat zit in elk product ([Hidalgo, 2015](#)). Het resultaat van de reflectiemethode is een economische complexiteitsindex (ECI) voor zowel nationale economieën als de producten die zij produceren.

Twee fundamentele begrippen bepalen de complexiteit van landen: (i) de diversiteit van producten die een nationale economie produceert en (ii) de ubiquiteit van deze producten. De methode maakt gebruik van dimensie reducerende technieken en lineaire algebra om de economische draagkracht van regio's te bepalen ([Hidalgo, 2021](#)). Aan de hand van deze technieken vermijdt deze methode keuzes te moeten maken over welke variabelen bepalend zijn voor de economische ontwikkeling: de complexiteitsmaatstaf gebruikt een beperkt aantal geobserveerde variabelen om de geografische verdeling van productief vermogen binnen een economie te beschrijven zonder assumpties te maken over het belang van deze variabelen, in tegenstelling tot meer traditionele empirische studies ([Hidalgo, 2021](#); [Hidalgo & Hausmann, 2009](#)). Volgens Hidalgo en Hausmann ([2009](#)) zullen landen met een groter productief vermogen meer gespecialiseerde en complexe producten kunnen produceren. Het fenomeen van deze geografische specialisatie is in Europa zowel nationaal ([Amiti, 1999](#)) als regionaal ([Ezcurra et al., 2006](#)) in kaart gebracht.

In de traditionele visie op handel specialiseren de rijkste landen zich in een beperkt aantal specifieke geavanceerde producten. De economische complexiteitstheorie vertrekt daarentegen van de idee dat een land alle producten die ze kan produceren, gegeven haar technologisch niveau, ook zal produceren, waarbij landen met een grotere complexiteit een grotere diversiteit hebben aan activiteiten. Dit leidt tot een hoge diversiteit in de export van welvarende landen ([Tacchella et al., 2012](#); [Cristelli et al., 2013](#); [Petràlia et al., 2017](#)). Een land met een hoge complexiteit zal namelijk in staat zijn om zich meer te

specialiseren, maar deze specialisatie zal niet beperkt worden tot een aantal gesofisticeerde producten. Dit fenomeen wordt ook opgemerkt op regionaal niveau ([Sbardella et al., 2017](#)).

Uit de literatuur blijkt dat de economische complexiteitsmaatstaven een goede voorspeller zijn voor economische groei op internationaal niveau ([Hausmann et al., 2014](#); [Hidalgo & Hausmann, 2009](#); [Tacchella et al., 2012](#); [Ourens, 2013](#)). Een natuurlijke extensie is om ze toe te passen op regionaal niveau. Deze extensie roept de vraag op of de methode om economische complexiteit te berekenen voor landen zomaar toegepast kan worden op regio's. Om dit na te gaan evalueren we of de assumpties die gemaakt worden om landen te vergelijken (het internationale niveau) ook kunnen aangenomen worden op regionaal niveau. Om economische complexiteit tussen landen te vergelijken moet de veronderstelling gemaakt worden dat het productief vermogen dat nodig is om een gegeven product te produceren gelijkaardig is tussen verschillende landen ([Fritz & Manduca, 2021](#)). Daarenboven dienen de technologieën die gebruikt worden om *eenzelfde* product te vervaardigen gemiddeld minder te verschillen dan de technologieën die gebruikt worden om twee *verschillende* producten te vervaardigen in dezelfde regio ([Hartmann et al., 2017](#)). Deze veronderstellingen zijn op het regionale niveau meer geloofwaardig aangezien de variatie in economische, sociale en politieke structuren minder groot zal zijn dan op internationaal niveau. De variatie in de economische prestaties tussen regio's blijkt ook vergelijkbaar te zijn met de variatie in economische prestaties tussen landen ([Fritz & Manduca, 2021](#)).

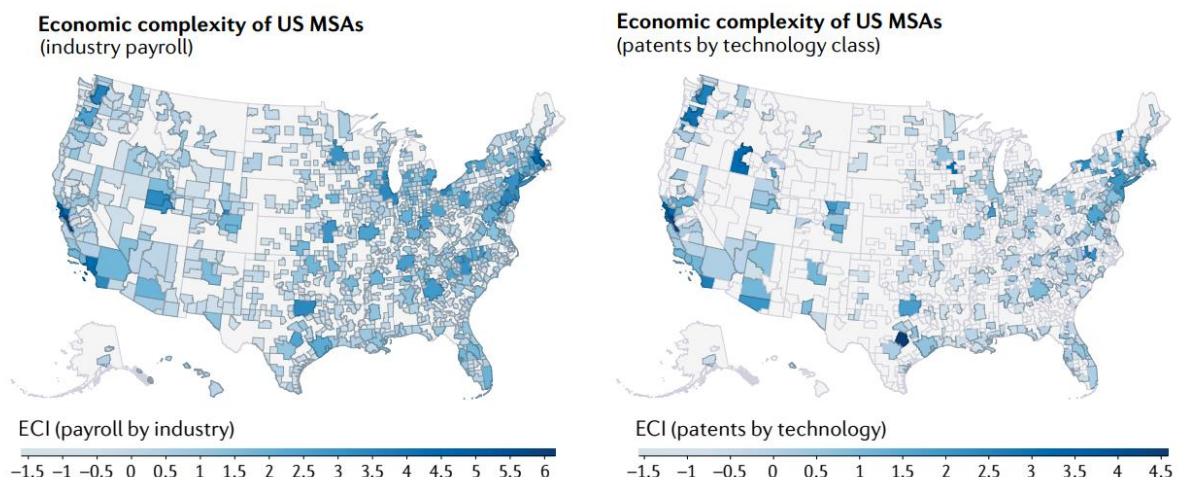
Dankzij de flexibiliteit van de economische complexiteitsmethode zijn er verscheidene studies gepubliceerd die niet de export als economische activiteit gebruiken, maar sectorale tewerkstellingsdata. Exportgegevens zijn onvermijdelijk gelimiteerd tot verhandelbare goederen, maar tewerkstellingsgegevens bevatten informatie over meerdere sectoren, inclusief dienstensectoren en niet-verhandelbare goederen. De inclusie van deze diensten en niet-verhandelbare goederen is van belang omdat ze een grote rol spelen in de economie van ontwikkelde landen ([Buera & Kaboski, 2012](#)). Dienstensectoren spelen ook vaak een ondersteunende rol in de productie, waardoor gegevens over tewerkstelling meer informatie bevatten over het economisch draagvlak dan enkel exportdata.

Chávez et al. ([2017](#)) stellen aan de hand van tewerkstellingsdata een economische complexiteitsindex op voor de staten van Mexico voor de periode van 1998 tot 2013 en vinden een positieve correlatie tussen de mate van economische complexiteit van een staat en zijn welvaart. Ze vinden tevens dat een toename in de complexiteitsindex met één standaarddeviatie gepaard gaat met een toename in de jaarlijkse economische groei van 0.4%. Dit is in lijn met de nationale bevindingen over de relatie tussen groei en economische complexiteit ([Hausmann et al., 2014](#); [Hidalgo & Hausmann, 2009](#); [Tacchella et al., 2012](#); [Ourens, 2013](#)). Fritz en Manduca ([2021](#)) benutten tewerkstellingsgegevens uit grootstedelijke gebieden in de Verenigde Staten om een complexiteitsindex op te stellen voor de periode 1998-2015.

In een cross-sectioneel onderzoek vinden ze dat steden met een hogere economische complexiteit een significant hoger per capita inkomen hebben. Deze relatie werd sterker doorheen de jaren en kende een piek in 2007, waarna de associatie tussen complexiteit en inkomen afnam. Dankzij hun vermogen om economische structuren samen te vatten, zijn de complexiteitsmaatstaven ook mee opgenomen in beleidsinitiatieven zoals de Europese slimme specialisatiestrategie ([Balland et al., 2019](#)).

2.3 Enkele bestaande voorbeelden

De eerste toepassingen van de complexiteitstheorie maakten gebruik van exportgegevens om de complexiteit van landen met elkaar te vergelijken. Op de website van de [Atlas of Economic Complexity](#)¹ staan verschillende complexiteitsrankings, landenprofielen en economische groeiprognoses, die op basis van de complexiteitstheorie (o.b.v. exportdata) werden opgesteld. Landen als Japan, Zwitserland en Duitsland zijn de meest complexe landen. België stond in 2020 op de 22^{ste} plaats op een totaal van 133 landen.



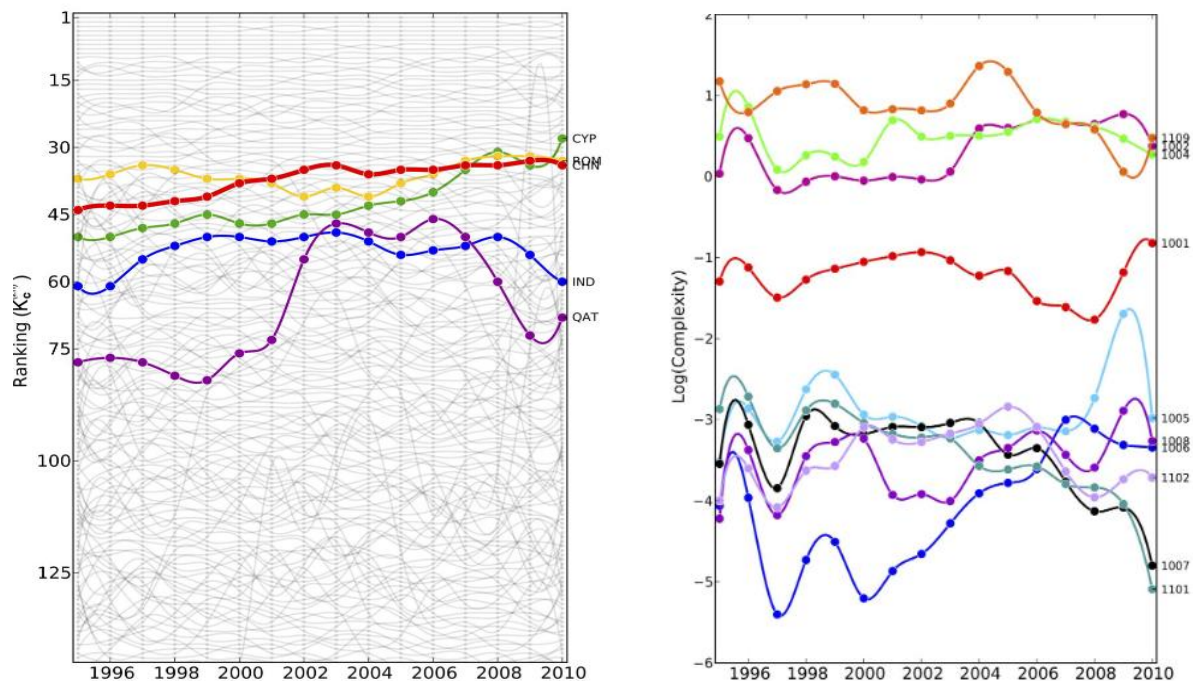
Figuur 1: Kaart met de complexiteitsindex in de VS o.b.v. tewerkstelling (links) en patenten (rechts).

Bron: Hidalgo ([2021](#))

De complexiteitsindex kan ook berekend worden o.b.v. tewerkstellingsdata of patenten. Figuur 1 geeft bijvoorbeeld een grafische voorstelling van de verschillen in complexiteit voor de verschillende regio's in de Verenigde Staten (MSA staat voor Metropolitan Statistical Area) ([Hidalgo, 2021](#)). De analyses werden gemaakt o.b.v. tewerkstelling (links) en patenten (rechts). In beide gevallen bevinden de meest complexe regio's zich aan de oost- en westkust.

¹ De Atlas of economic complexity is een initiatief van de Harvard Growth Lab, zie <https://atlas.cid.harvard.edu>.

Figuur 2 (links) toont op basis van Cristelli et al. (2013) dat deze methode niet enkel toelaat om geografische verschillen in economische complexiteit in een bepaald jaar te analyseren, maar dat het ook mogelijk is om tijdsevoluties in complexiteit te beschouwen aan de hand van een complexiteitsdashboard, in dit geval voor een grote groep van landen. Figuur 2 (rechts) illustreert ten slotte dat het niet enkel mogelijk is om de complexiteit van geografische entiteiten zoals landen of gemeenten te berekenen, maar ook voor sectoren, zodat ook nagegaan kan worden welke sectoren gekenmerkt worden door de hoogste of laagste mate van complexiteit.



Figuur 2: Grafiek met de tijdsevolutie in de complexiteitsindex voor verschillende landen (links) en verschillende HS2007-sectoren (rechts) o.b.v. internationale handelsgegevens.

Bron: Cristelli et al. (2013)

2.4 Economische complexiteitsindex (ECI)

Op basis van deze literatuur wordt eerst een methodologie ontwikkeld om een economische complexiteitsindex te berekenen aan de hand van de tewerkstellingsdata op vestigingsniveau voor de gemeenten van Vlaanderen (inclusief Brussel). Deze data zijn beschikbaar via de Rijksdienst voor Sociale Zekerheid. Het doel van de methodologie is om uit deze gegevens het productieve vermogen van elke gemeente af te leiden en uit te drukken in een numerieke complexiteitsmaatstaf. Deze index kan als een kwantitatieve basis gebruikt worden om het productief vermogen van Vlaamse gemeentes te monitoren. Het feit dat deze data gebaseerd zijn op het vestigingsniveau zorgt voor een representatieve weergave van de verdeling van de economische activiteit overheen de regio's. Sectie 4.2 gaat verder in

op het belang van vestigingsdata. Hiernaast wordt ook de NACE-BEL categorie gerapporteerd van de vestiging. Dit is de Belgische versie van de Statistische nomenclatuur van de Economische activiteiten in de Europese Gemeenschap en vormt het referentiekader voor de productie en verspreidingen van statistieken met betrekking tot economische activiteiten in België. Wat de geografische indeling betreft wordt er gewerkt op het niveau van de gemeenten op basis van de postcodes uit de RSZ-databank. Per gemeente zal voor elke aanwezige industrie het aantal tewerkgestelden berekend worden. Deze dataset zal de basis vormen voor de berekeningen van de economische complexiteitsindex.

2.5 Kennis complexiteitsindex (KCI)

Daarnaast wordt ook een alternatieve kennis complexiteitsindex uitgewerkt, die de complexiteit van Vlaamse gemeentes in kaart brengt aan de hand van patentactiviteiten. De data over patenten is afkomstig van het Expertisecentrum Onderzoek en Ontwikkelingsmonitoring (ECOOM, 2022). In deze dataset worden de gepubliceerde patenten gerapporteerd voor ondernemingen, universiteiten en non-profit organisaties van de overheid samen met hun postcode en publicatienummer. Deze patentdata stellen ons in staat om, analoog aan de ECI, een complexiteitsmaatstaf te berekenen op basis van het aantal patenten in een bepaalde regio en sector.

3. Methodologie

Deze sectie zet de berekeningswijze van de complexiteitsindex uiteen. Kort samengevat kan de index berekend worden door de informatie die vervat zit in diversiteit en ubiquiteit te combineren in een matrix (de inputmatrix, zie 3.1) en de ECI of KCI te berekenen op basis van de reflectiemethode (zie 3.2) of aan de hand van lineaire algebra (zie 3.3). Diversiteit wordt hierbij gedefinieerd als de hoeveelheid industrieën en sectoren die er aanwezig zijn in een gemeente. Ubiquiteit wordt gedefinieerd als de hoeveelheid gemeenten die een bepaalde industrie² bezitten. Sectie 3.4 toont een grafische weergave en sectie 3.5 werkt een fictief motiverend voorbeeld uit.

3.1 De inputmatrix

De economische complexiteitsindex wordt berekend aan de hand van een inputmatrix die weergeeft of een economische activiteit aanwezig is in een gemeente, op basis van de hoeveelheid tewerkstelling of patenten in de betreffende industrie. De inputmatrix M_{gi} bestaat uit rijen die de gemeenten van Vlaanderen weergeven en kolommen voor hun economische activiteiten. Elk element in deze matrix geeft weer of een gemeente g gespecialiseerd is in activiteit i of niet. De aanwezigheid van een bepaalde activiteit wordt gemeten met de Locatie Quotiënt (LQ). De LQ is een maatstaf voor de industriële specialisatie van een regio en identificeert welke economische activiteiten relevant zijn voor een gebied. Ze wordt berekend aan de hand van de sectorale tewerkstelling in een regio tegenover een grotere geografische entiteit, in dit geval de gehele Vlaamse economie. Laat X_{gi} het aantal tewerkgestelden weergeven voor een activiteit i in gemeente g , dan wordt de LQ voor een industrie geformuleerd als:

$$LQ_{gi} = \frac{X_{gi}}{\sum_i X_{gi}} / \frac{\sum_g X_{gi}}{\sum_{g,i} X_{gi}}$$

Een element in de inputmatrix M_{gi} is gelijk aan 1 als de Locatie Quotiënt groter of gelijk is aan 1 en anders gelijk aan 0. Dit is het geval als het tewerkstellingsaandeel van sector i in gemeente g groter is dan het totale tewerkstellingsaandeel van sector i in de totale tewerkstelling.

$$M_{gi} = 1 \text{ als } LQ_{gi} \geq 1$$

Door het gebruik van tewerkstellingsdata op gemeenteniveau kan deze definitie echter leiden tot de ondervertegenwoordiging van lokale industrieën die bestaan uit niet-verhandelbare sectoren en diensten in grote of meer diverse regio's. Neem bijvoorbeeld apotheken (NACE-BEL code 44730) in de stad Antwerpen: deze sector heeft een locatie quotiënt kleiner dan 1. Hierdoor zal ze niet mee opgenomen worden in de inputmatrix volgens bovenstaande specificatie. Er zijn echter in 2020 408

² Industrie, sector en activiteit worden in dit rapport gebruikt als synoniemen.

voltijds equivalente werknemers tewerkgesteld in Antwerpen in deze sector. De stad bezit dus wel degelijk vaardigheden en kennis om zijn inwoners te voorzien van geneesmiddelen. Eén manier om dit op te lossen is door een industrie als aanwezig te beschouwen in een regio indien er een bepaalde hoeveelheid werknemers, x , aanwezig is, onafhankelijk van de Locatie Quotiënt:

$$M_{gi} = 1 \text{ als } LQ_{gi} \geq 1 \text{ of } X_{gi} \geq x$$

3.2 De reflectiemethode

De initiële diversiteit en ubiquiteit kunnen worden gemeten door respectievelijk te sommeren over de rijen en de kolommen van de inputmatrix M_{gi} :

$$Diversiteit = k_{g,0} = \sum_i M_{gi} \quad (1)$$

$$Ubiquiteit = k_{i,0} = \sum_g M_{gi} \quad (2)$$

Diversiteit is de initiële complexiteitsmaatstaf voor de regio's. Ze meet het aantal activiteiten waarin een gemeente gespecialiseerd is. Als een regio een grote diversiteit heeft, zal ze een relatief groot productief vermogen bezitten, wat af te leiden valt uit het feit dat ze bovengemiddeld actief is in een groot aantal economische activiteiten. Niet alle activiteiten zijn echter even complex. Een activiteit die in veel gemeenten terug te vinden is, zal makkelijker te repliceren zijn en minder productieve kennis vergen dan een activiteit waarin maar enkele gemeenten gespecialiseerd zijn. De ubiquiteit verzamelt deze informatie: ze geeft het aantal gemeenten weer die in een bepaalde activiteit gespecialiseerd zijn. De ubiquiteit schat de 'verwaarloosbaarheid' van de economische waarde van economische activiteiten in met als centrale idee dat hoe meer gemeenten actief zijn in bepaalde sectoren, hoe lager de economische toegevoegde waarde van de betrokken activiteiten.

In de reflectiemethode van Hidalgo en Hausmann ([2009](#)) zal de uiteindelijke complexiteit berekend worden door informatie over diversiteit en ubiquiteit te combineren over N iteraties (stappen):

$$k_{g,N} = \frac{1}{k_{g,0}} \sum_i M_{gi} * k_{i,N-1} \quad (3)$$

$$k_{i,N} = \frac{1}{k_{i,0}} \sum_g M_{gi} * k_{g,N-1} \quad (4)$$

In iteratie $N=1$ zal $k_{g,1}$ de gemiddelde ubiquiteit uitdrukken voor de activiteiten waarin een gemeente g gespecialiseerd is. Gelijkaardig zal $k_{i,1}$ de gemiddelde diversiteit weergeven van de gemeenten die gespecialiseerd zijn in een activiteit i . Elke iteratie voor $k_{g,N}$ zal tot een nauwkeurigere schatting komen van de complexiteit van een gemeente g door gebruik te maken van informatie over de complexiteit van activiteiten waarin gemeente g gespecialiseerd is. Iedere iteratie van $k_{g,N}$ neemt namelijk informatie op van $k_{i,N-1}$. $k_{g,2}$ zal de informatie uit $k_{i,1}$ gebruiken en zal de gemiddelde diversiteit bevatten van gemeenten met een vergelijkbare set aan activiteiten als gemeente g . Gelijkaardig zal elke iteratie voor $k_{i,N}$ een fijnere schatting geven voor de complexiteit van een activiteit i door gebruik te maken van informatie over de diversiteit van gemeenten die gespecialiseerd zijn in een activiteit i . Het proces van iteraties stopt wanneer er geen verdere informatie onttrokken kan worden. Dit gebeurt wanneer er geen veranderingen meer optreden in de rangschikking van gemeenten en/of activiteiten.

De reflectiemethode heeft twee gekende problemen, zie onder meer Caldarelli et al. (2012). De eerste moeilijkheid is dat de interpretatie van even en oneven iteraties van elkaar verschillen. Zo omschrijft $k_{g,1}$ de gemiddelde ubiquiteit van de economische activiteiten in gemeente g uit stap 0 in vergelijking (1). Gemeenten met vaker voorkomende (en dus minder waardevolle) activiteiten hebben daarom een hogere gemeten complexiteit in stap $N = 1$. Maar de empirische evidentie leert ons dat dergelijke gemeenten doorgaans weinig gediversifieerd zijn, daar vooral gediversifieerde gebieden in staat zijn om zeldzame activiteiten te ontplooien, wat voor deze gemeenten een lage complexiteit oplevert in stap 0. Deze negatieve correlatie tussen even en oneven complexiteitswaarden zet zich ook door bij hogere iteraties (stappen). Meer algemeen geven complexiteitsmaatstaven voor gemeenten, $k_{g,N}$, in *even* ($N = \text{even}$) stappen de mate van diversiteit van hun economische activiteiten weer en in *oneven* stappen de gegeneraliseerde ubiquiteit van deze activiteiten. Analoog schatten de maatstaven voor sectoren gemeenten, $k_{i,N}$, in *even* ($N = \text{even}$) stappen de ubiquiteit van deze sectoren in terwijl oneven stappen de diversiteit van de gemeenten waar ze voorkomen in kaart brengen. Dit probleem wordt voornamelijk ondervangen door enkel waarden voor even of oneven stappen te beschouwen, afhankelijk van de interesse in diversiteit of ubiquiteit. De economische complexiteit van gemeenten, ECI_g , stemt dan overeen met (genormaliseerde versies van) *even* iteraties voor $k_{g,N}$, $ECI_g = k_{g,N}^{(2n)}$; voor industrieën stemt ze overeen met *oneven* iteraties voor $k_{i,N}$, $ECI_i = k_{i,N}^{(2n+1)}$.

Een tweede moeilijkheid is de keuze van het aantal stappen, N , die doorlopen moeten worden om tot een finale indicator te komen. Bijkomende moeilijkheid daarbij is dat de berekende complexiteitswaarden voor elke gemeente en elke sector convergeren naar een identieke waarde naarmate het aantal stappen toeneemt. Het volstaat dus niet om stappen (3) en (4) een groot aantal keer te herhalen, omdat de berekende complexiteit daardoor identiek wordt voor alle gemeenten en sectoren. De pragmatische oplossing is om stappen (3) en (4) te herhalen tot de rangschikking van gemeenten en sectoren naar hun complexiteit niet langer veranderd en stabiel blijft. Hidalgo en Hausmann (2009) stellen dat dit doorgaans het geval is na $N=16$ stappen.

3.3 ECI op basis van lineaire algebra

Omwillen van bovenstaande complicaties werd een alternatieve methode ontwikkeld om de complexiteit van locaties en producten/sectoren te berekenen aan de hand van lineaire algebra. Door vergelijking (4) toe te voegen aan vergelijking (3) bekomen we dat:

$$k_{g,N} = \frac{1}{k_{g,0}} \sum_i M_{gi} * \frac{1}{k_{i,0}} \sum_g M_{gi} * k_{g',N-2} \quad (5)$$

$$k_{g,N} = \sum_{g'} k_{g',N-2} * \sum \frac{M_{gi} M_{g'i}}{k_{g,0} k_{i,0}} \quad (6)$$

$$k_{g,N} = \sum_{g'} \tilde{M}_{gg'} k_{g',N-2} \quad (7)$$

Aan vergelijking (7) is voldaan wanneer $k_{g,N} = k_{g,N-2} = 1$. De oplossing kan berekend worden aan de hand van de eigenvector van $\tilde{M}_{gg'}$ gelinkt aan de grootste eigenwaarde. Omdat $\tilde{M}_{gg'}$ een stochastische matrix is zal dit een eigenvector van enen zijn, aangezien voor $n \rightarrow \infty$ alle berekende complexiteitswaarden convergeren naar eenzelfde waarde. Daardoor zal de informatie over economische complexiteit in de eigenvector van de tweede grootste eigenwaarde zitten, aangezien dat de vector is die de gemeenten het beste verdeelt op basis van de activiteiten die erin aanwezig zijn.

$K_G = \text{Eigenvector van } \tilde{M}_{gg'}, \text{ met de tweede grootste eigenwaarde}$

Om tot de uiteindelijke economische complexiteitsindex van de gemeenten te komen wordt K_G genormaliseerd om verschillen in schaal te elimineren:

$$ECI_g = \frac{K_g - \bar{K}_g}{std(K_g)}$$

Een gemeente die een ECI_g groter dan nul heeft, zal dus complexer zijn dan de gemiddelde complexiteit overheen alle gemeenten. Analoog door vergelijking (3) toe te voegen aan (4) kan de economische complexiteitsindex voor de industrieën gevonden worden:

$$ECI_i = \frac{K_i - \bar{K}_i}{std(K_i)}$$

Met

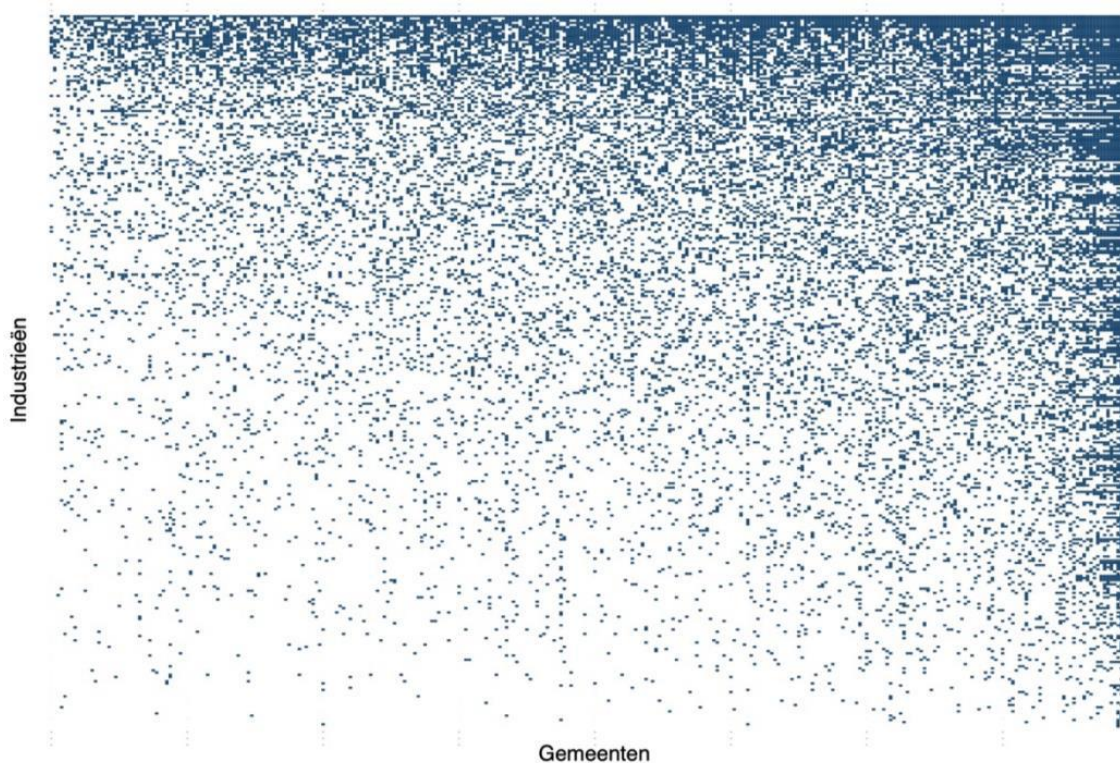
$K_i = \text{Eigenvector van } \tilde{M}_{ii}, \text{ met de tweede grootste eigenwaarde}$

Een laatste potentieel relevante normalisatie is voor bevolking of oppervlakte. Als twee gemeenten eenzelfde niveau van complexiteit behalen, maar de eerste gemeente doet dit met de helft van de oppervlakte of tewerkgestellten van de tweede gemeente, suggereert dit dat de eerste gemeente een groter productievermogen per capita heeft. In het achterhoofd houdende dat appendix 11 van Hidalgo en Hausmann (2009) afraadt om te normaliseren met bevolking of oppervlakte omdat dit lineariteit veronderstelt, kunnen ook complexiteitswaarden per capita informatief zijn.

3.4 Grafische weergave voor de Vlaamse economie

Figuur 3 illustreert de grafische representatie van de getransformeerde inputmatrix voor Vlaamse gemeenten en industrieën in 2020, met een punt voor elke industrie die voorkomt in een bepaalde gemeente. Dit komt overeen met de inputmatrix uit sectie 3.1. en tabel 3 uit het voorbeeld in sectie 3.5 (zie verder). De industrieën zijn weergegeven in de rijen en gerangschikt volgens hun ubiquiteit van onder naar boven. De industrie die onderaan staat komt dus het minst vaak voor in de verzameling van gemeenten. De gemeenten zijn weergegeven in de kolommen en gerangschikt volgens hun diversiteit van links naar rechts. De gemeente die uiterst rechts staat heeft dus tewerkstelling in het grootste aantal industrieën. Het resultaat is een typische driehoekige structuur van de matrix. Dit betekent dat diverse gemeenten industrieën bezitten die zowel een hoge als lage ubiquiteit hebben. Zoals reeds aangegeven hebben meer diverse gemeenten daardoor een lagere gemiddelde ubiquiteit. Minder diverse gemeenten zijn voornamelijk enkel gespecialiseerd in industrieën met een hoge ubiquiteit. Deze driehoekige structuur beschrijft de geografische verdeling van activiteiten gebaseerd op tewerkstelling en is cruciaal voor de berekening van de economische complexiteit. Ze zorgt ervoor dat er betekenisvolle variatie is in diversiteit en ubiquiteit waaruit informatie onttrokken kan worden om het productief vermogen van gemeenten in Vlaanderen samen te vatten in één index.

Een aantal zaken beïnvloeden hoe prominent deze structuur zich voordoet. Ten eerste wordt de complexiteitsindex in dit onderzoek berekend aan de hand van de tewerkstellingsdata op vestigingsniveau. In tegenstelling tot exportdata bevat deze ook diensten en lokale sectoren. Door deze inclusie zullen een aantal sectoren mee opgenomen worden in de matrix die buiten beschouwing gelaten zouden worden indien er enkel gekeken zou worden naar de vervaardiging van verhandelbare goederen. Aan de ene kant zijn er activiteiten die veel voorkomen in alle gemeenten en dus een hoge ubiquiteit hebben, zoals bijvoorbeeld bouwmarkten (NACE-BEL code 47521), apotheken (NACE-BEL code 47730) en detailhandels in winkels waarbij voedingsmiddelen overheersen (NACE-BEL code 47114). Aan de andere kant zullen er ook lokale industrieën mee opgenomen worden die gespecialiseerd zijn en dus een lage ubiquiteit hebben zoals bijvoorbeeld callcenters (NACE-BEL code 82200) en officieel hoger onderwijs (NACE-BEL code 85421).

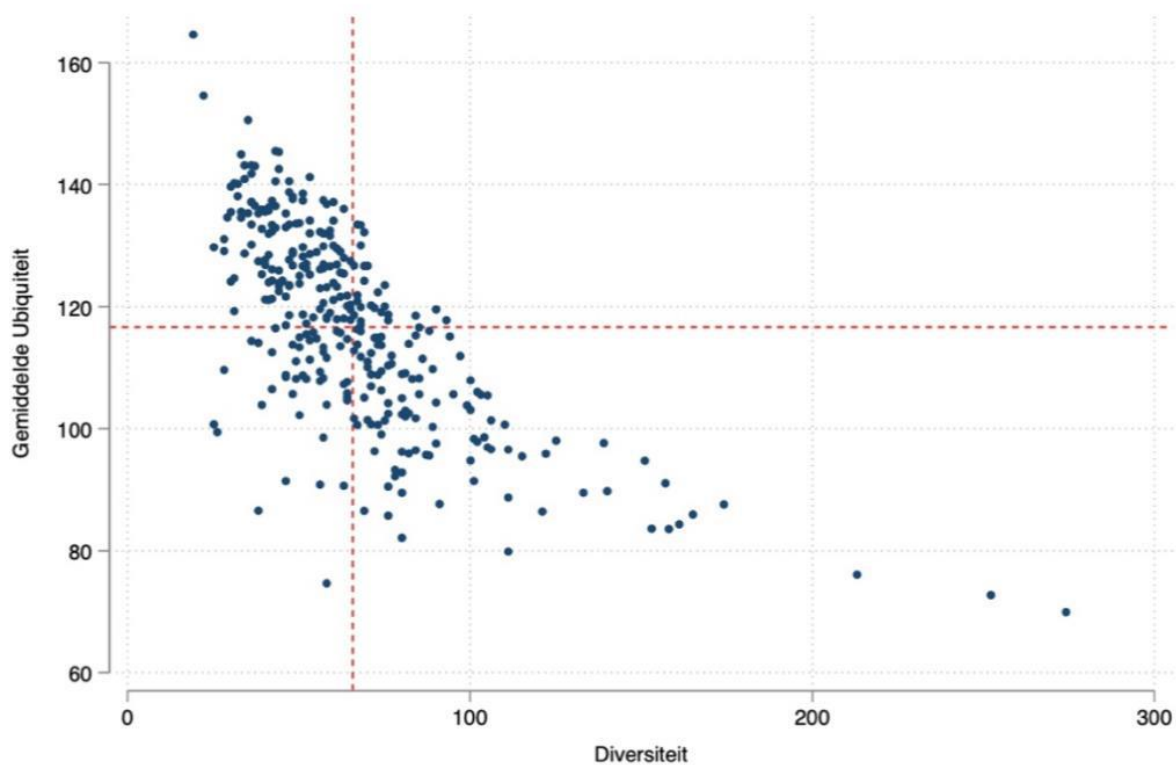


Figuur 3: Structuur van de inputmatrix (situatie in jaar 2020), elk puntje staat voor een industrie die in een gemeente voorkomt, industrieën met een locatie quotiënt groter dan 1 of minstens 20 werknemers worden mee in rekening genomen. Gemeenten zijn gerangschikt van een lage naar hoge diversiteit. Industrieën zijn gerangschikt van een lage naar hoge ubiquiteit.

Bron: eigen berekening o.b.v. RSZ data

Ten tweede speelt de juiste specificatie van de methode om de inputmatrix te construeren ook een rol. In sectie 3.1 werd aangehaald dat een activiteit aanwezig is in een gemeente als ze een locatie quotiënt groter dan 1 heeft of wanneer er een minimum aantal (bijvoorbeeld $x=20$) werknemers aanwezig zijn in

een gemeente. De reden hiervoor is om ondervertegenwoordiging van veelvoorkomende industrieën te vermijden in diverse gemeenten. Het toevoegen van de extra voorwaarde dat een industrie ook aanwezig is in een gemeente wanneer er tenminste 20 personen in tewerkgesteld zijn zal de prominentie van de driehoekige structuur verhogen. In de appendix (figuur A) wordt de inputmatrix weergegeven waarbij deze extra voorwaarde niet aanwezig is. Vergeleken met figuur 3 valt op dat de driehoekige structuur hier minder duidelijk te zien is. Het is ook mogelijk om rekening te houden met activiteiten vanaf één tewerkgesteld persoon. De grafische weergave van deze inputmatrix is weergegeven in de appendix, figuur B. De driehoekige structuur is zeer duidelijk aanwezig in deze matrix. Het gevaar hier is echter dat deze methode te gevoelig is voor industrieën met een lage tewerkstelling. De drempelwaarde van 20 is tot op zekere hoogte arbitrair maar probeert een afweging te maken tussen de volgende twee fenomenen: de ondervertegenwoordiging van veelvoorkomende industrieën in diverse regio's enerzijds en anderzijds de overschatting van de aanwezigheid van een industrie indien een te lage drempelwaarde gebruikt zou worden.



Figuur 4: Relatie tussen diversiteit en ubiquiteit (2020)
Bron: eigen berekening o.b.v. RSZ data

Een direct gevolg van de driehoekige structuur van de inputmatrix is te zien in figuur 4. Hier worden de gemeenten tegen elkaar afgezet op twee assen voor 2020. De x-as geeft de diversiteit van de gemeente weer (of $k_{g,0}$, zie sectie 3.2 vergelijking 1). Op de y-as staat de gemiddelde ubiquiteit van alle sectoren die in de gemeente voorkomen, dus hoe lager de y-waarde, hoe zeldzamer de activiteiten gemiddeld

zijn. Dit komt overeen met de eerste iteratie van de reflectiemethode uit sectie 3.2 voor de gemeenten ($k_{g,1}$ uit vergelijking 3). Figuur 4 bevat dus slechts een beperkte hoeveelheid informatie uit het netwerk van gemeenten en activiteiten (enkel $k_{g,0}$ en $k_{g,1}$). De verticale en horizontale rode stippellijnen geven de gemiddelde waarde weer voor deze twee variabelen. Er is een negatieve correlatie tussen deze twee variabelen op te merken. De meeste gemeenten bevinden zich in het kwadrant linksboven of in het kwadrant rechtsonder. Hoe hoger de diversiteit is in een gemeente, hoe lager de gemiddelde ubiquiteit. Deze bevinding stemt overeen met de algemene literatuur over complexiteit. De reden hiervoor kan zijn dat regio's met een hoge diversiteit een grotere bron aan organisationeel, infrastructureel en kenniskapitaal te benutten om meer gespecialiseerde industrieën te faciliteren die slechts op een beperkt aantal locaties kunnen ontstaan. Ze genieten van een productief vermogen of complexiteit waaruit ze competitieve voordelen halen. Het zou ook een aanwijzing kunnen zijn dat complexe kennis zich moeilijk kan verspreiden overheen regio's en daarom enkel aanwezig is een beperkt aantal gemeenten. Dit fenomeen is gevonden door Balland & Rigby (2017) voor de Verenigde Staten op basis van patentdata.

3.5 Een fictief voorbeeld

Deze sectie werkt een eenvoudig fictief voorbeeld uit, geïnspireerd op Hidalgo en Hausmann (2009), om de interpretatie en de beperkingen van de reflectiemethode verder te duiden. Het voorbeeld vertrekt van 4 gemeenten, $g \in [1,2,3,4]$, met activiteiten in 4 sectoren, $i \in [1,2,3,4]$, waarbij de volgende inputmatrix, M_{gi} , de sectorale tewerkstelling in elke gemeente weergeeft. Daaruit blijkt dat er in totaal 3555 mensen tewerkgesteld zijn, zie tabel 1.

Tabel 1: Inputmatrix

	Sector 1	Sector 2	Sector 3	Sector 4
Gemeente 1	50	1000	640	1500
Gemeente 2	0	110	0	0
Gemeente 3	0	0	75	80
Gemeente 4	0	0	0	100

Aan de hand van de formule uit sectie 3.1. kunnen we deze inputmatrix gebruiken om de locatie quotiënten, LQ_{gi} , voor deze economie te berekenen als volgt:

Tabel 2: Locatie quotiënten

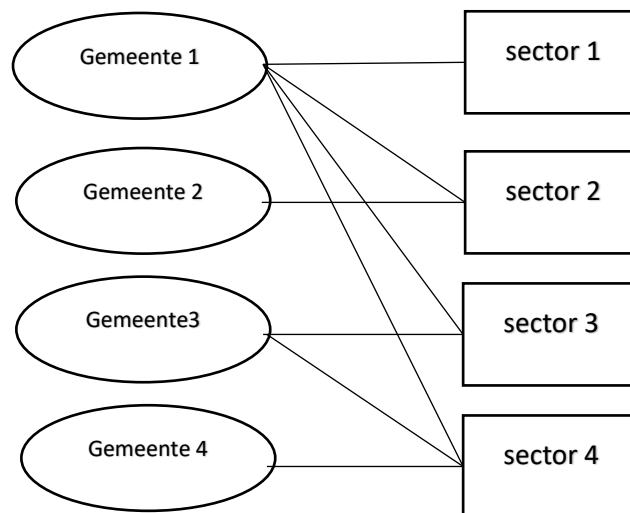
	Sector 1	Sector 2	Sector 3	Sector 4
Gemeente 1	$\frac{50}{50 + 1000 + 640 + 1500} \sqrt{\frac{50 + 0 + 0 + 0}{3555}} = 1.1$	1	1	1
Gemeente 2	0	3.2	0	0
Gemeente 3	0	0	2.41	1.09
Gemeente 4	0	0	0	2.12

Dit laat vervolgens toe om de getransformeerde inputmatrix met economische sectoren die relevant zijn op te stellen, door enkel sectoren met een locatiequotiënt hoger dan 1 te weerhouden. In dit eenvoudige voorbeeld is dat het geval voor alle sectoren, zoals te zien in tabel 3.

Tabel 3: Getransformeerde inputmatrix

	Sector 1	Sector 2	Sector 3	Sector 4
Gemeente 1	1	1	1	1
Gemeente 2	0	1	0	0
Gemeente 3	0	0	1	1
Gemeente 4	0	0	0	1

Figuur 5 geeft een kort overzicht van het economisch netwerk. Merk op dat gemeente 1 het sterkst gedifferentieerd is en actief is in alle sectoren, daar waar gemeentes 2 en 4 slechts actief zijn in één sector. De enige activiteit van gemeente 2 is echter relatief zeldzaam, aangezien ze enkel nog voorkomt in de meest gediversifieerde gemeente 1, terwijl gemeente 4 actief is in een vaak voorkomende sector, waar 3 gemeenten actief zijn. De initiële complexiteitsmaatstaf voor de gemeenten en sectoren kunnen vervolgens berekend worden aan de hand van vergelijkingen (1) en (2), zoals weergegeven in tabel 4.



Figuur 5: Schematische weergave van de getransformeerde inputmatrix

Tabel 4: Initiële diversiteit en ubiquiteit

Gemeente	Diversiteit, $k_{g,0}$	Sector	Ubiquiteit, $k_{i,0}$
Gemeente 1	$1+1+1+1 = 4$	Sector 1	$1+0+0+0$
Gemeente 2	1	Sector 2	2
Gemeente 3	2	Sector 3	2
Gemeente 4	1	Sector 4	3

Tabel 4 bevestigt dat gemeente 1 de meest diverse gemeente is, gevolgd door gemeente 3. Bovendien

bevestigt de tabel dat sector 2, waar gemeente 2 zich in specialiseert, tamelijk zeldzaam is en een lagere ubiquiteit heeft dan bijvoorbeeld sector 3. Bovendien is de meeste gediversifieerde gemeente 1 ook als enige actief in sector 1, die de laagste ubiquiteit (hoogste zeldzaamheid) heeft.

Aan de hand van vergelijkingen (3) en (4) kunnen verschillende reflecties berekend worden van de maatstaven voor gemeenten en sectoren. Tabel 5 rapporteert resultaten voor iteraties $N = 1$ en 2.

Tabel 5: Reflectiemethode voor $N=1$ en 2

Gemeente	$k_{g,1}$	$k_{g,2}$	Sector	$k_{i,1}$	$k_{i,2}$
Gemeente 1	$(1/4)(1+2+2+3)=2$	$(1/4)(4+2.5+3+2.33)=2.96$	Sector 1	$(1/1)(4+0+0+0)=4$	$(1/1)(2+0+0+0)=2$
Gemeente 2	2	2,5	Sector 2	2,5	2
Gemeente 3	2,5	2,67	Sector 3	3	2,25
Gemeente 4	3	2,33	Sector 4	2,33	2,5

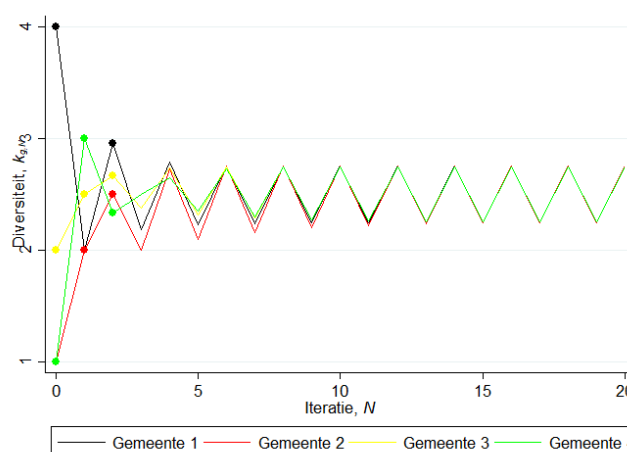
Wat betreft de complexiteit voor gemeenten, $k_{g,N}$, geven oneven iteraties zoals toegelicht in sectie 3.2. de gemiddelde ubiquiteit van de economische activiteiten weer, waar de kolom voor $k_{g,1}$ in tabel 5 verduidelijkt dat de eerste twee gemeenten actief zijn in meer zeldzame sectoren en daarom een lagere gemiddelde ubiquiteit kennen. Even iteraties schatten de diversiteit van de gemeenten in en houden steeds meer rekening met de ubiquiteit van de activiteiten waarin gemeenten actief zijn. Daarom rangschikt de methode gemeente 2 boven gemeente 4 in stap $N=2$, hoewel ze beiden slechts actief zijn in één sector, aangezien gemeente 2 zich specialiseert in een zeldzame activiteit die slechts terug te vinden is in een meer gediversifieerde gemeente, gemeente 1. Dit suggereert dat gemeente 2, ondanks haar beperkte diversiteit, wel degelijk productieve capaciteiten heeft om relatief geavanceerde economische activiteiten te ondersteunen. Dit is niet het geval voor gemeente 4, die zich specialiseert in sectoren die zowel in gediversifieerde als minder gediversifieerde gemeenten voorkomen, wat suggereert dat deze activiteiten minder geavanceerde kennis en/of infrastructuur vergen. Hoewel gemeenten 2 en 4 dus exact 1 economische activiteit ontplooiën, is de complexiteitsmethode in staat beide gemeenten te differentiëren en uit de bredere informatie uit het productienetwerk af te leiden dat gemeente 2 wellicht een meer complexe productiestructuur bevat.

Analoog geeft iteratie 1 voor de ubiquiteit aan dat gemeenten die actief zijn in sector 1 (in dit geval, enkel de meest diverse gemeente 1) doorgaans zeer gediversifieerd zijn. Sector 4 komt vaker voor in minder gediversifieerde gemeenten, waardoor het wellicht een eenvoudiger en/of minder restrictief productieproces kent. De even iteratie voor $N=2$ geeft een meer fijnmazigere inschatting van de ubiquiteit van elke sector, waarbij sectoren 1 en 2 nu vergelijkbaar worden qua ubiquiteit rekening houdend met de vergelijkbare gemiddelde diversiteit van de actieve gemeenten 1 en 2. Onderstaande tabel 6 verduidelijkt de verschillende interpretaties voor iteraties $N=0, 1$ en 2.

Tabel 6: Interpretatie van de eerste 2 iteraties

Definitie	Concept	Interpretatie
$k_{g,0}$	Diversiteit	Aantal activiteiten in gemeente g
$k_{i,0}$	Ubiquiteit	Aantal gemeenten actief in sector i
$k_{g,1}$	Gemiddelde ubiquiteit	Gemiddelde ubiquiteit van de activiteiten in gemeente g
$k_{i,1}$	Gemiddelde diversiteit	Gemiddelde diversiteit van gemeenten actief in sector i
$k_{g,2}$	Gegeneraliseerde diversiteit	Gemiddelde diversiteit van gemeenten met gelijkaardige activiteiten als gemeente g
$k_{i,2}$	Gegeneraliseerde ubiquiteit	Gemiddelde ubiquiteit van de activiteiten in gemeenten actief in sector i

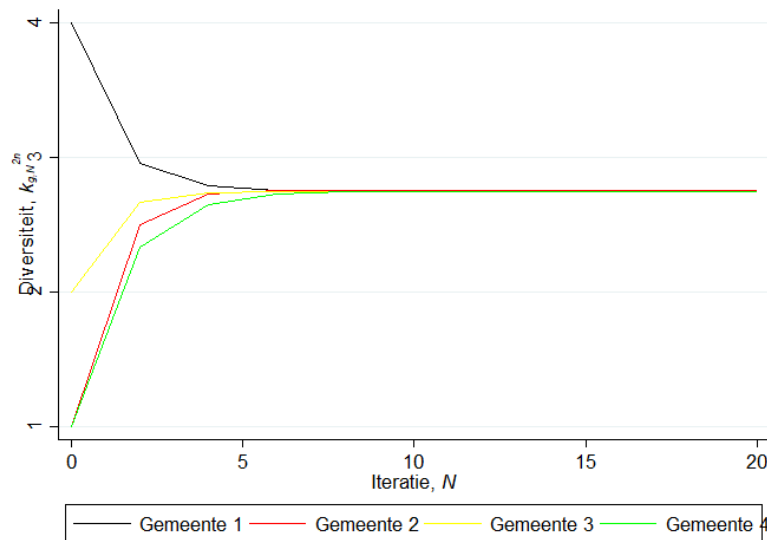
Verdere verfijningen voor berekende complexiteitswaarden voor de 4 gemeenten en sectoren kunnen bekomen worden door vergelijkingen (3) en (4) iteratief toe te passen, om steeds meer rekening te houden met de diversiteit en ubiquiteit in het netwerk. De resultaten van de eerste 20 iteraties voor diversiteit, $k_{g,N}$, staan weergegeven in figuur 6. De resultaten voor de eerste twee stappen komen exact overeen met de gerapporteerde gegevens in tabellen 4 en 5, zoals aangegeven met de gekleurde bollen. Deze figuur toont twee zaken. Ten eerste is duidelijk te zien dat even en oneven iteraties verschillende interpretaties kennen, zoals toegelicht in sectie 3.2. Meer concreet wordt daar toegelicht dat even iteraties de diversiteit van de gemeenten inschat, rekening houdend met de ubiquiteit van hun activiteiten, terwijl oneven stappen hun ubiquiteit inschatten, wat de negatieve correlatie verklaart. Dit maakt de figuur moeilijk te interpreteren. Ten tweede toont de figuur duidelijk dat naarmate de iteraties toenemen, alle complexiteitswaarden convergeren naar eenzelfde waarde en dus uiteindelijk aan informatie inboeten. Daarom wordt het wenselijke aantal iteraties bepaald opdat de rangschikking van gemeenten naar complexiteit niet langer wijzigt, zie ook figuur 6 hieronder.



Figuur 6: Diversiteit, resultaten voor de eerste 20 iteraties

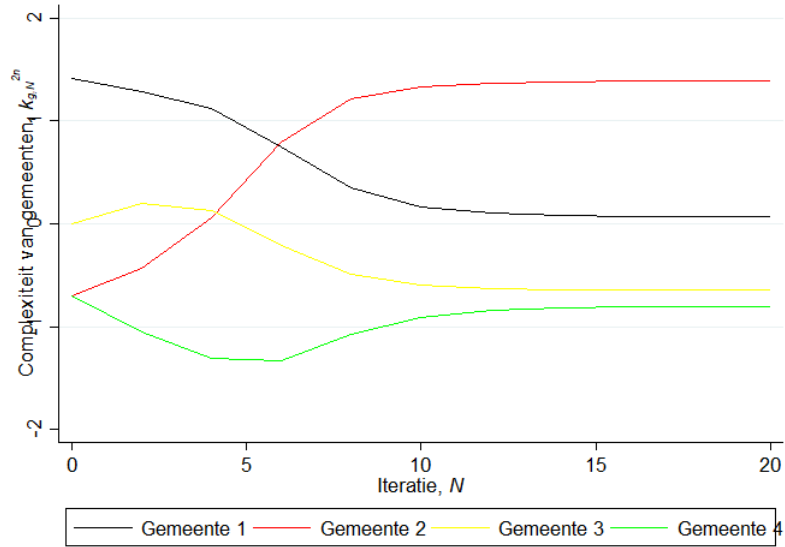
Figuur 7 beperkt zich tot de complexiteitsresultaten voor even iteraties, $k_{g,N}^{(2n)}$, wat Hidalgo en Hausmann (2009) de economische complexiteit van de gemeenten zouden noemen. Figuur 7 is dus identiek aan figuur 6, behalve dat ze enkel resultaten voor even iteraties toont. Deze geven een meer

eenduidig beeld en illustreren hoe gemeente 1 doorgaans als de meest gediversifieerde regio lijkt te worden gerangschikt. Daarnaast duikt gemeente 3 stelselmatig als tweede meest diverse regio op. Gemeente 4, die net als gemeente 2 slechts actief is in één sector, maar in een sector met hogere ubiquiteit (lagere zeldzaamheid), belandt op de laatste plaats qua diversiteit.



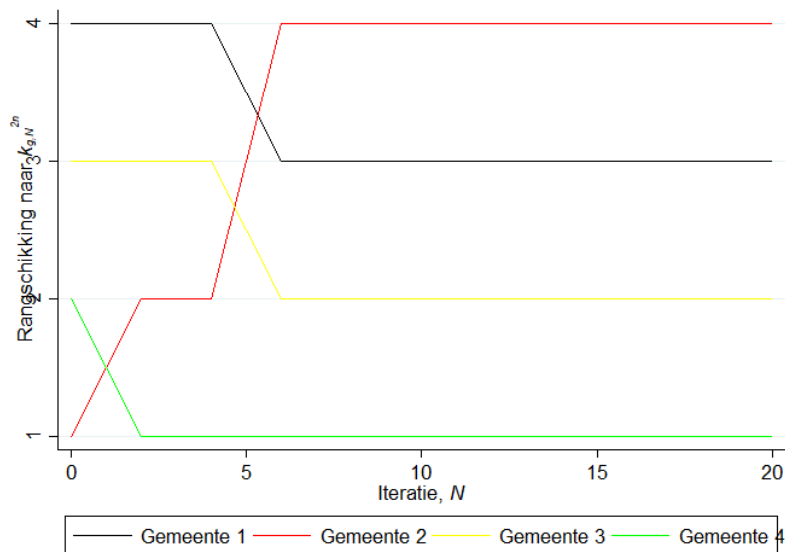
Figuur 7: Diversiteit, resultaten voor de 10 eerste even iteraties

Om de diversiteitsindex $k_{g,N}^{(2n)}$ om te zetten in een complexiteitsindex, ECI_g , wordt deze conform sectie 3.3. gestandaardiseerd, zoals weergegeven in figuur 8. Op deze figuur is duidelijker te zien hoe gemeente 2 vanaf iteratie 5 gerangschikt wordt als de meest complexe gemeente, omdat het zich louter specialiseert in een sector die doorgaans voorkomt in de meest gediversifieerde gemeenten. Bovendien stabiliseren de schattingen zich vanaf iteratie 10, waarna alle beschikbare informatie vervat zit in de schattingen. De rangschikking stabiliseert vanaf de zesde iteratie. In dit eenvoudige voorbeeld blijken 6-10 iteraties dus te volstaan om de finale complexiteitsindices voor de gemeenten te berekenen.



Figuur 8: De economische complexiteit van gemeenten

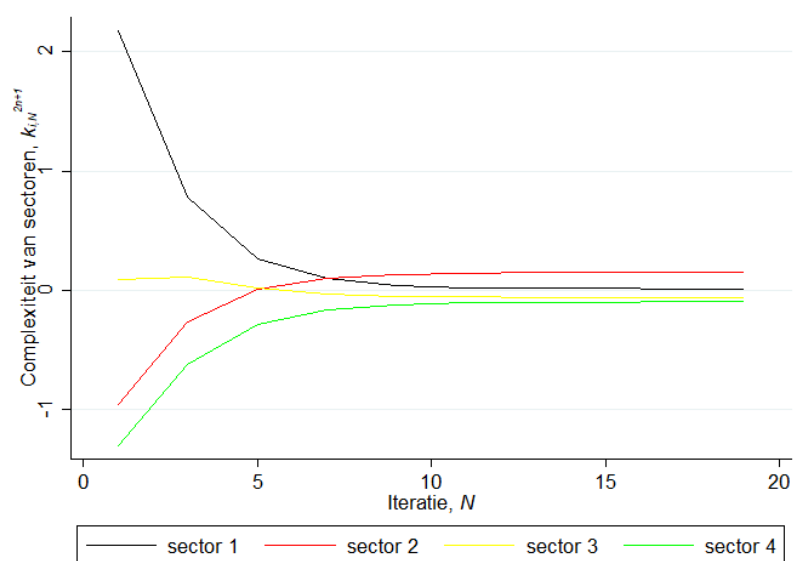
Dit laatste is ook te zien in figuur 9, die de gemeenten rangschikt naar hun gemeten complexiteit in figuur 8. De figuur toont duidelijk dat deze rangschikking stabiliseert vanaf de zesde iteratie.



Figuur 9: Rangschikking van gemeenten naar economische complexiteit

Ten slotte toont figuur 10 de analoge schattingen voor de economische complexiteit van sectoren, ECI_i , aan de hand van de oneven iteraties, $k_{i,N}^{(2n+1)}$, en na de standaardisatie uit sectie 3.3. De figuur toont dat de reflectiemethode al na een zevental iteraties tot de consensus komt dat sector 2 de hoogste complexiteit kent, ze komt slechts in 2 gemeenten voor, waarvan er 1 gemeente is met een hoge

diversiteit. Sector 4 komt het vaakst voor, wat duidt op een hoge ubiquiteit en resulteert in een lage complexiteit (niet-complexe producten kunnen door veel regio's geproduceerd worden).



Figuur 10: Economische complexiteit van sectoren

4. De geografie van bestaande clusterinitiatieven

4.1 Speerpuntclusters en cluster lidmaatschap

Sinds 2017 werden in totaal zeven speerpuntclusters opgezet. Elke cluster wordt gekenmerkt door een sector overschrijdend lidmaatschap. Tegelijk heeft elke cluster zijn eigen focus, of 'strategisch domein'. Concreet onderscheiden we de volgende clusters: Catalisti (chemie en kunststoffen), SIM (materialen), VIL (logistiek), Flanders' Food (agrovoeding), Flux 50 (energie en bouw), De Blauwe Cluster (blauwe economie) en MEDVIA (innovatieve gezondheidszorg)³. De speerpuntclusters opereren in een 'triple helix' structuur waarbij ondernemingen, kennisinstellingen en de overheid nauw samenwerken. Het doel van de speerpuntclusters is om 'bottom-up' een lange termijn competitiviteitsprogramma te ontwikkelen en dit via samenwerkingsinitiatieven op vlak van onderzoek en innovatie uit te voeren⁴.

Eens een cluster is opgericht kunnen bedrijven jaarlijks vrijblijvend kiezen om lid te worden van één of meerdere clusters. Bedrijven die lid willen worden betalen een jaarlijkse lidmaatschapsbijdrage. Deze bijdragen zijn verschillend van cluster tot cluster en kunnen ook afhangen van de grootte van het bedrijf of de betrokkenheid bij de clusterwerking.

In opdracht van EWI stelt ECOOM-STORE jaarlijks economische en innovatie kernindicatoren op. Hiertoe worden de btw-nummers van de clusterleden gekoppeld aan de Belfirst en PATSTAT gegevensdatabanken. ECOOM-STORE is verantwoordelijk om het clusterlidmaatschap te vertalen naar een reeks relevante btw-nummers. In een eerste stap wordt hiervoor gebruik gemaakt van de facturatiegegevens van betalende leden. Aan deze lijst worden, in samenspraak met de cluster, de relevante verbonden ondernemingen toegevoegd (dit zijn btw-nummers die eenzelfde *Global Ultimate Owner* hebben als het betalende lid en die betrokken zijn bij de clusterwerking).

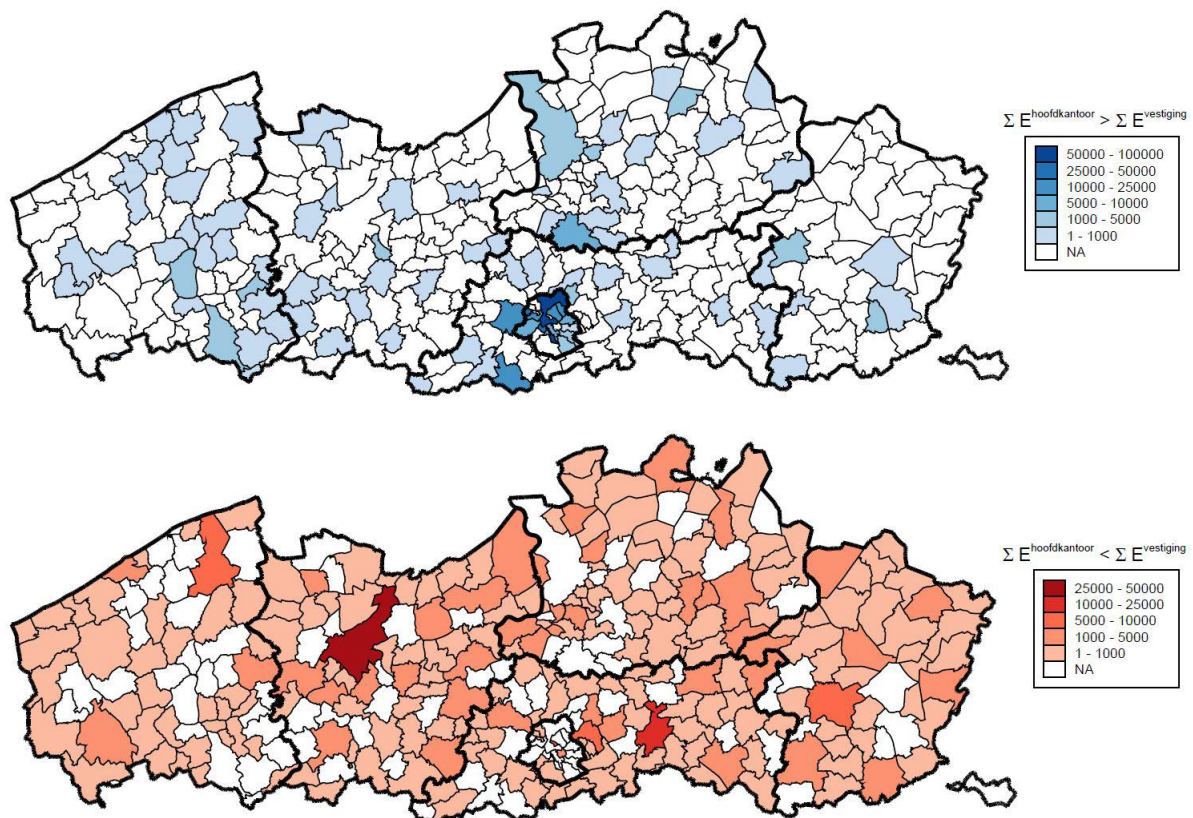
4.2 Maatschappelijke zetel versus vestiging

De analyses naar complexiteit gebeuren op basis van de RZS-gegevens op vestigingsniveau. Concreet bevat deze databank voor elk ondernemingsnummer de verschillende vestigingen die tot dezelfde onderneming behoren met hun respectievelijke locatie en aantal werknemers (uitgedrukt in voltijdse equivalenten). Op deze manier kan een onderscheid gemaakt worden tussen de locatie van de Maatschappelijke Zetel (die bij het ondernemingsnummer hoort) en de locatie van één of meerdere vestigingen waar de economische activiteit of O&O plaatsvindt.

³ De beleidsrapporten die STORE voor elk van deze clusters heeft opgesteld, kunnen op de STORE 2016-2020 website teruggevonden worden: Publicaties — Steunpunt Economie en Ondernemen (steunpunt-economie-ondernemen.be)

⁴ Meer informatie over de speerpuntclusters kan op de VLAIO website teruggevonden worden: [Speerpuntclusters | Agentschap Innoveren en Ondernemen \(vlaio.be\)](https://www.vlaio.be/Agentschap-Innoveren-en-Ondernemen)

Het belang hiervan werd reeds aangekaart in het rapport van Baert en Reynaerts (2017). Figuur 11 toont de gemeentes die gekenmerkt worden door een overschatting (bovenaan, blauw) en onderschatting (onderaan, rood) van de lokale tewerkstelling indien de locatie van de hoofdzetel als referentie genomen wordt. De figuur toont aan dat er disproportioneel veel hoofdkantoren in Brussel gevestigd zijn. Omgekeerd zien we ook dat er veel vestigingen in Oost-Vlaanderen liggen, wiens maatschappelijke zetel elders ligt. Om een correcte geografische representatie te verkrijgen, zowel wat de economische complexiteitsindex betreft als wat het clusterlidmaatschap betreft, is het cruciaal om de berekeningen op het niveau van de vestiging uit te voeren.



Figuur 11: Het belang van vestigingsdata bij het berekenen van agglomeratie-effecten. De figuur toont de gemeentes die gekenmerkt worden door een overschatting (figuur bovenaan, blauw) en onderschatting (figuur onderaan, rood) van de lokale tewerkstelling indien de locatie van de hoofdzetel als referentie genomen wordt.

Bron: Baert, L., Reynaerts, J. (2017)

4.3 Complementariteit met geografische clusters

Op basis van de economische complexiteitsindex kunnen we de complexiteit en dus de schatting voor het productieve vermogen weergeven per postcode. We kunnen eveneens de graad van clusterlidmaatschap (het percentage van vte's t.o.v. de totale cluster) weergeven per postcode. Hierbij merken we op dat beide maatstaven berekend worden op basis van vestigingsdata, hetgeen een zo

nauwkeurig mogelijk weergave van de reële tewerkstelling geeft. In het geval van de kennis complexiteitsindex beschikken we enkel over het aantal patenten aangevraagd door de bedrijven (de Maatschappelijke Zetel). Deze spreiding kan ook vergeleken worden met de spreiding van de clustertewerkstelling, maar dit zal een enigszins vertekend beeld opleveren. De complementariteit tussen geografische clusters (weergegeven o.b.v. de economische complexiteitsindex) en de clusterinitiatieven, zal grafisch worden weergegeven m.b.v. figuren analoog aan figuur 1.

4.4 Mogelijke bijkomende clusterinitiatieven

Naast de analyse welke regio's het meest complex zijn, kunnen we ook analyseren welke sectoren bijdragen tot een hoge complexiteitsindex. Door deze sectoren te vergelijken met de strategische domeinen van de bestaande speerpuntclusters kunnen we mogelijke lacunes ontdekken en aanbevelingen doen naar potentieel toekomstige clusterinitiatieven. Voor deze analyses doen we beroep op de economische complexiteitsindex (in tegenstelling tot de kennis complexiteitsindex).

5. Verdere toepassingen

Nadat we de geografie van de complexe regio's in kaart hebben gebracht, kunnen we hier nog verdere toepassingen aan koppelen. Eén voor de hand liggende toepassing is een dashboard om de complexiteit van de Vlaamse gemeenten en sectoren doorheen de tijd en ten opzichte van elkaar weer te geven, zoals in figuur 2. Twee bijkomende mogelijkheden worden hier kort aangehaald. Ten eerste kunnen we nagaan in welke regio's er nog een bijkomend potentieel tot groei is. Op basis van de geografie van bestaande en toekomstige bedrijfsterreinen kunnen we bijkomende inzichten leveren in het potentieel van deze bedrijfsterreinen. Daartoe kan een methodologie opgesteld worden om sectoren te identificeren die nauw aansluiten bij het lokale economische weefsel en relevant kunnen zijn om vorm te geven aan de optimale invulling van beschikbare bedrijfsterreinen. Ten tweede kunnen we de complementariteit bestuderen tussen de complexiteit van de regio's en de geografie van grote en kleine bedrijven. Dit biedt een inzicht in het belang van grote bedrijven (of meer concreet grote vestigingen) voor het vergroten van complexiteit in bepaalde regio's.

Referenties

- Amiti, M. (1999). Specialization patterns in Europe. *Weltwirtschaftliches Archiv* 135, 573–593.
- Baert, L., Reynaerts, J. (2017). De effecten van agglomeratie-externaliteiten op productiviteit in -Vlaanderen, Beleidsrapport STORE 17-013.
- Balland, P.-A., Boschma, R., Crespo, J., & Rigby, D. L. (2019). Smart specialization policy in the European Union: relatedness, knowledge complexity and regional diversification. *Regional Studies*, 53(9), 1252-1268.
- Balland, P. A., Broekel, T., Diodato, D., Giuliani, E., Hausmann, R., O'Clery, N., & Rigby, D. (2022). The new paradigm of economic complexity. *Research Policy*, 51(3), 104450.
- Buera, F. J., & Kaboski, J. P. (2012). The Rise of the Service Economy. *American Economic Review*, 102(6), 2540-2569.
- Caldarelli, G., Cristelli M., Gabrielli A., Pietronero L., Scala A., Taccella A. (2012). A Network Analysis of Countries' Export Flows: Firm Grounds for the Building Blocks of the Economy. *PLOS ONE*, 7(10)
- Chávez, J. C., Mosqueda, M. T., & Gómez-Zaldívar, M. (2017). Economic Complexity and Regional Growth Performance: Evidence from the Mexican Economy. *The Review of Regional Studies*, 47(2), 201-219.
- Cristelli, M., Gabrielli, A., Tacchella, A., Caldarelli, G., & Pietronero, L. (2013). Measuring the Intangibles: A Metrics for the Economic Complexity of Countries and Products. *PLOS ONE*, 8(8).
- Durlauf, S. N., & Quah, D. T. (1999). Chapter 4 The new empirics of economic growth. In *Handbook of Macroeconomics* (Vol. 1, pp. 235-308).
- Durlauf, S. N., Johnson, P. A., & Temple, J. R. (2005). Chapter 8 Growth Econometrics. In *Handbook of Economic Growth* (Vol. 1, pp. 555-677).
- Ezcurra, R., Pascual, P., & Rapún, M. (2006). Regional Specialization in the European Union. *Regional Studies*, 40(6), 601-616.
- Fritz, B. S., & Manduca, R. A. (2021). The Economic Complexity of US Metropolitan Areas. *Regional Studies*, 55(7), 1299-1310.
- Hartmann, D., Guevara, M. R., Figueroa, C. J., Aristarán, M., & Hidalgo, C. (2017). Linking Economic Complexity, Institutions, and income inequality. *World Development*, 93, 75-93.
- Hausmann, R., Hidalgo, C., Bustos, S., Coscia, M., Simoes, A., & Yildirim, M. A. (2014). *The Atlas of Economic Complexity: Mapping Paths to Prosperity*. MIT Press, Cambridge, MA.
- Hidalgo, C., & Hausmann, R. (2009). The building blocks of economic complexity. *PNAS*, 10570-10575.
- Hidalgo C. (2015) *Why information grows. The evolution of Order, from Atoms to Economies*. Basic Books.
- Hidalgo, C. (2021). Economic complexity theory and applications. *Nature Reviews Physics*, 3(2), 92-113.
- Ourens, G., (2013). Can the Method of Reflections help predict future growth?, LIDAM Discussion Papers IRES 2013008, Université catholique de Louvain, Institut de Recherches Economiques et Sociales (IRES).
- Petalia, S., Balland, P.A. and Morrison, A. (2017), "Climbing the Ladder of Technological Development," *Research Policy* 46(5), 956 – 969.

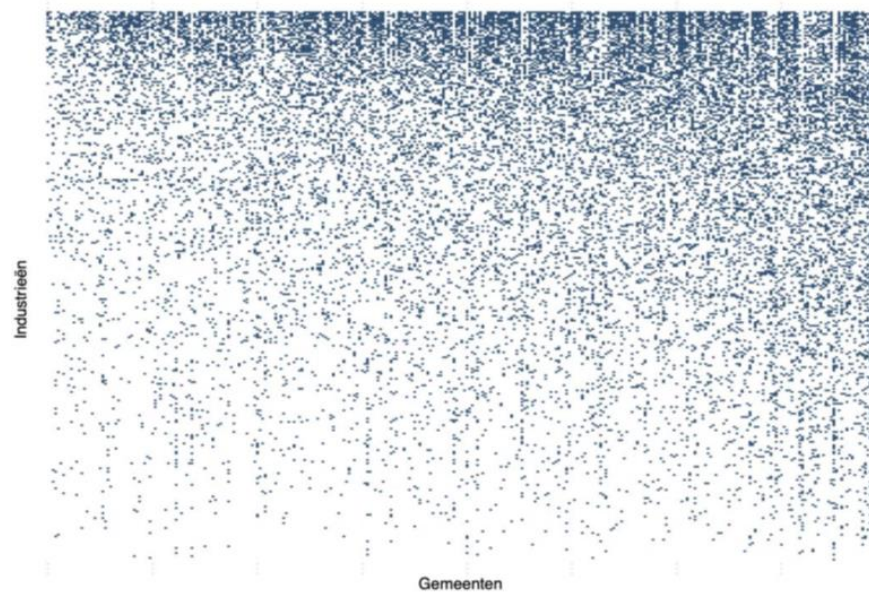
Sbardella , A., Pugliese, E., & Pietrnero, L. (2017). Economic development and wage inequality: A complex system analysis. PLOS ONE, 12(9).

Schillebeeckx, S., (2022) Economische Complexiteit en Welvaart: Een casestudie voor de Vlaamse Gemeenschap, Masterproef, Master in de economische wetenschappen, KU Leuven.

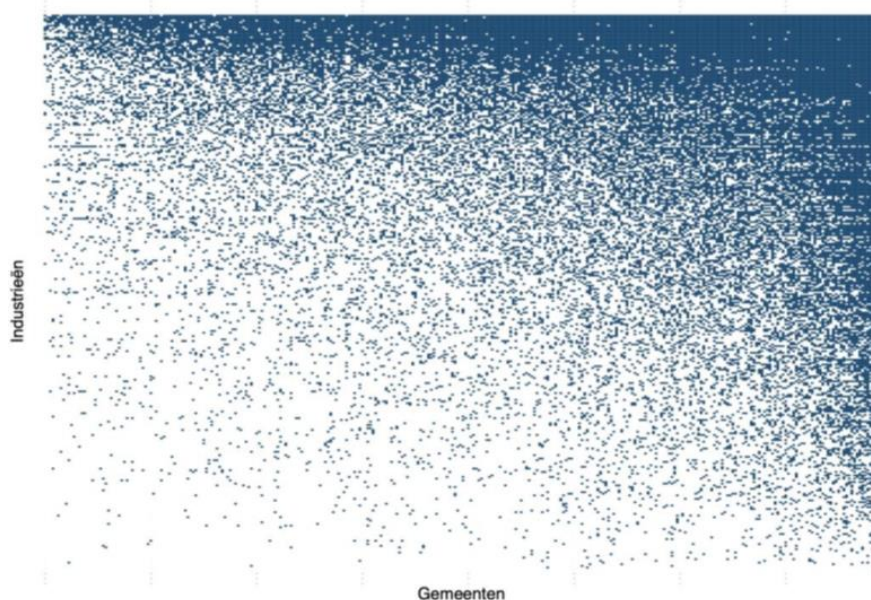
Tacchella, A., Cristelli, M., Caldarelli , G., Gabrieli, A., & Pietronero, L. (2012). A New Metrics for a Countries 'Fitness and Products' Complexity. Scientific Reports, 2(1).

Appendix

Inputmatrices: Grafische weergave van de verschillen in diversiteit en ubiquiteit wanneer er verschillende tewerkstellingsvoorwaarden worden gehanteerd bij het bepalen welke industrieën er worden meegenomen in de analyse.



*Figuur A: Inputmatrix zonder extra tewerkstellingsvoorwaarden (enkel industrieën met locatie quotiënt groter dan 1 worden meegenomen)
Bron: eigen berekening o.b.v. RSZ data*



*Figuur B: Inputmatrix met tewerkstellingsvoorwaarde van 1 persoon (elke industrie waarin tenminste 1 persoon tewerkgesteld is, wordt meegenomen)
Bron: eigen berekening o.b.v. RSZ data*