Center for Collective Learning

centerforcollectivelearning.org

Economic Complexity: How AI is helping us Understand Sustainable Economic Development

César A. Hidalgo
Center for Collective Learning,
Corvinus University of Budapest
Toulouse School of Economics
Manchester University

Thomas Thwaites

The world works not because a few people know a lot, but because many people know a little.

Economic complexity is about understanding how that knowledge comes together.

Economic complexity

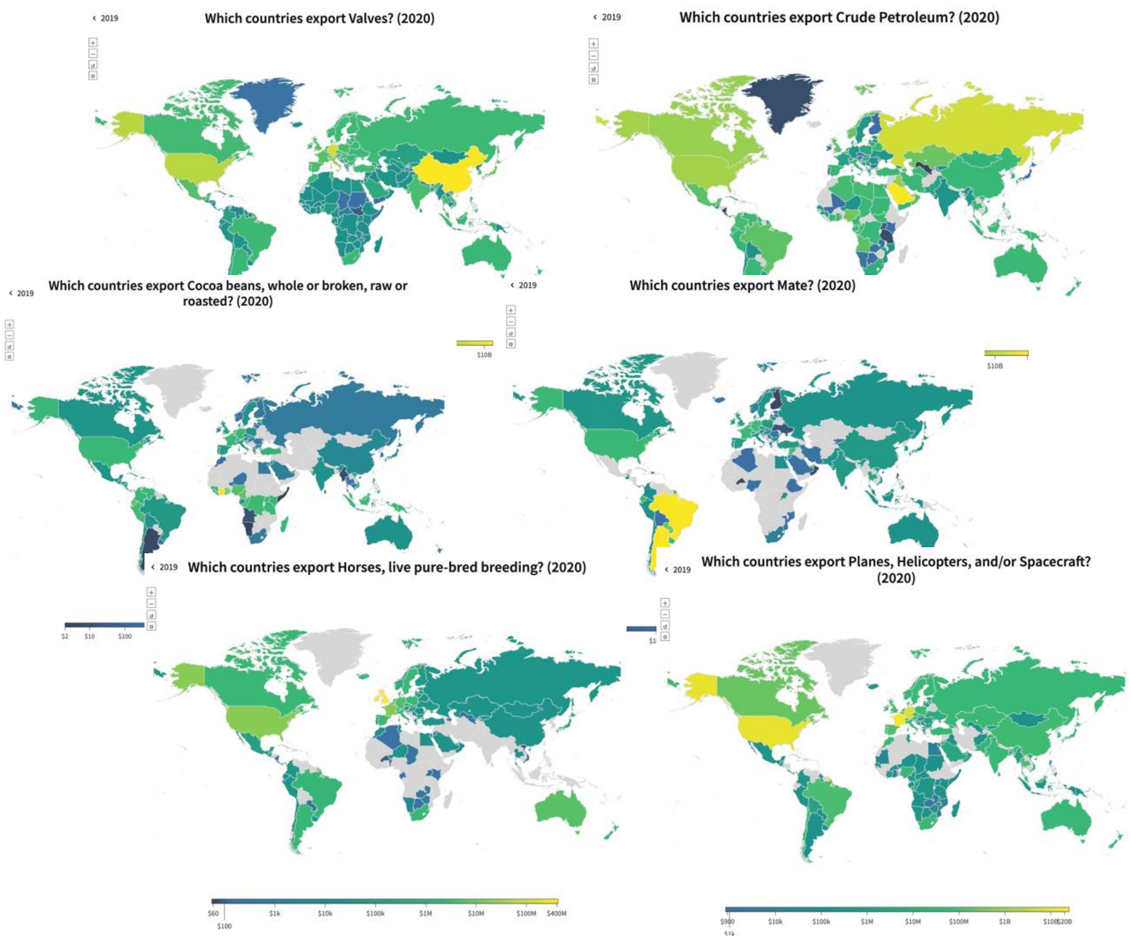
machine learning

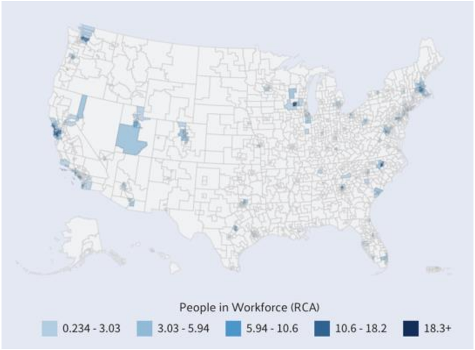
+

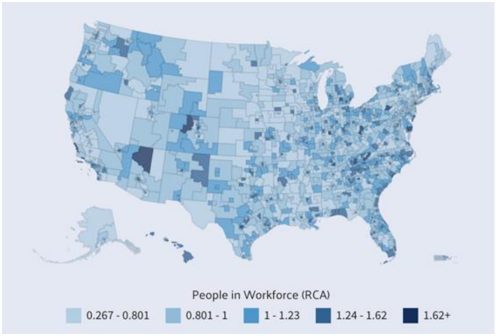
economic data

=

development outcomes



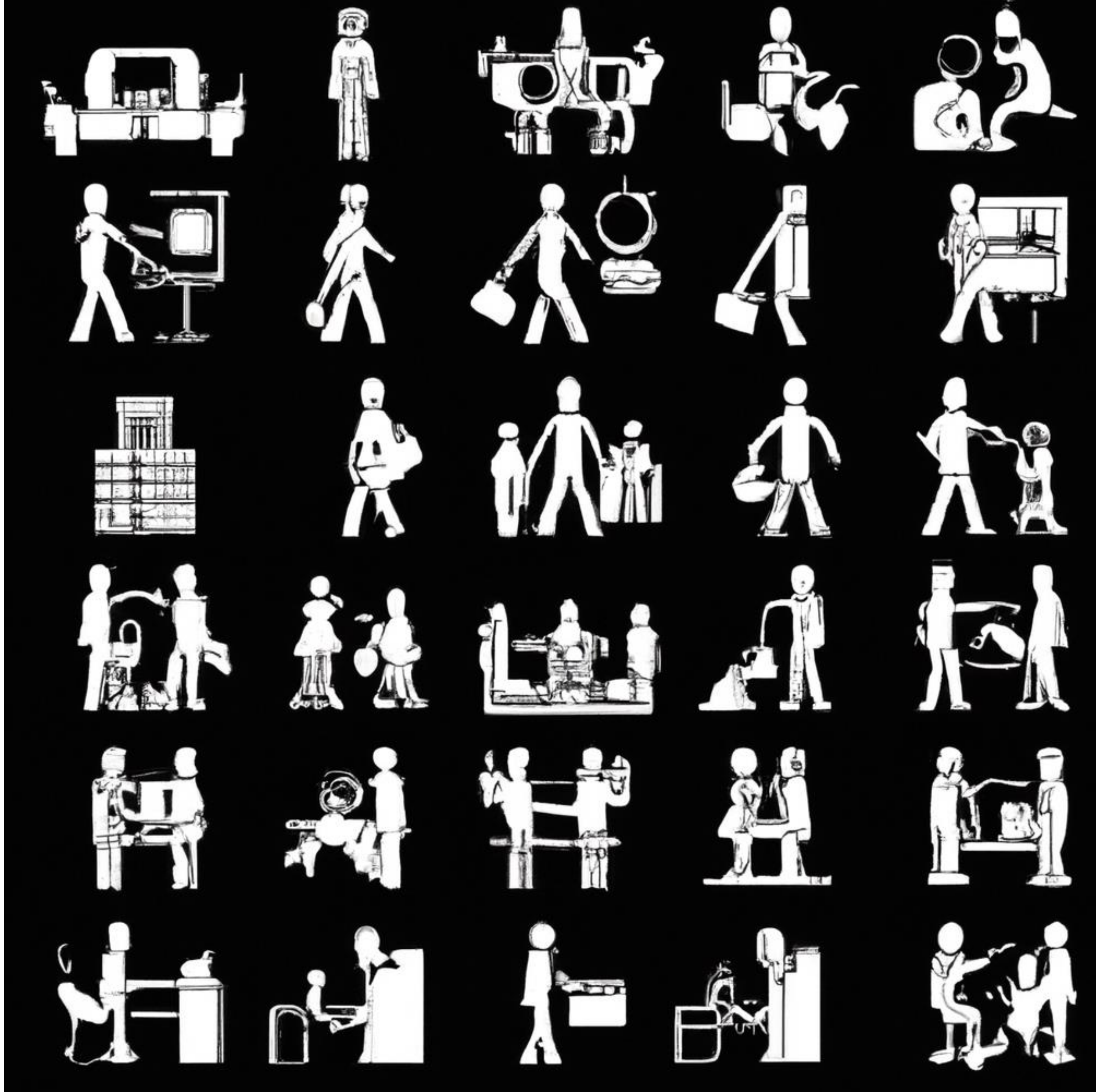




Starting in 2006-2007

Why Machine Learning

Because factors of production, and in particular **knowledge**, are highly specific and **non-fungible** (not interchangeable).



chat.openai.com/chat

+ New chat

Phone models discussed.

AI Methods in Economics

Digital trade importance.

RQ: Papers, Ideas / T: Suggest

AI and Economics Revolution

Recommend TV Shows

New chat

New chat

Clear conversations

Upgrade to Plus NEW

Dark mode

Updates & FAQ

Log out

ChatGPT

Examples	Capabilities	Limitations
"Explain quantum computing in simple terms" →	Remembers what user said earlier in the conversation	May occasionally generate incorrect information
"Got any creative ideas for a 10 year old's birthday?" →	Allows user to provide follow-up corrections	May occasionally produce harmful instructions or biased content
"How do I make an HTTP request in Javascript?" →	Trained to decline inappropriate requests	Limited knowledge of world and events after 2021

ChatGPT Feb 13 Version. Free Research Preview. Our goal is to make AI systems more natural and safe to interact with. Your feedback will help us improve.

The best thing about AI is its ability to

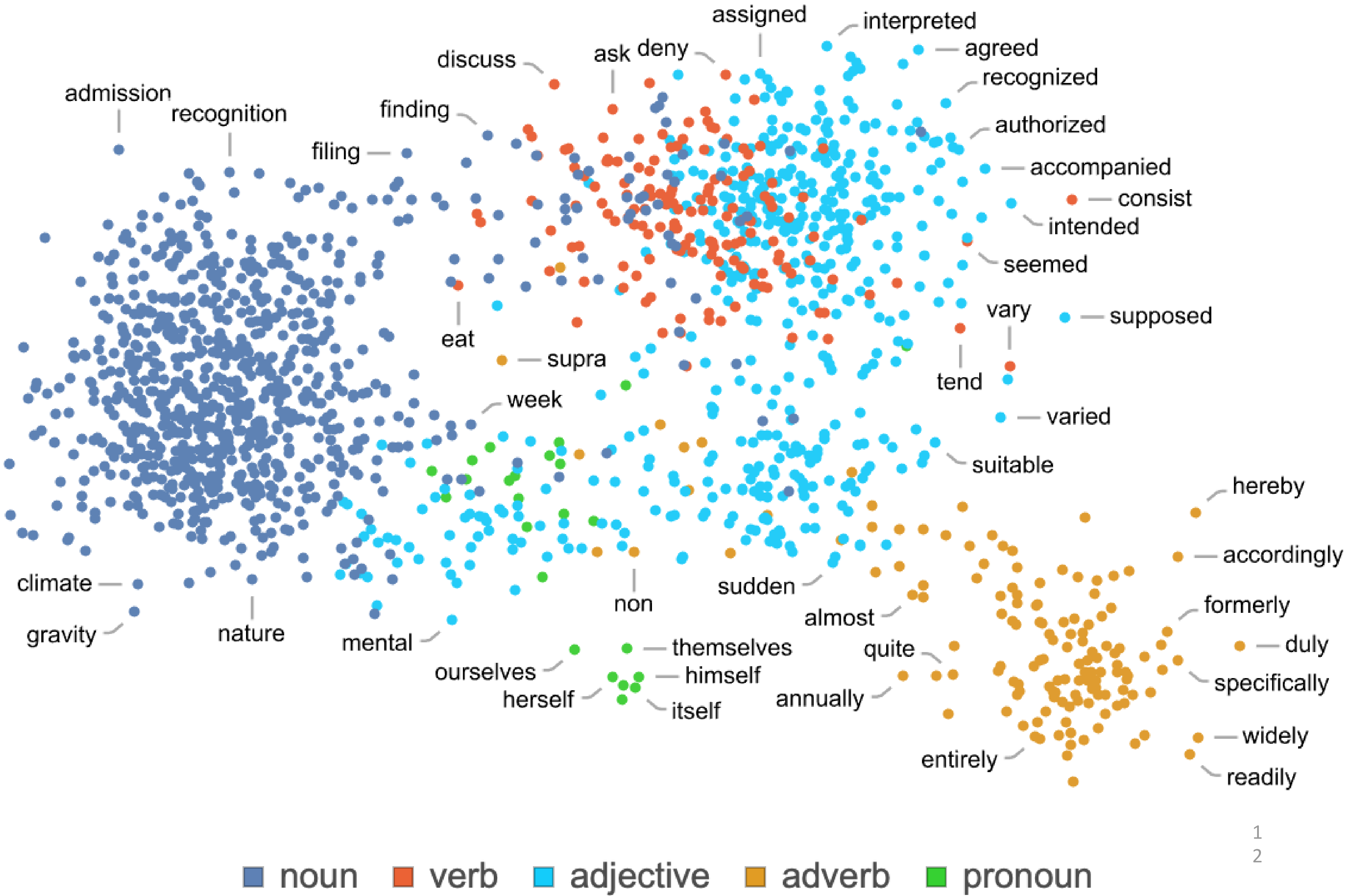
learn	4.5%
predict	3.5%
make	3.2%
understand	3.1%
do	2.9%

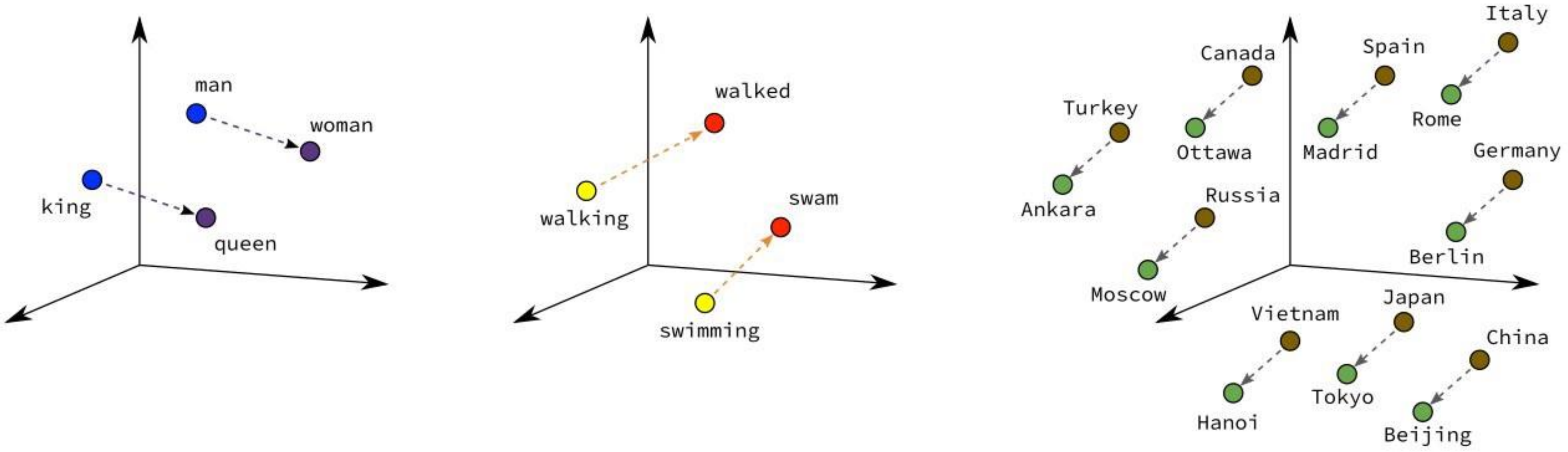
https://writings.stephenwolfram.com/2023/02/what-is-chatgpt-doing-and-why-does-it-work/

Verb, Nouns, Adjectives, and Adverbs List

Verbs	Nouns	Adjectives	Adverbs
accuse	accusation	accusing	accusingly
argue	argument	arguable	arguably
characterize	character	characteristic	characteristically
condition	condition	conditional	conditionally
darken	dark, darkness	dark, darkened	darkly
destroy	destruction	destructive	destructively
drink	drink, drunkenness	drunk, drunken	drunkenly

Word Embeddings Provide Semantic Representations That Transcend Parts of Speech Grammar

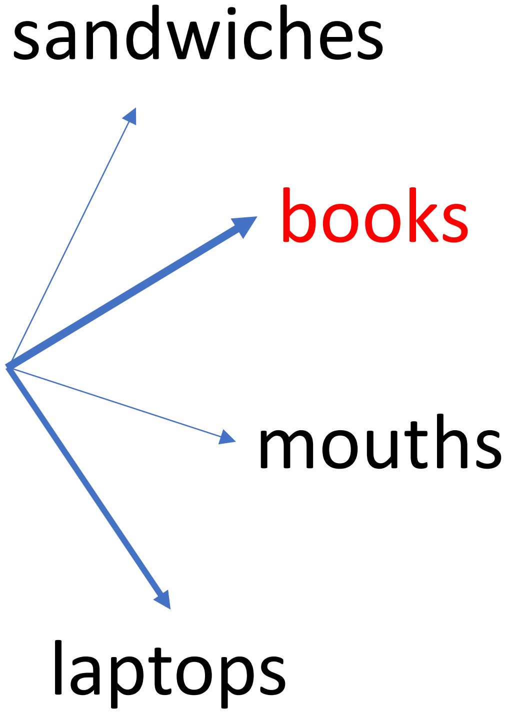




Male-Female Verb Tense Country-Capital

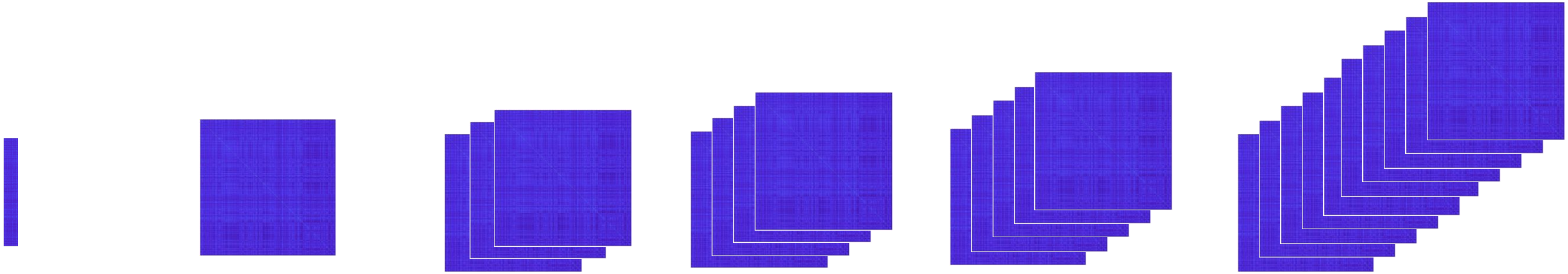
Attention!

At the **library**, the **students** opened their….



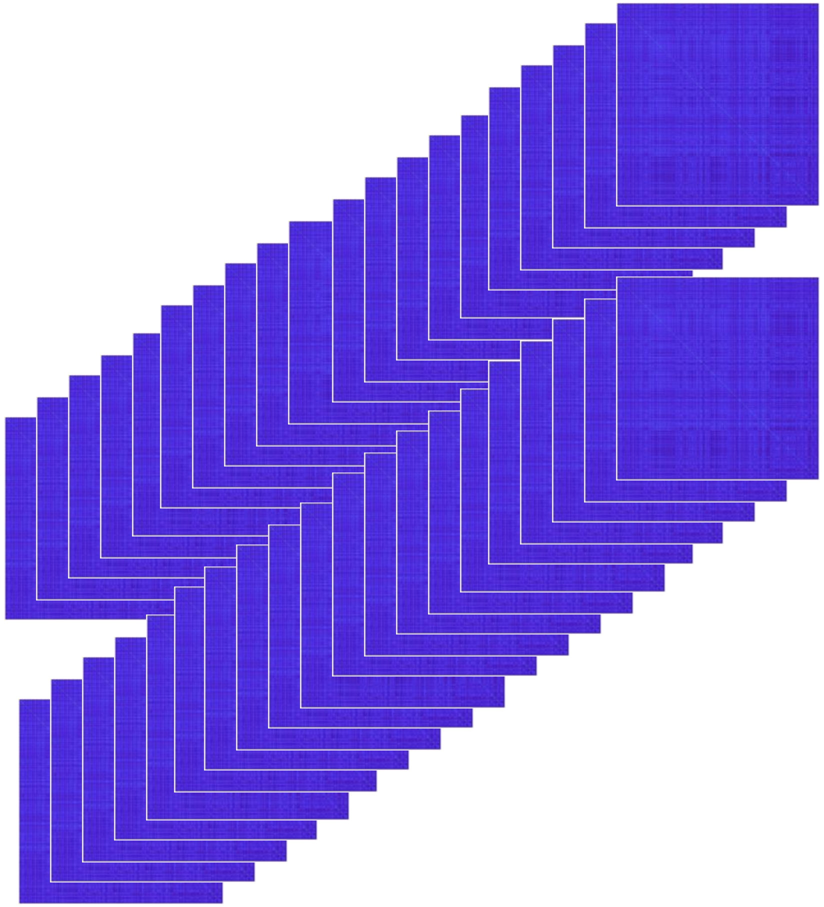
Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. **"Attention is all you need."** *Advances in neural information processing systems* 30 (2017).

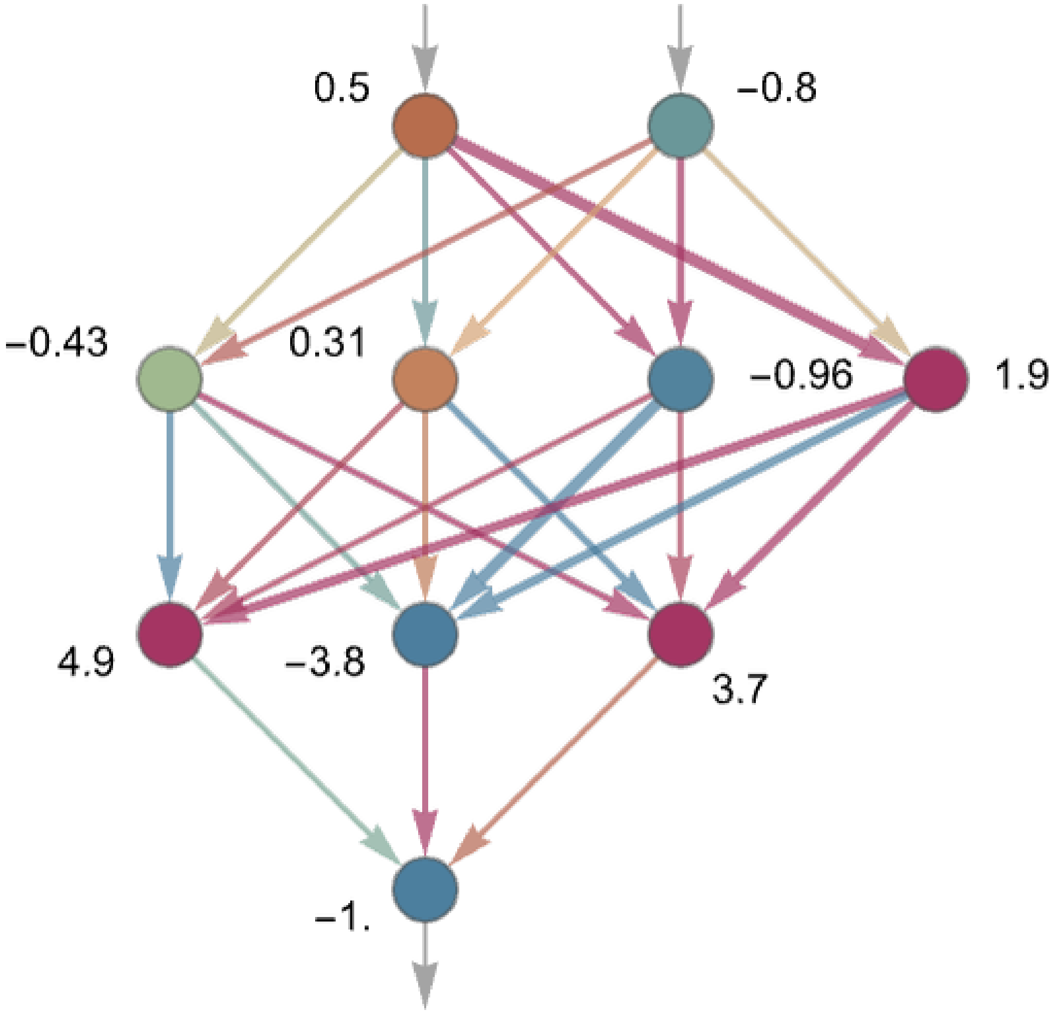
It is a BIG problem!



Words	Bigrams	Trigrams	4-grams	5-grams		20-grams
10,000 words	100 million pairs	1 trillion triplets	10,000 trillion 4-grams	100 million trillion 5-grams	…	20-grams
10^4	10^8	10^{12}	10^{16}	10^{20}		10^{80}

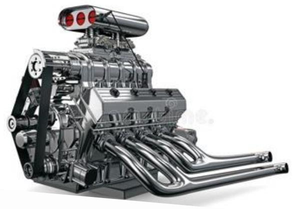
Use neural networks to approximate these functions





With only a “few“ billion parameters

“Parts of Speech” Economics



Manufacture | Capital Intensive

Agriculture | Capital Intensive

Agriculture | Labor Intensive

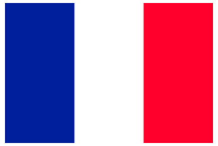
Manufacture | Labor Intensive

NLP, LLMs

Just like we can count the number of words in each sentence or paragraph, and their co-occurrences, to create representations of their semantic meaning, we can count the number of economic activities that are present across cities, regions, and countries to create representations of the knowledge embedded in them.

Economic Complexity

Spark Ignition Engines, Tobacco, Engine Parts, Aircraft Parts, Vaccines, Plywood, Tractors, Coffee, Frozen Bovine Meat, etc…



Spark Ignition Engines, Engine Parts, Aircraft Parts, Aircraft, Wheat, Wine, Perfumes, Vaccines, etc…

Crude Petroleum, Refined Petroleum, Petroleum Gases, Wheat, Aircraft Parts, etc.

But……
Who cares about Economic Complexity?

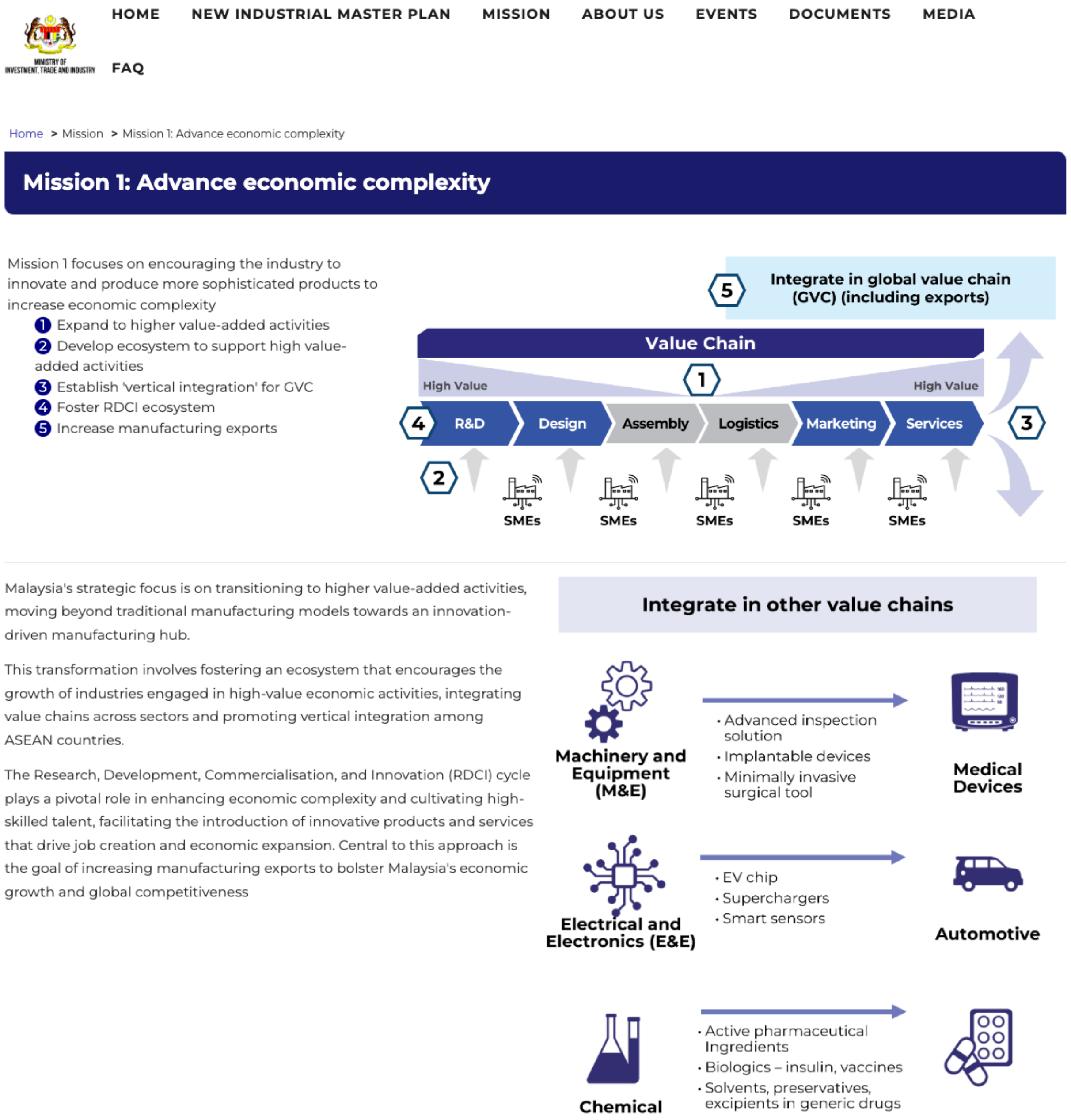
COMPLEXITY ANALYSIS STUDY OF MALAYSIA'S MANUFACTURING INDUSTRIES

2014 || FINAL REPORT

For further information, please contact:

Director General
ECONOMIC PLANNING UNIT
Prime Minister's Department
Block B5 & B6, Complex B
Federal Government Administrative Centre
62502 PUTRAJAYA

www.epu.gov.my



Mission 1: Advance economic complexity

Mission 1 focuses on encouraging the industry to innovate and produce more sophisticated products to increase economic complexity

1. Expand to higher value-added activities
2. Develop ecosystem to support high value-added activities
3. Establish 'vertical integration' for GVC
4. Foster RDCI ecosystem
5. Increase manufacturing exports

Malaysia's strategic focus is on transitioning to higher value-added activities, moving beyond traditional manufacturing models towards an innovation-driven manufacturing hub.

This transformation involves fostering an ecosystem that encourages the growth of industries engaged in high-value economic activities, integrating value chains across sectors and promoting vertical integration among ASEAN countries.

The Research, Development, Commercialisation, and Innovation (RDCI) cycle plays a pivotal role in enhancing economic complexity and cultivating high-skilled talent, facilitating the introduction of innovative products and services that drive job creation and economic expansion. Central to this approach is the goal of increasing manufacturing exports to bolster Malaysia's economic growth and global competitiveness


Kingdom of Saudi Arabia
VIZBUILDER
EN | AR
DATA SAUDI
بيانات السعودية
وزارة الاقتصاد والتخطيط
MINISTRY OF ECONOMY & PLANNING
Kingdom of Saudi Arabia
A unified platform to present and analyze the latest economic and social data for the Kingdom and its regions in visually interactive ways that facilitate understanding of the Saudi economic landscape.
EXPLORE REGIONS ON THE MAP
Explore our reports
KINGDOM

GOBIERNO DE MÉXICO

Trámites Gobierno

Data MÉXICO

Vizbuilder ES | EN

Data MÉXICO

EXPLORA, VISUALIZA, COMPARA, Y DESCARGA DATOS MEXICANOS

ECONOMÍA
SECRETARÍA DE ECONOMÍA

¡Más de 12.000 perfiles para descubrir!

Ej. Ciudad de México, Monterrey BUSCAR

¿Qué es DataMéxico?

DataMéxico permite la integración, visualización y análisis de datos públicos para fomentar la innovación, inclusión y diversificación de la economía mexicana.

PERFILES

Explore México mediante datos económicos, sociales y ocupacionales a través de visualizaciones interactivas personalizables.

COMPLEJIDAD ECONÓMICA

Conozca el nivel de desarrollo industrial y económico en México a múltiples niveles geográficos.

VIZ BUILDER

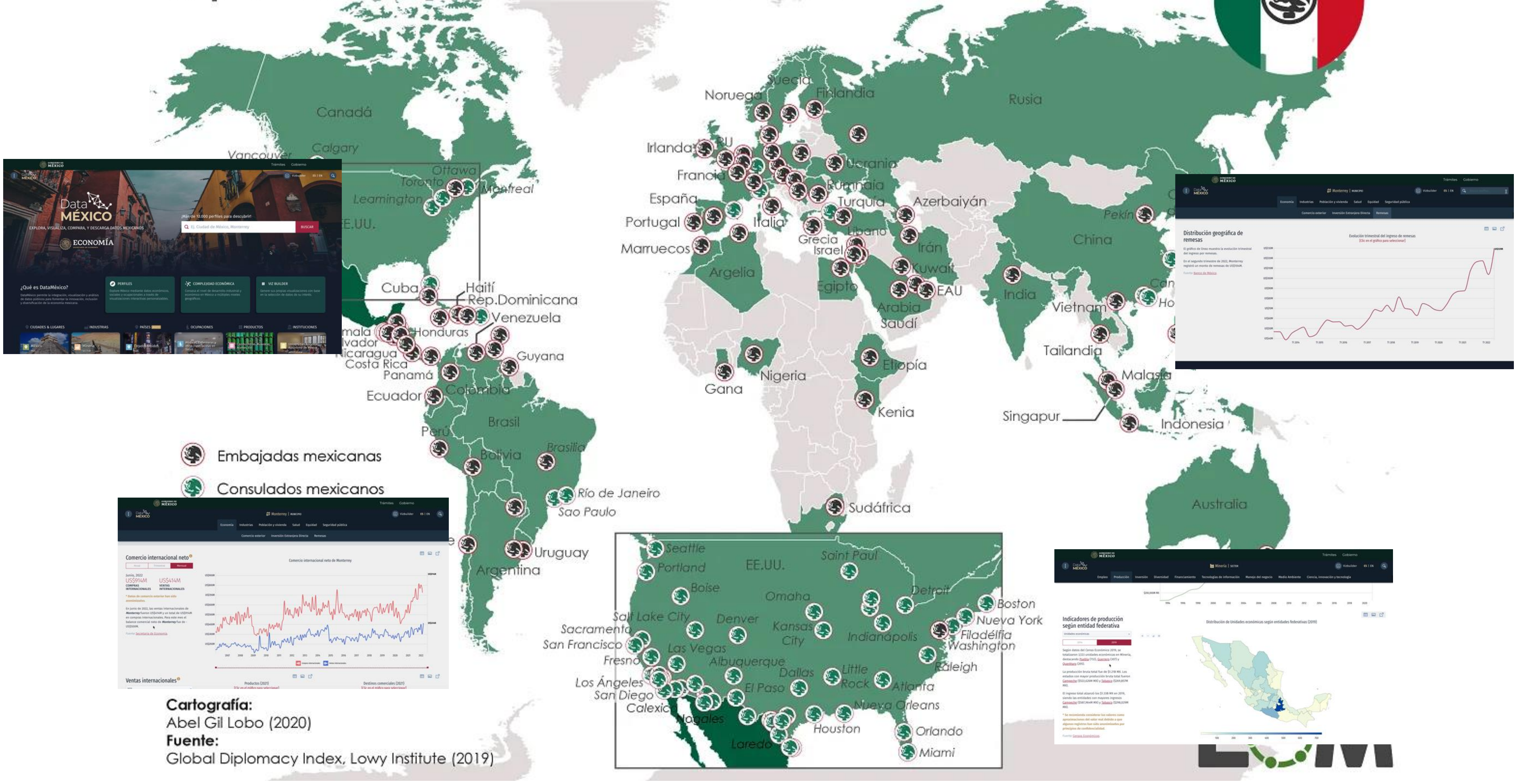
Genere sus propias visualizaciones con base en la selección de datos de su interés.

CIUDADES & LUGARES | INDUSTRIAS | PAÍSES ¡NUEVO! | OCUPACIONES | PRODUCTOS | INSTITUCIONES

México | Minería | Estados Unidos | Médicos, Enfermeras y otros Especialistas en Salud | Combustibles Minerales, Aceites, Etc | Universidad Nacional Autónoma De México

La diplomacia de México

Representacion mexicana en el mundo



Embajadas mexicanas

Consulados mexicanos

Cartografía:
Abel Gil Lobo (2020)
Fuente:
Global Diplomacy Index, Lowy Institute (2019)

OEC

COTEC

LA COMPLEJIDAD ECONÓMICA EN ESPAÑA

DataMÉXICO

statistics estonia

Observatorio Institucional

HEALTHY Communities NC

DATA USA

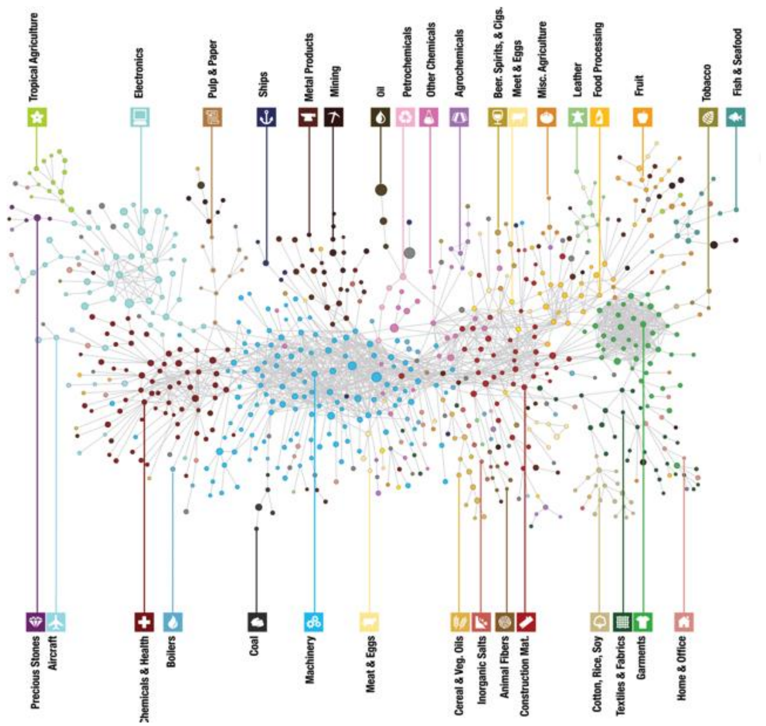
ITP Producción

DATA AFRICA

DataChile

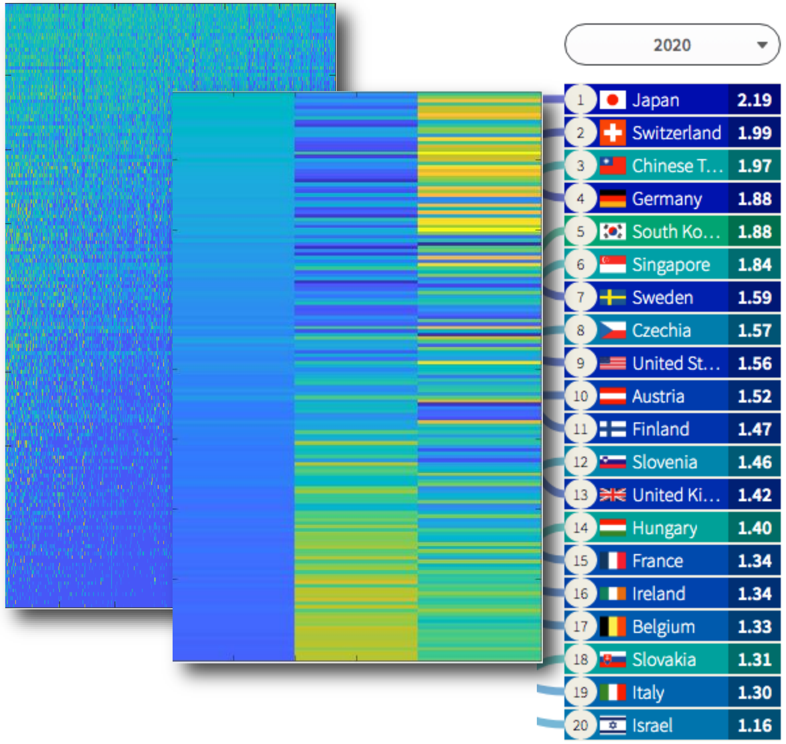
Two Basic Ideas in Economic Complexity

Relatedness



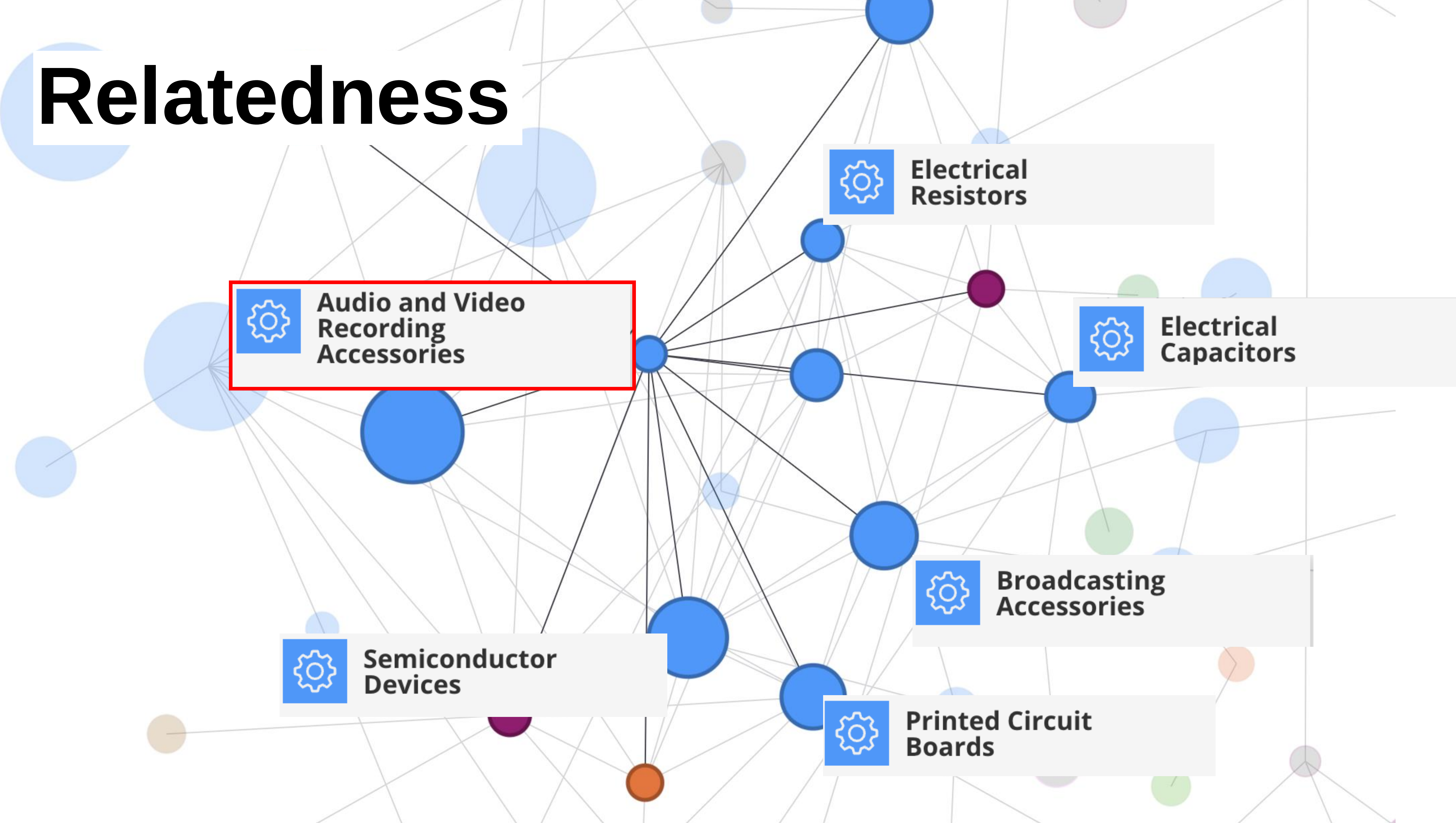
Hidalgo et al. Science (2007)

Complexity Indexes



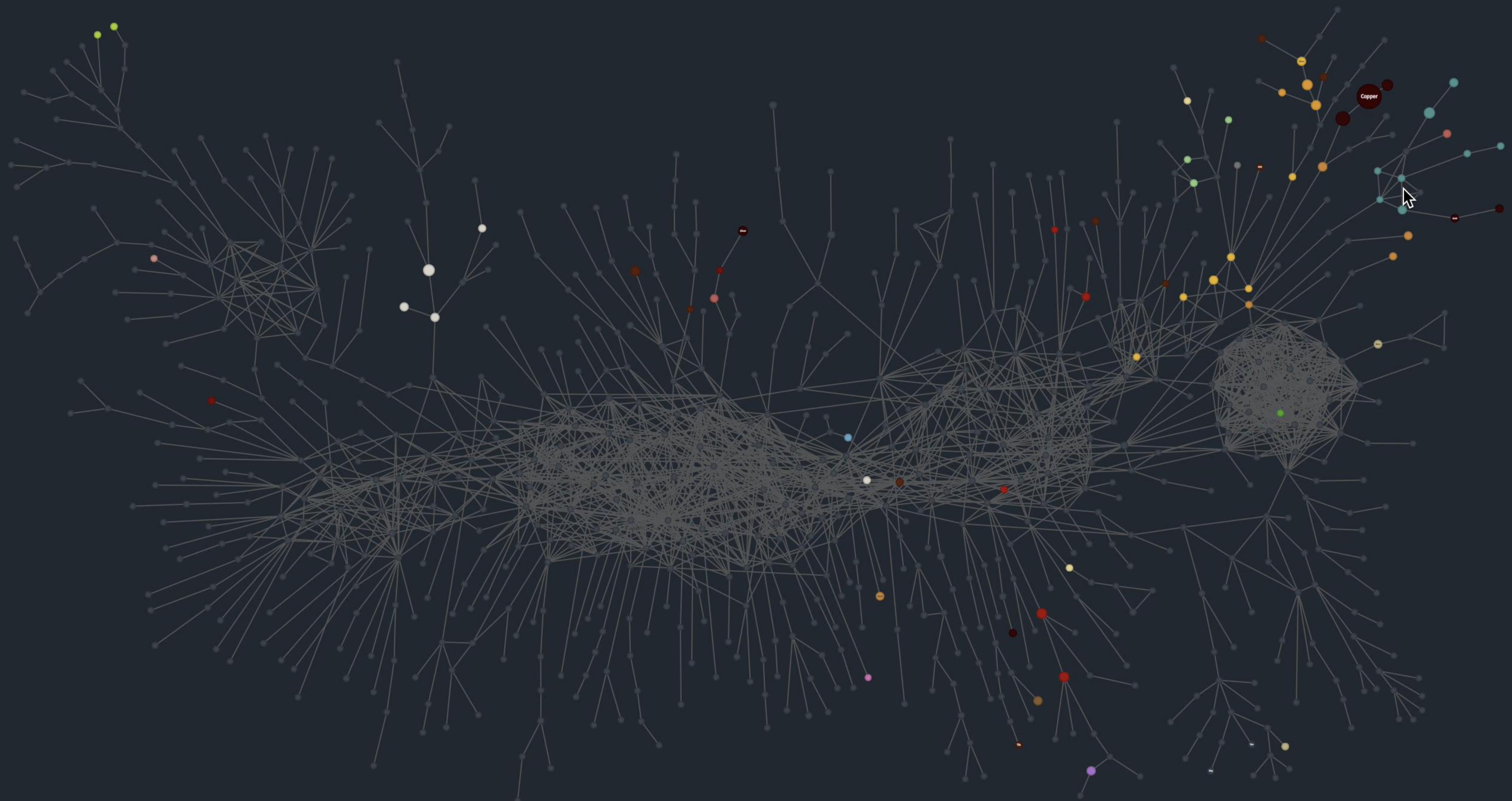
Hidalgo & Hausmann. PNAS (2009)

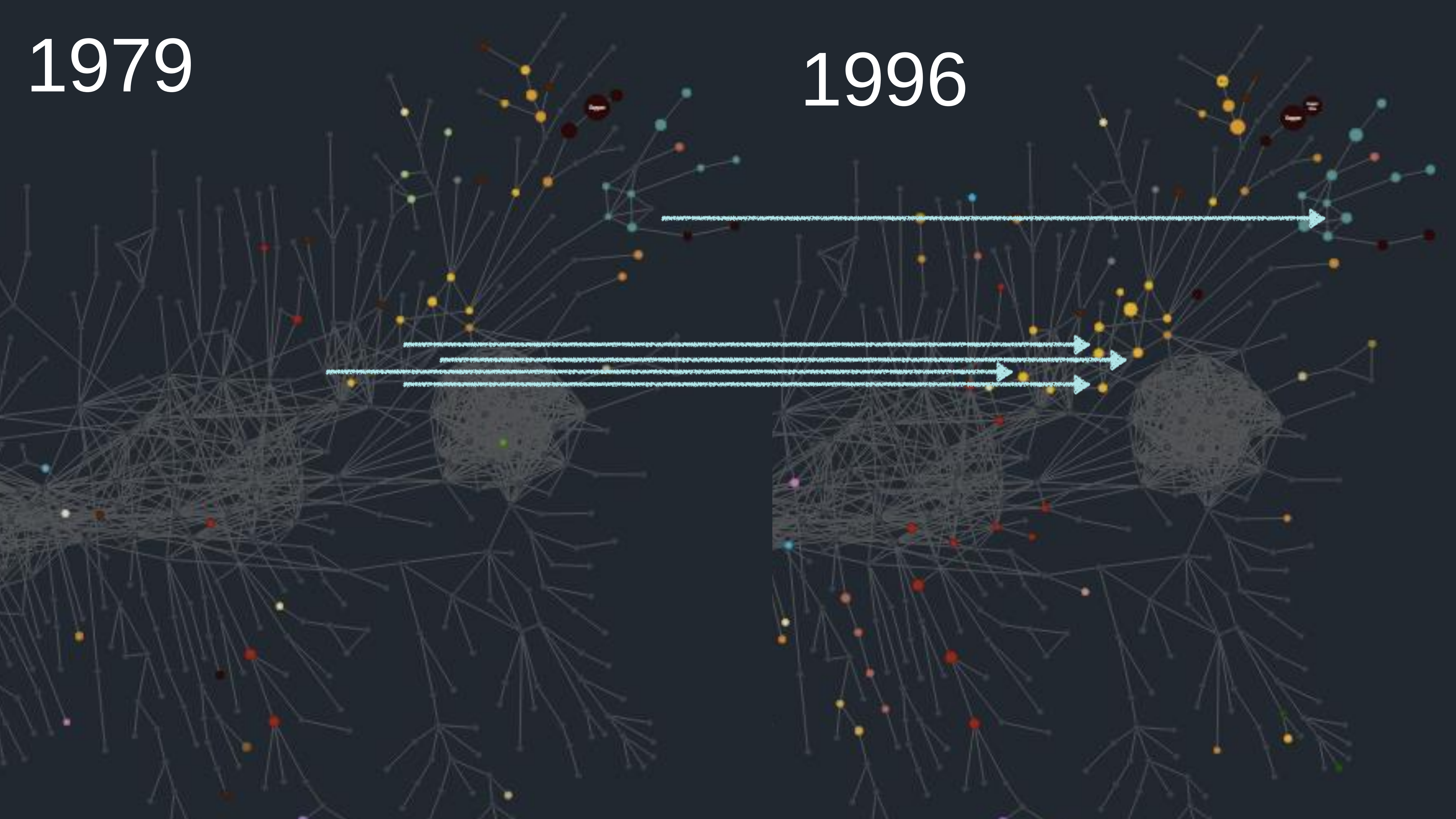
Relatedness



What are the export opportunities of Chile? (1979)

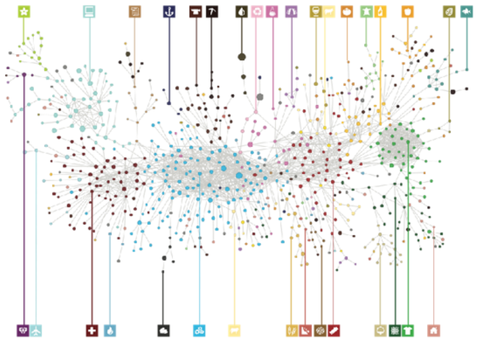
TOTAL: $3.67B



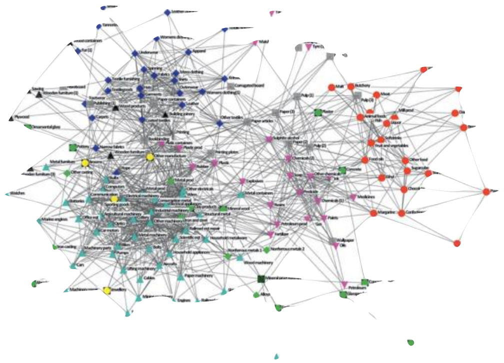

1979
1996

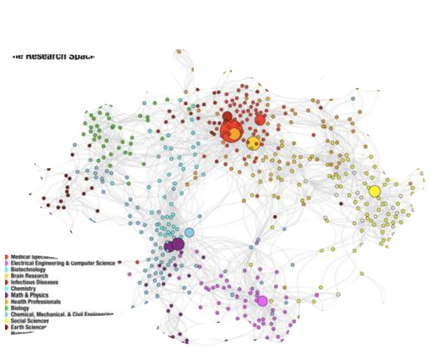
THE PRINCIPLE OF RELATEDNESS

Products

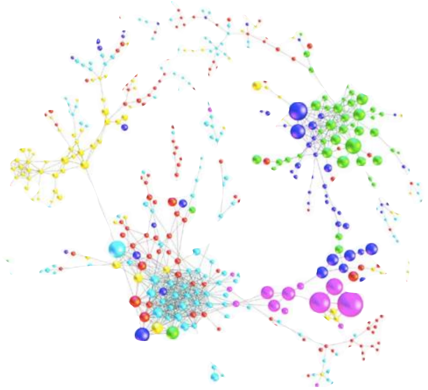


Industries





Research Areas



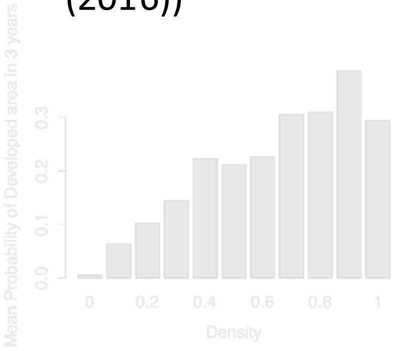
Patents

Products

(Hidalgo et al 2007)

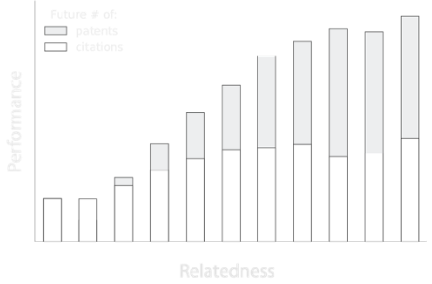
Research Areas

(Guevara et al. (2016))



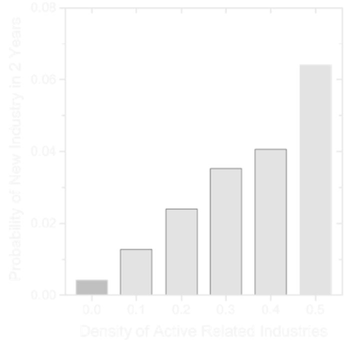
Patents

(Kogler et al. (2013), Boschma et al. (2015), Alstott et al. (2016))

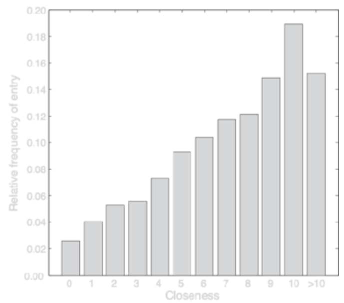


Industries

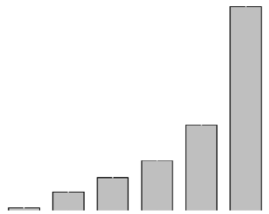
BRAZIL
(Jara-Figueroa et al. 2018)



SWEDEN:
(Neffke, Henning, Boschma 2011)



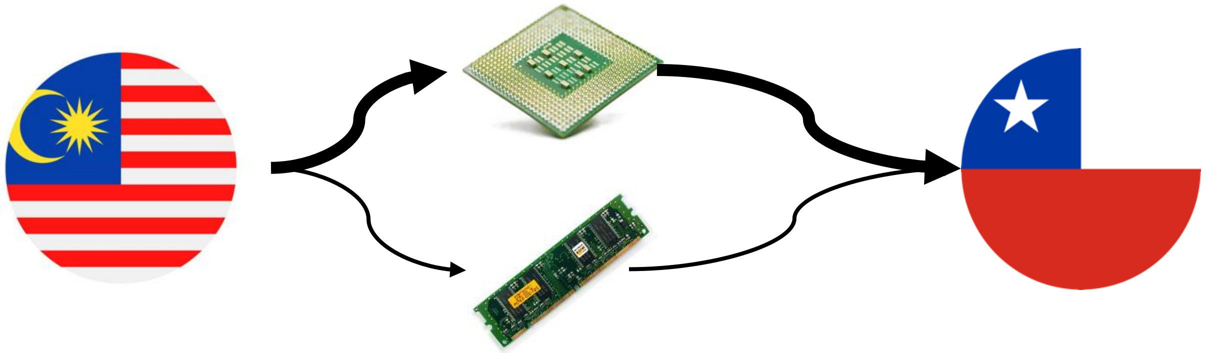
CHINA:
(He et al. 2017
Gao et al. 2021)



César A. Hidalgo, Pierre-Alexandre Balland, Ron Boschma, Mercedes Delgado, Maryann Feldman, Koen Frenken, Edward Glaeser, Dieter F. Kogler, Andrea Morrison, Frank Neffke, David Rigby, Scott Stern, Siqi Zheng, Canfei He, & Shengjun Zhu. "The Principle of Relatedness" CCS 2018. Springer Proceedings in Complexity. Springer (2018)

Bilateral Relatedness

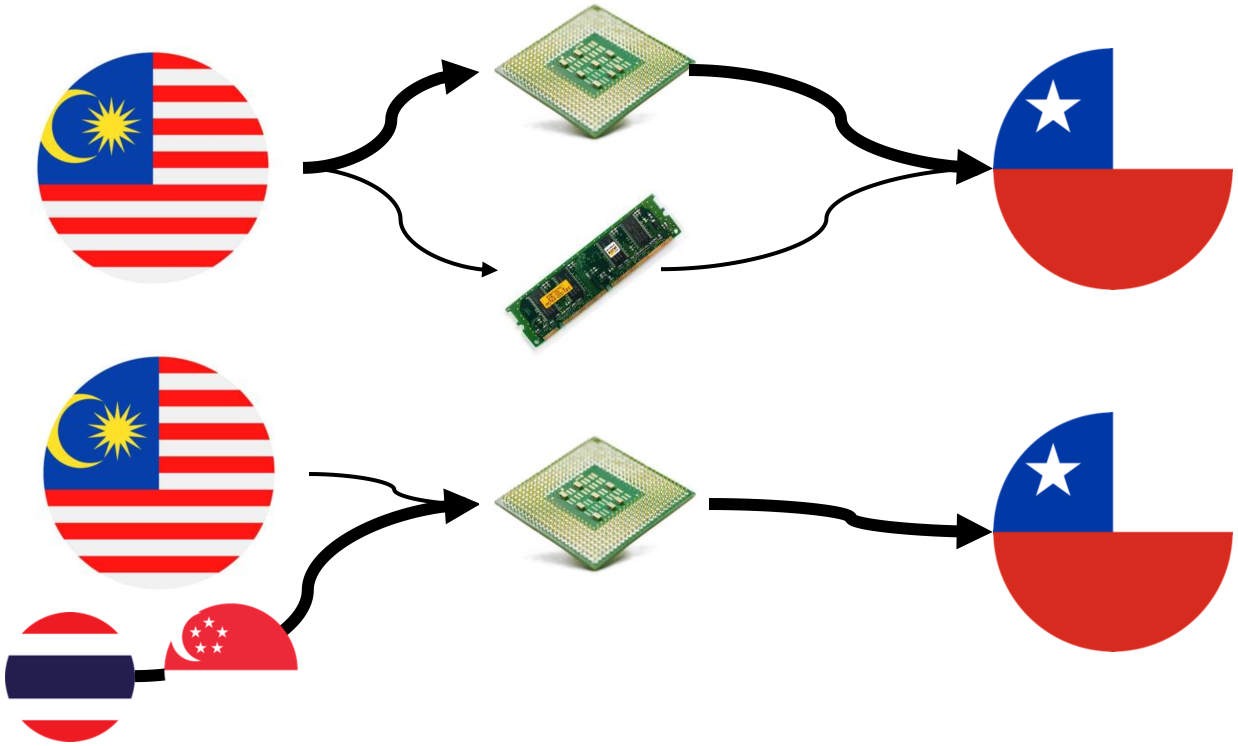
Product Relatedness:
Enter same market with similar product.



Jun, Bogang, Aamena Alshamsi, Jian Gao, and César A. Hidalgo. "Bilateral relatedness: knowledge diffusion and the evolution of bilateral trade." *Journal of Evolutionary Economics* 30 (2020): 247-277.

Bilateral Relatedness

Product Relatedness:
Enter same market with similar product.



Exporter Relatedness:
Enter same market as geographic neighbor.

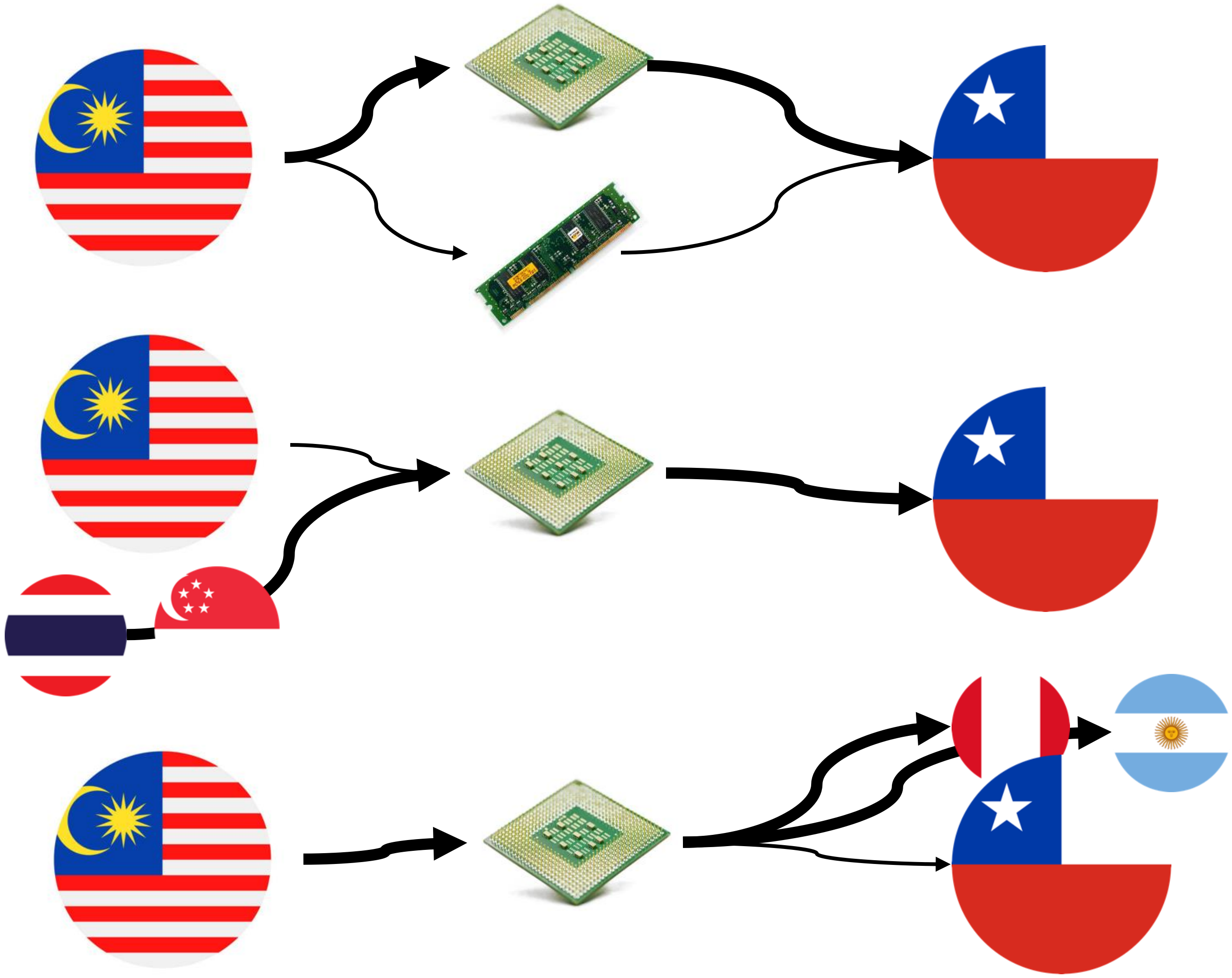
Jun, Bogang, Aamena Alshamsi, Jian Gao, and César A. Hidalgo. "Bilateral relatedness: knowledge diffusion and the evolution of bilateral trade." *Journal of Evolutionary Economics* 30 (2020): 247-277.

Bilateral Relatedness

Product Relatedness: Enter same market with similar product.

Exporter Relatedness: Enter same market as geographic neighbor.

Importer Relatedness: Enter the geographic neighbor of a current market



Jun, Bogang, Aamena Alshamsi, Jian Gao, and César A. Hidalgo. "Bilateral relatedness: knowledge diffusion and the evolution of bilateral trade." *Journal of Evolutionary Economics* 30 (2020): 247-277.

Potential Exports

TYPE

Product | **Country**

COUNTRY

GERMANY

FLOW

Exports | Imports

DEPTH

Section | HS2 | **HS4**

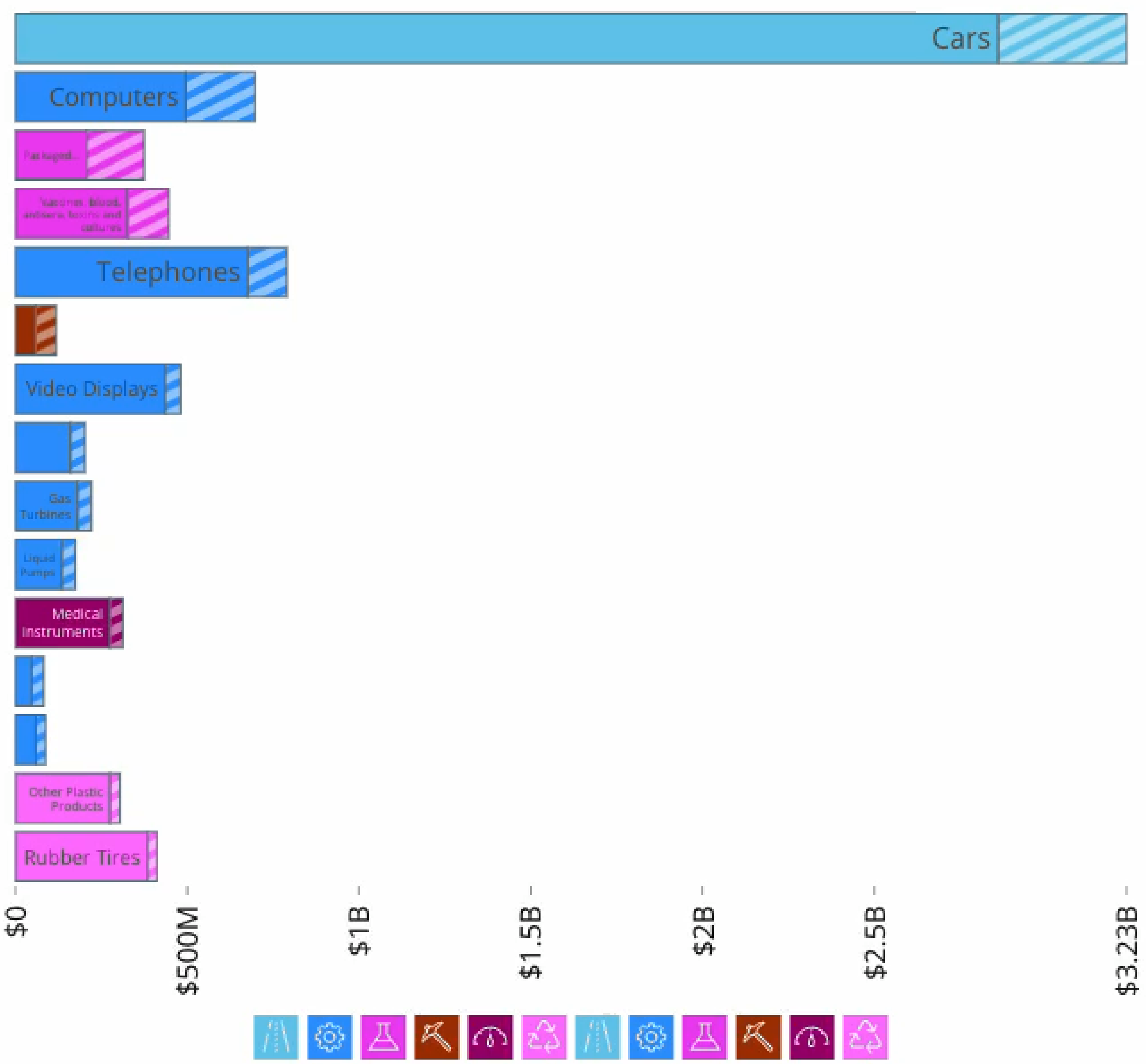
VIEW

Top 10 | **Top 15** | Top 20

SORT BY

Highest Potential | Most Saturated

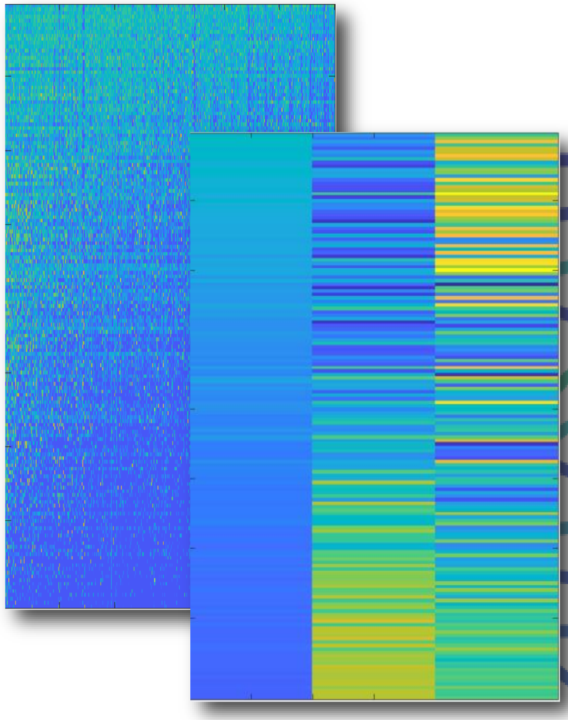
Exports with highest growth potential (Hungary - Germany)

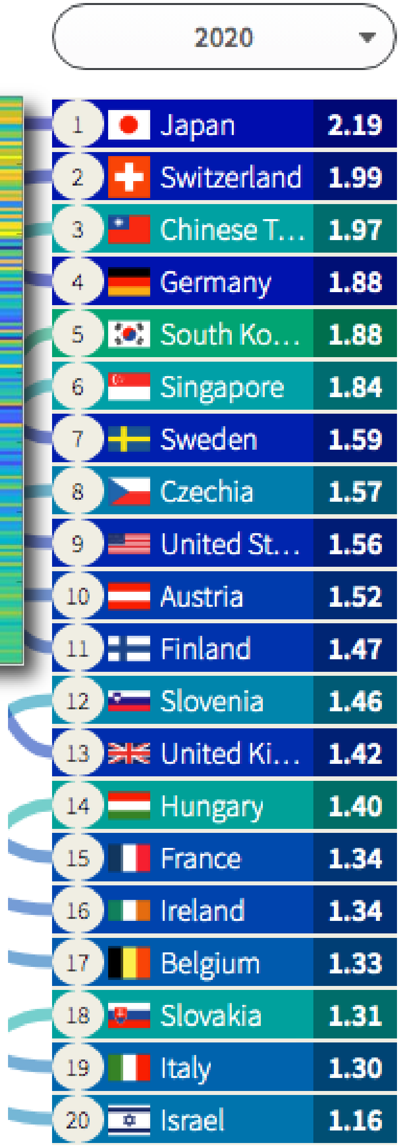


oec.world

Economic Complexity

The use of dimensionality reduction techniques (e.g. SVD) to summarize the sophistication of productive structures.





Economic Complexity Explains

Economic Growth

Hidalgo and Hausmann, 2009; Chávez et al., 2017; Domini, 2019; Hausmann et al., 2014; Koch, 2021; Lo Turco and Maggioni, 2020; Ourens, 2012; Stojkoski et al., 2016

Inequality

Hartmann et al., 2017,Barza et al., 2020; Ben Saâd and Assoumou-Ella, 2019; Chu and Hoang, 2020; Fawaz and Rahnama-Moghadamm, 2019

Emissions

Can and Gozgor, 2017; Dordmond et al., 2020; Fraccascia et al., 2018; Hamwey et al., 2013; Lapatinas et al., 2019; Mealy and Teytelboym, 2020; Neagu, 2019; Romero and Gramkow, 2021

Economic Complexity

Knowledge of a place
is the knowledge of the activities
present in it

$$K_c = f(M_{cp}, K_p),$$

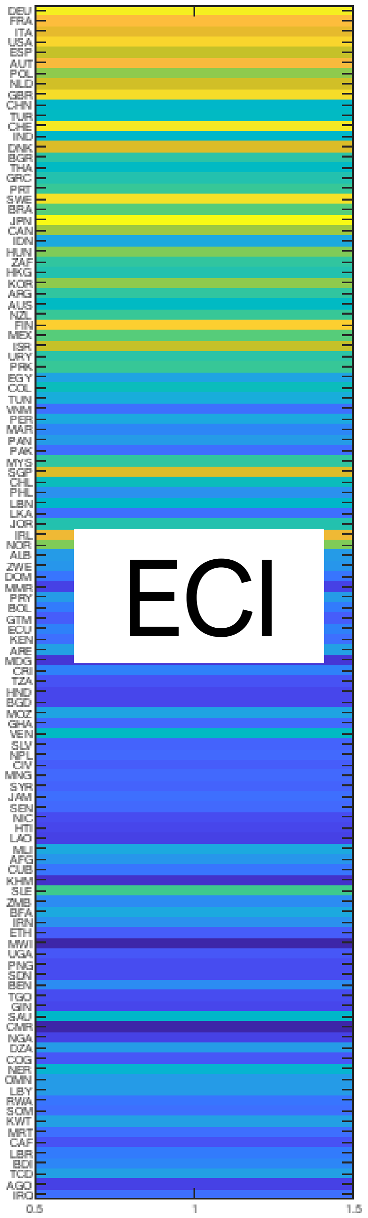
Knowledge of an activity
is the knowledge of the places
where it is present

$$K_p = g(M_{cp}, K_c),$$

$$K_c = f\left(M_{cp}, g(M_{cp}, K_c)\right),$$

$$\widetilde{M}_{cc'} K_{c'} = \lambda K_c$$

Knowledge can be estimated
as the solution to a linear
eigenproblem



Hidalgo CA, Nature Review Physics (2021), PNAS (2009)

When f and g are defined as simple averages….

$$K_c = \frac{1}{M_c}\sum_p M_{cp} K_p$$

$$K_p = \frac{1}{M_p}\sum_c M_{cp} K_{c'}$$

$$\widetilde{M}_{cc'} = \frac{1}{M_c}\sum_p \frac{M_{cp} M_{c'p}}{M_p}$$

The "easy way" to estimate Kc and Kp is to simply iterate the mapping, starting with Kp=Mp and Kc=Mc. The mapping converges after about 20 iterations.

Hidalgo CA, Nature Review Physics (2021), PNAS (2009)

The Economic Complexity Index as an optimization problem

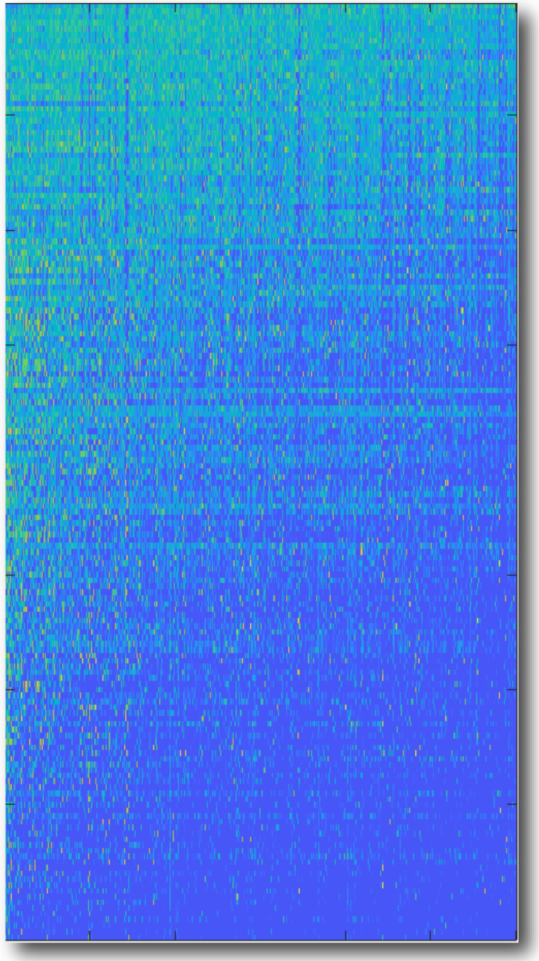
Let V_i be a vector whose entries describe a location (e.g. country) or an activity (e.g. product). Let A_{ij} be a matrix connecting locations and activities. Then ECI is a solution that minimizes the cost function U:

$$U(\vec{V}) = \frac{1}{4} \sum_{i,j} A_{ij} (V_i - V_j)^2$$

Servedio, Vito DP, Alessandro Bellina, Emanuele Calò, and Giordano De Marzo. "Economic Complexity in Mono-Partite Networks." *arXiv preprint arXiv:2405.04158* (2024).

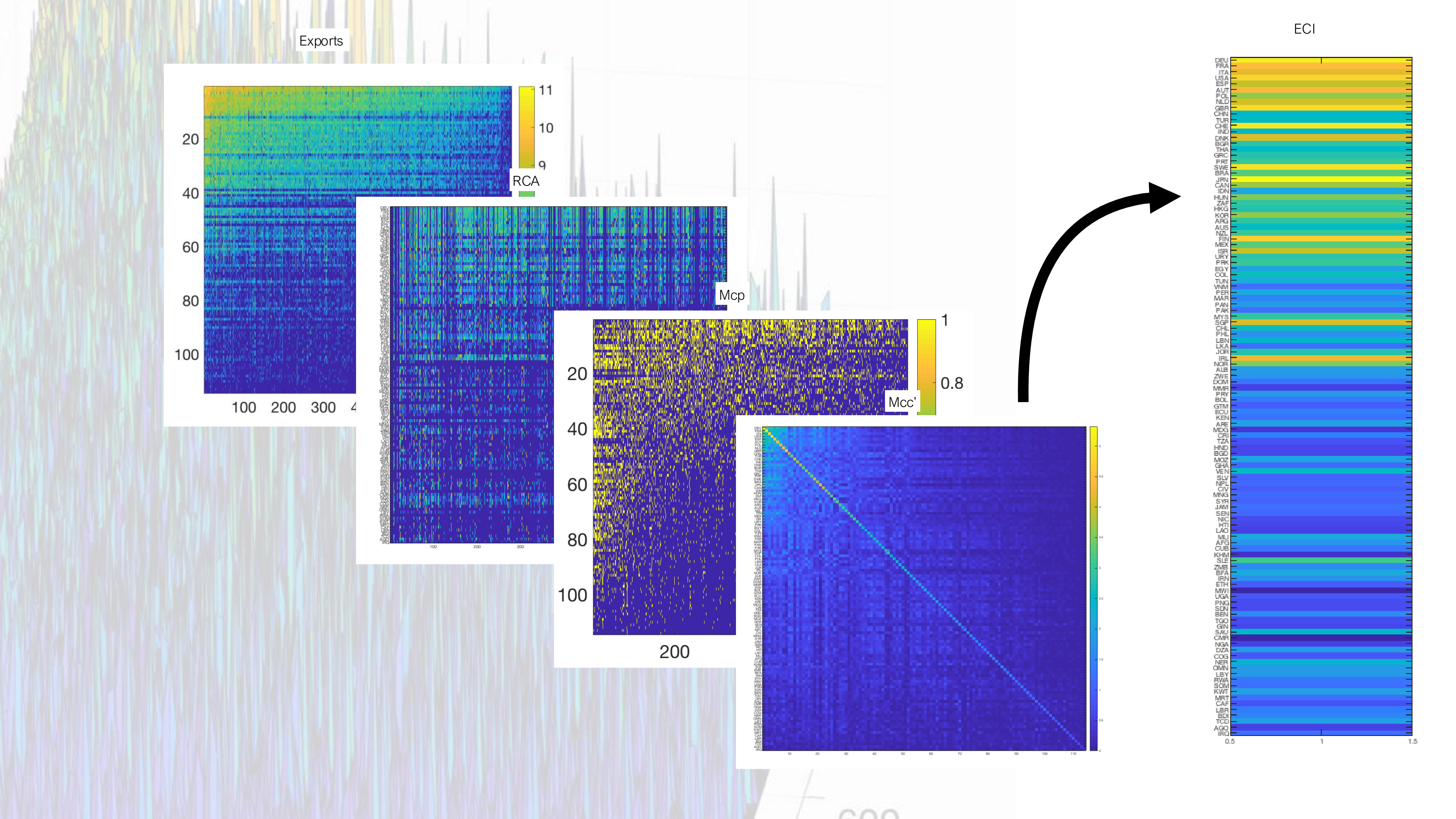
But is not that easy!

Units of observation are not comparable!

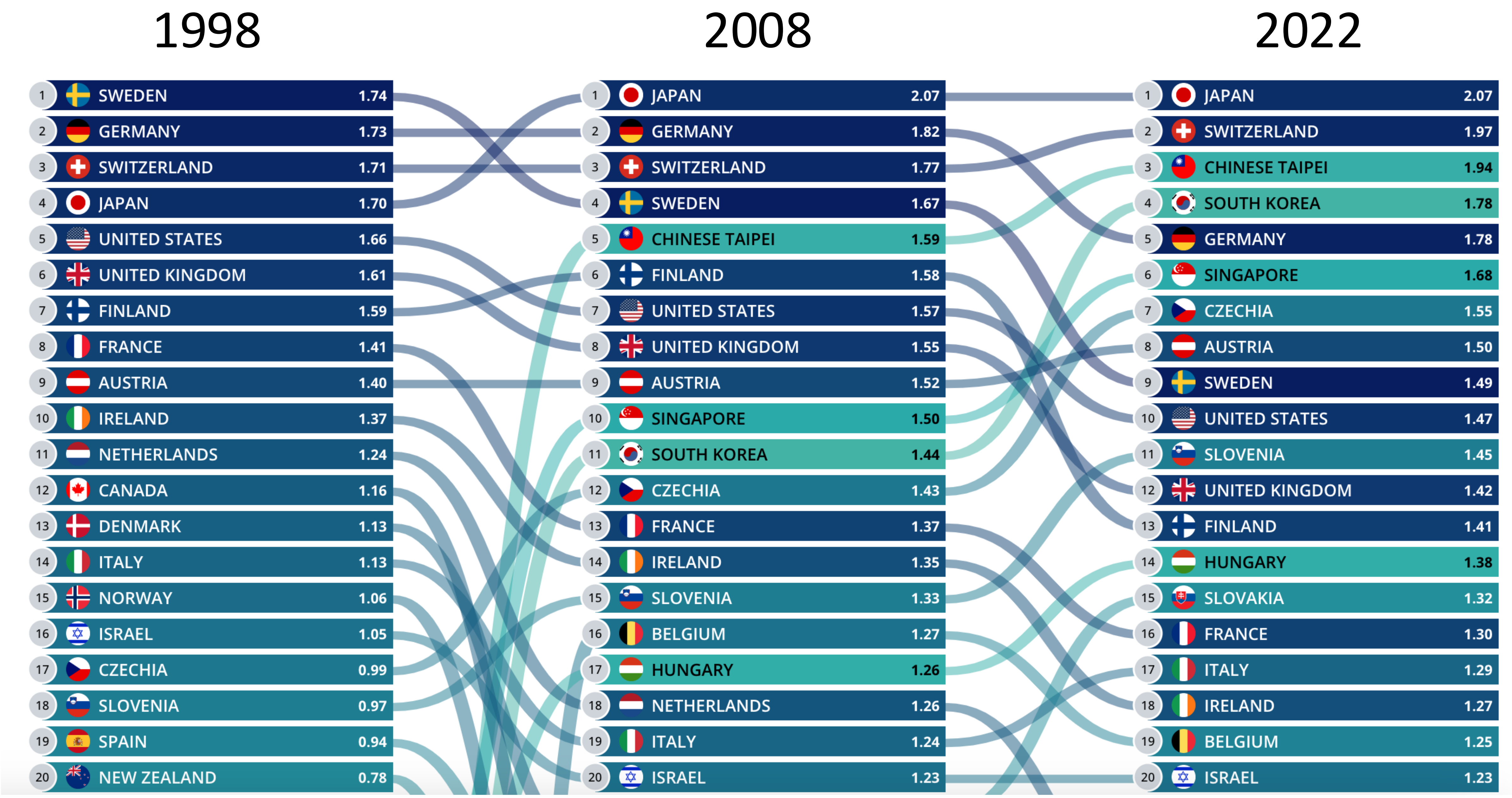


China & USA ~ 15 to 20 trillion GDP

Macedonia ~ 0.0012 trillion GDP


Exports
11
10
9
20
40
60
80
100
100
200
300
RCA
Mcp
1
0.8
20
40
60
80
100
200
Mcc'
ECI
0.5
1
1.5

Economic Complexity Index (Trade)



1998

10	IRELAND	1.37
11	NETHERLANDS	1.24
12	CANADA	1.16
13	DENMARK	1.13
14	ITALY	1.13
15	NORWAY	1.06
16	ISRAEL	1.05
17	CZECHIA	0.99
18	SLOVENIA	0.97
19	SPAIN	0.94
20	NEW ZEALAND	0.78
21	SLOVAKIA	0.74
22	RUSSIA	0.71
23	BELARUS	0.70
24	BRAZIL	0.68
25	UKRAINE	0.67
26	SINGAPORE	0.61
27	AUSTRALIA	0.61
28	HUNGARY	0.57
29	CHINESE TAIPEI	0.55
30	POLAND	0.53
31	SOUTH AFRICA	0.52
32	SOUTH KOREA	0.47

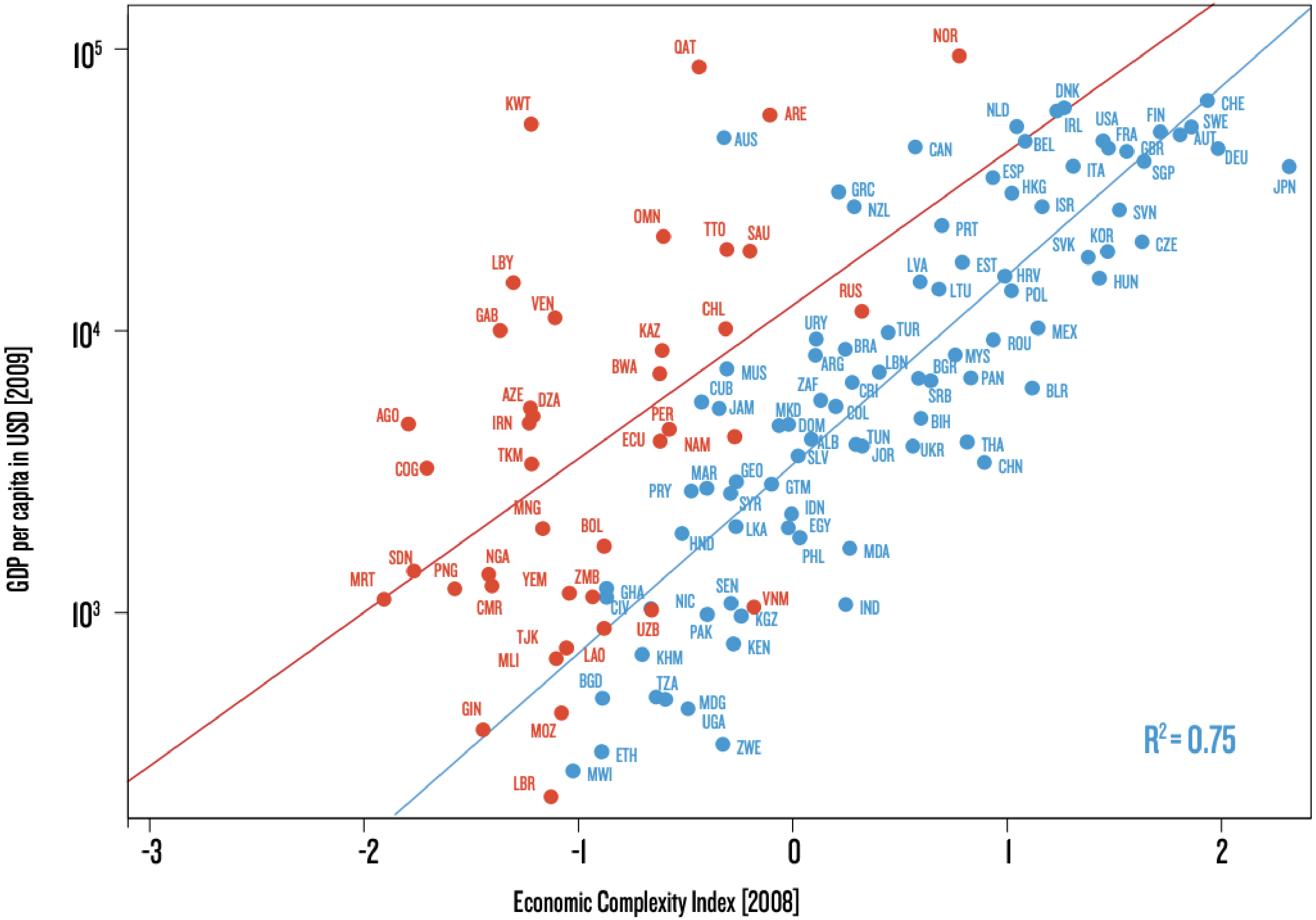
2008

10	SINGAPORE	1.50
11	SOUTH KOREA	1.44
12	CZECHIA	1.43
13	FRANCE	1.37
14	IRELAND	1.35
15	SLOVENIA	1.33
16	BELGIUM	1.27
17	HUNGARY	1.26
18	NETHERLANDS	1.26
19	ITALY	1.24
20	ISRAEL	1.23
21	SLOVAKIA	1.10
22	DENMARK	1.04
23	MEXICO	0.96
24	POLAND	0.96
25	CANADA	0.90
26	SPAIN	0.87
27	NORTH KOREA	0.87
28	MALAYSIA	0.83
29	NORWAY	0.83
30	BELARUS	0.82
31	BRAZIL	0.76
32	UKRAINE	0.70

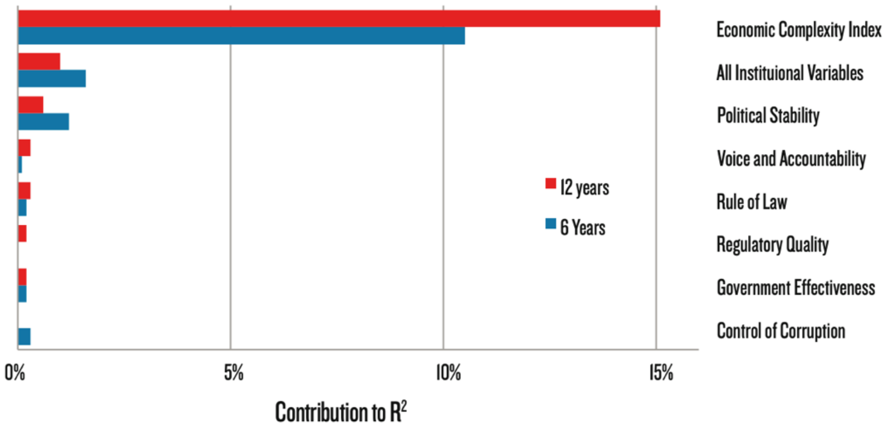
2022

10	UNITED STATES	1.47
11	SLOVENIA	1.45
12	UNITED KINGDOM	1.42
13	FINLAND	1.41
14	HUNGARY	1.38
15	SLOVAKIA	1.32
16	FRANCE	1.30
17	ITALY	1.29
18	IRELAND	1.27
19	BELGIUM	1.25
20	ISRAEL	1.23
21	HONG KONG	1.18
22	CHINA	1.12
23	NETHERLANDS	1.11
24	MALAYSIA	1.09
25	MEXICO	1.08
26	ROMANIA	1.03
27	DENMARK	1.01
28	POLAND	0.99
29	THAILAND	0.96
30	LITHUANIA	0.93
31	CANADA	0.92
32	BELARUS	0.82

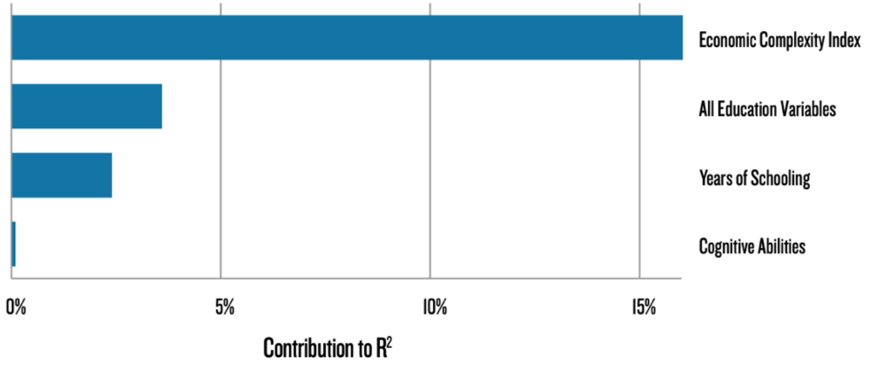
Economic Complexity Explains Future Economic Growth



Explains more growth than institutions



Explains more growth than education



Atlas of Economic Complexity (Puritan Press 2011, MIT Press 2014)

Explorations in Economic History
Volume 83, January 2022, 101421

Patterns of specialization and economic complexity through the lens of universal exhibitions, 1855-1900

Giacomo Domini

	(3)	(4)
	Next-century growth	**Next-century growth**
ECI	0.509***	0.400***
	(0.154)	(0.128)
GDP per capita	-0.597***	-0.542**
	(0.161)	(0.219)
Constant	-0.128	-0.092**
	(0.121)	(0.043)
Country fixed effects	No	Yes
N of observations	96	96
N of countries	33	33
N of time periods	5	5
Adjusted R^2	0.221	0.770

Domini, Giacomo. *Explorations in Economic History* 83 (2022): 101421.

Economic Complexity Explains Variations in Income Inequality

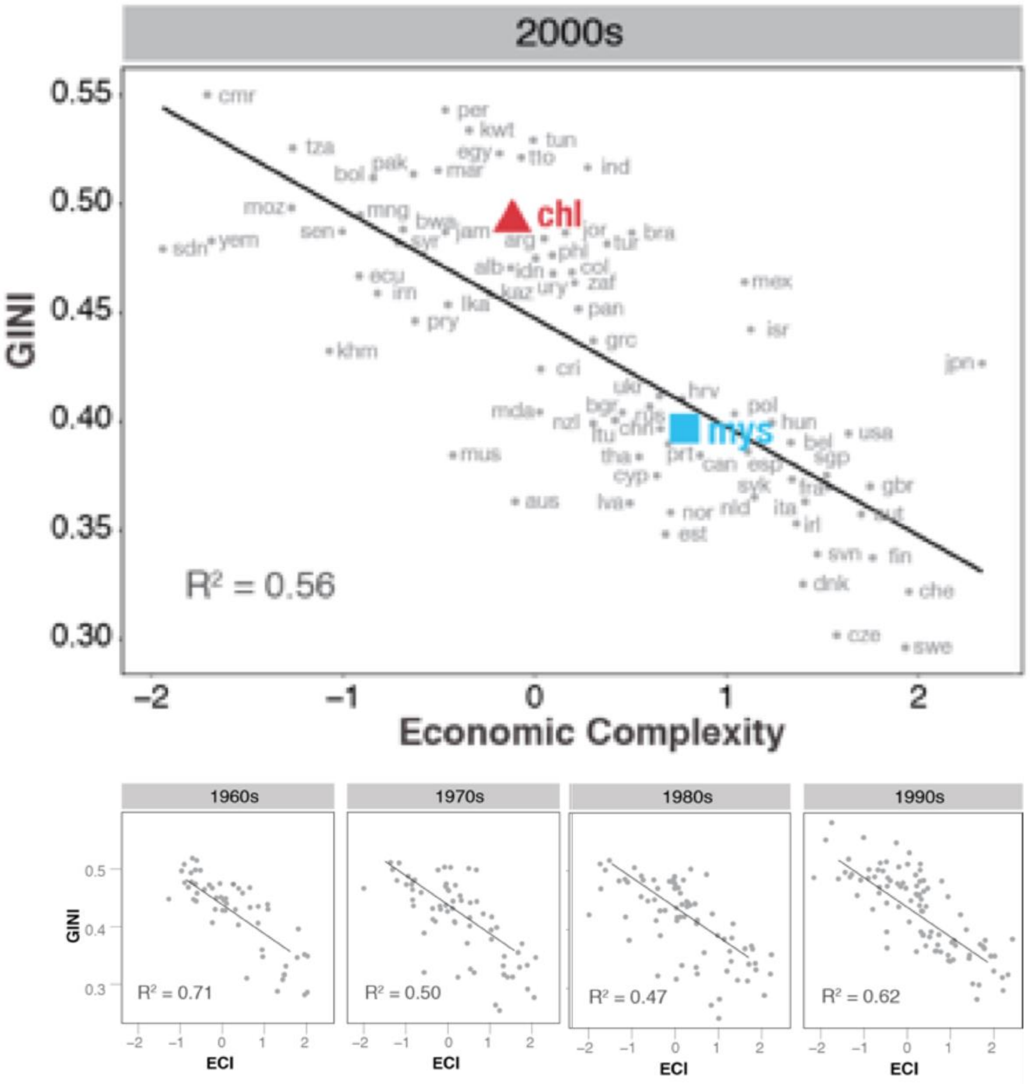


Table 10. *Cross-section regression results*

	Dependent variable: GINI EHII					
	(1)	(2)	(3)	(4)	(5)	(6)
ECI	−0.040***					−0.036***
	(0.007)					(0.007)
Fitness Index		−0.023***				
		(0.005)				
Entropy			−0.025***			
			(0.005)			
HHI				0.146***		0.058
				(0.044)		(0.044)
ln(GDP PPP pc)	0.067**	0.036	0.086***	0.065**	0.059*	0.068**
	(0.028)	(0.029)	(0.029)	(0.031)	(0.032)	(0.028)
ln(GDP PPPpc)^2	−0.004**	−0.002	−0.005***	−0.004**	−0.004*	−0.004**
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Schooling	−0.005***	−0.008***	−0.008***	−0.008***	−0.009***	−0.005***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
ln population	0.007**	0.010***	0.008***	0.004	0.0001	0.007***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Rule of law	−0.013	−0.008	−0.013	−0.017	−0.016	−0.014
	(0.013)	(0.013)	(0.013)	(0.014)	(0.014)	(0.013)
Corruption control	0.011	0.009	0.011	0.019	0.027*	0.009
	(0.013)	(0.014)	(0.013)	(0.014)	(0.014)	(0.013)
Government effectiveness	0.002	−0.013	−0.007	−0.012	−0.022	0.003
	(0.017)	(0.017)	(0.017)	(0.018)	(0.018)	(0.017)
Political stability	−0.010	−0.011*	−0.014**	−0.017**	−0.017**	−0.011*
	(0.006)	(0.007)	(0.006)	(0.007)	(0.007)	(0.006)
Regulatory quality	−0.006	−0.006	0.001	−0.0002	−0.012	−0.002
	(0.012)	(0.013)	(0.013)	(0.014)	(0.014)	(0.013)
Voice and accountability	0.001	0.009	0.015*	0.011	0.006	0.004
	(0.008)	(0.008)	(0.008)	(0.008)	(0.009)	(0.008)
Constant	0.083	0.199	0.132	0.206	0.286**	0.071
	(0.130)	(0.131)	(0.132)	(0.138)	(0.141)	(0.130)
Observations	142	142	142	142	142	142
R^2	0.717	0.698	0.703	0.667	0.639	0.721
Adjusted R^2	0.693	0.672	0.678	0.639	0.612	0.695
Residual std. error	0.035	0.036	0.035	0.037	0.039	0.034
	(df = 130)	(df = 130)	(df = 130)	(df = 130)	(df = 131)	(df = 129)

Hartmann, Guevara, Jara-Figueroa, Aristaran, & Hidalgo,. World Development (2017)

Economic Complexity Explains Greenhouse Emission Intensity

ELSEVIER
Contents lists available at ScienceDirect
World Development
journal homepage: www.elsevier.com/locate/worlddev

Economic complexity and greenhouse gas emissions

João P. Romero [a,*], Camila Gramkow [b,1]

[a] Universidade Federal de Minas Gerais (UFMG), Center for Development and Regional Planning (Cedeplar), Brazil
[b] United Nations Economic Commission for Latin America and the Caribbean (ECLAC), Brazil and Chile

Table 2
Emission intensity fixed effects regressions.

Model	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
ECI	−0.0475	−0.0423	−0.0501	0.0676	−0.0709	−0.0840	0.0437	0.0912
	(0.119)	(0.136)	(0.125)	(0.0711)	(0.117)	(0.130)	(0.0879)	(0.0708)
Lagged ECI	−0.156**	−0.166*	−0.156**	−0.169**	−0.128*	−0.118	−0.166**	−0.137**
	(0.0763)	(0.0846)	(0.0777)	(0.0805)	(0.0749)	(0.0737)	(0.0721)	(0.0562)
Ln of GDP per capita	−0.470**	−0.450*	−0.472**	−0.628***	−0.438**	−0.491**	−0.382***	−0.408**
	(0.189)	(0.238)	(0.191)	(0.105)	(0.185)	(0.187)	(0.0956)	(0.172)
Ln of Agric. Share	0.172*	0.148	0.170*	0.138*	0.138	0.182*	0.143*	0.0678
	(0.0963)	(0.0994)	(0.0968)	(0.0792)	(0.0879)	(0.0931)	(0.0844)	(0.0778)
Ln of Openness	0.167**	0.171**	0.166**	0.151*	0.165**	0.174**	0.0594	0.0958
	(0.0768)	(0.0782)	(0.0736)	(0.0771)	(0.0742)	(0.0703)	(0.0626)	(0.0667)
Ln of Electricity Cons.		0.0112						0.158
		(0.125)						(0.110)
Ln of Urbanization			0.0280					−0.770***
			(0.247)					(0.232)
Ln of Sec. School Enrol.				0.0441				−0.00561
				(0.107)				(0.0922)
Ln of Population					0.253			0.419*
					(0.321)			(0.232)
Ln of Manuf. Share						0.114		−0.0526
						(0.0744)		(0.0660)
Ln of Patents							0.0000429	−0.00135
							(0.0217)	(0.0234)
Constant	9.977***	9.779***	9.900***	11.22***	5.635	9.774***	9.661***	4.991
	(1.589)	(1.690)	(1.769)	(0.847)	(5.466)	(1.725)	(0.836)	(3.752)
N. Obs.	485	469	485	439	485	469	383	344
Adj. R-sq.	0.358	0.359	0.357	0.515	0.361	0.406	0.636	0.728

Note: Dependent variable: Ln of GHG emissions (kilotons of CO2 equivalent) by units of output (billions of 2010 USD). Time dummies were included in all the regressions. Robust standard erros between brackets. Significance levels: *** = 1%; ** = 5%; * = 10%.
Source: Authors' elaboration.

Springer Link

Open Access | Published: 05 February 2021

Economic Complexity and Environmental Performance: Evidence from a World Sample

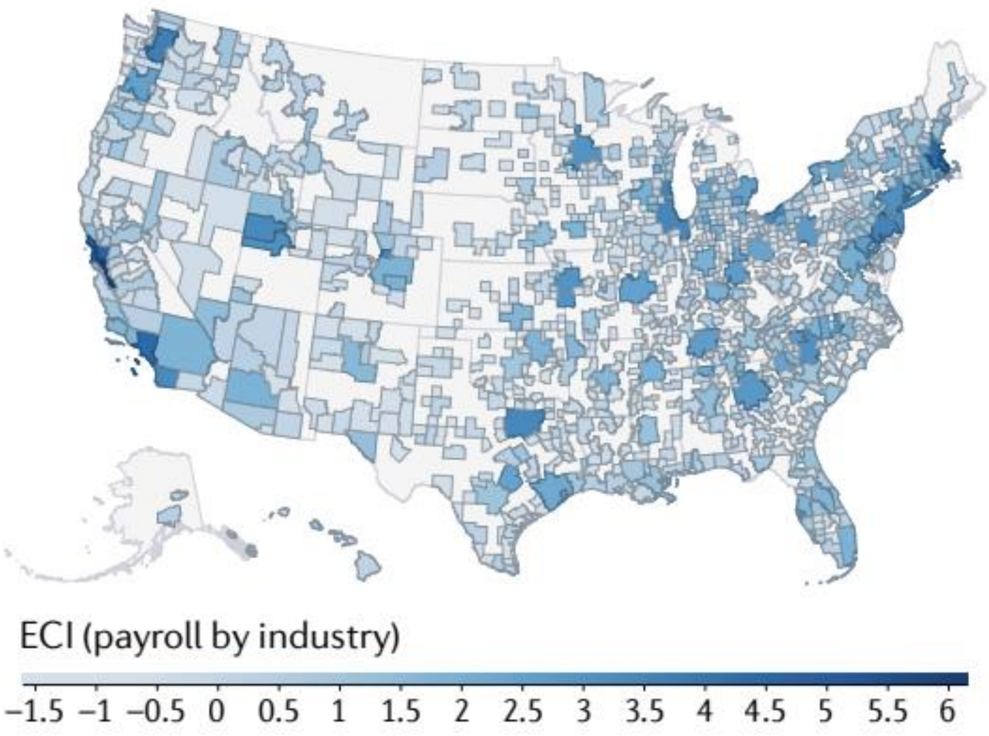
Eirini Boleti, Antonios Garas, Alexandra Kyriakou & Athanasios Lapatinas

Table 2 The effect of economic complexity on environmental performance: pooled OLS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ECI	6.414***	4.869***	5.167***	3.904***	3.584***	3.459***	4.071***	3.220***
	(0.331)	(0.337)	(0.343)	(0.378)	(0.408)	(0.403)	(0.422)	(0.39)
GDP per capita	7.805***	7.770***	7.532***	7.704***	6.891***	7.158***	6.541***	5.760***
	(0.321)	(0.306)	(0.305)	(0.437)	(0.478)	(0.476)	(0.584)	(0.576)
GDP per capita^2	0.443***	0.639***	0.658***	0.883***	0.651***	0.591***	0.14	-0.262
	(0.136)	(0.149)	(0.149)	(0.157)	(0.169)	(0.165)	(0.182)	(0.179)
Population			-0.006***	-0.006***	-0.006***	-0.006***	-0.005***	-0.006***
			(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Agriculture				-0.115***	-0.106***	-0.140***	-0.129***	-0.157***
				(0.038)	(0.038)	(0.039)	(0.046)	(0.044)
Industry				-0.057***	-0.028	-0.049**	0.005	0.012
				(0.022)	(0.022)	(0.023)	(0.027)	(0.026)
Corruption					1.248***	1.036***	2.144***	1.186***
					(0.377)	(0.371)	(0.383)	(0.354)
Trade						0.023***	0.006	0.015***
						(0.005)	(0.005)	(0.005)
Urban							0.028	0.026
							(0.02)	(0.02)
Education							-0.000***	-0.000***
							(0.000)	(0.000)
OECD								6.523***
								(0.774)
Observations	1283	1210	1210	1160	1160	1149	940	940
R-squared	0.814	0.855	0.857	0.865	0.866	0.87	0.89	0.9
F-statistic	555.8	525.3	521.8	479	466.7	460.2	483.9	526.2

Dependent variable: Environmental Performance Index (EPI). Main independent variable: Economic Complexity Index (ECI). Time fixed effects are included in all regressions. Regional dummies are also included: *europe*, *asia*, *oceania*, *north america*, *south america*. Robust standard errors in parentheses

b Economic complexity of US MSAs
(industry payroll)



Economic complexity of US MSAs
(patents by technology class)

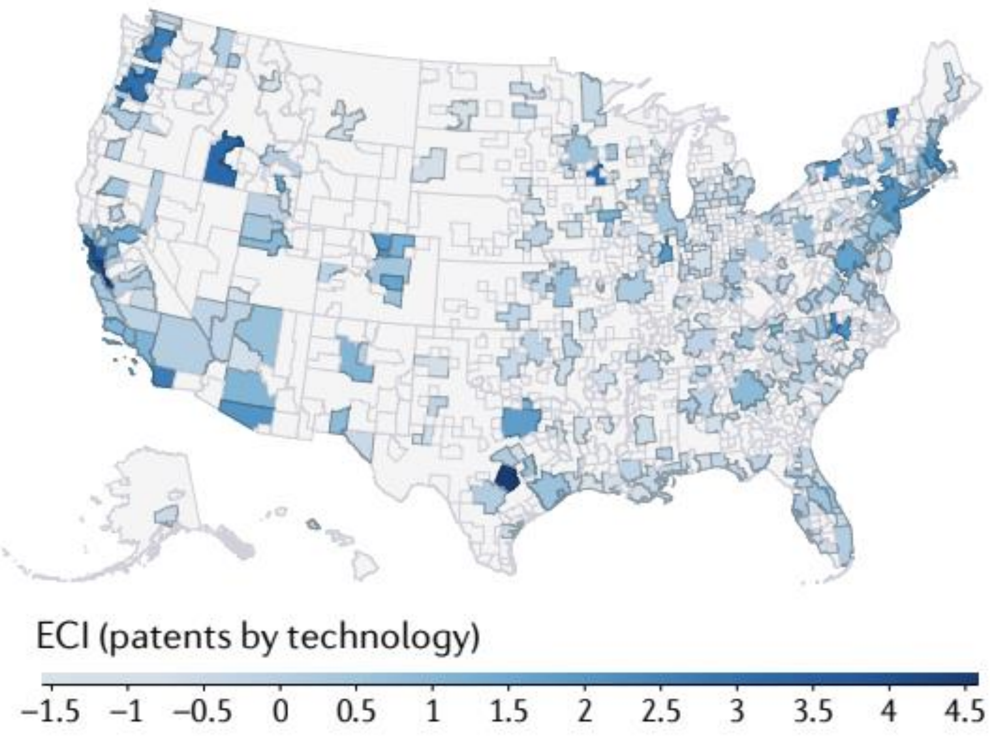
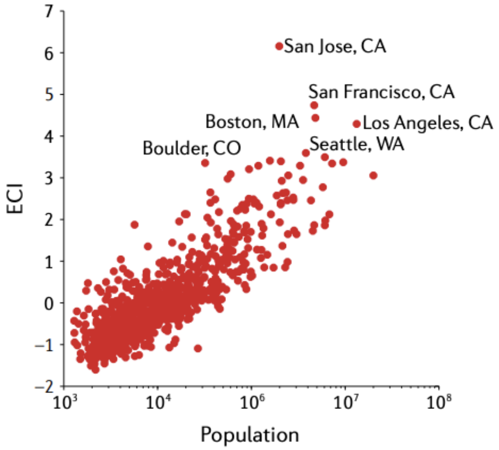


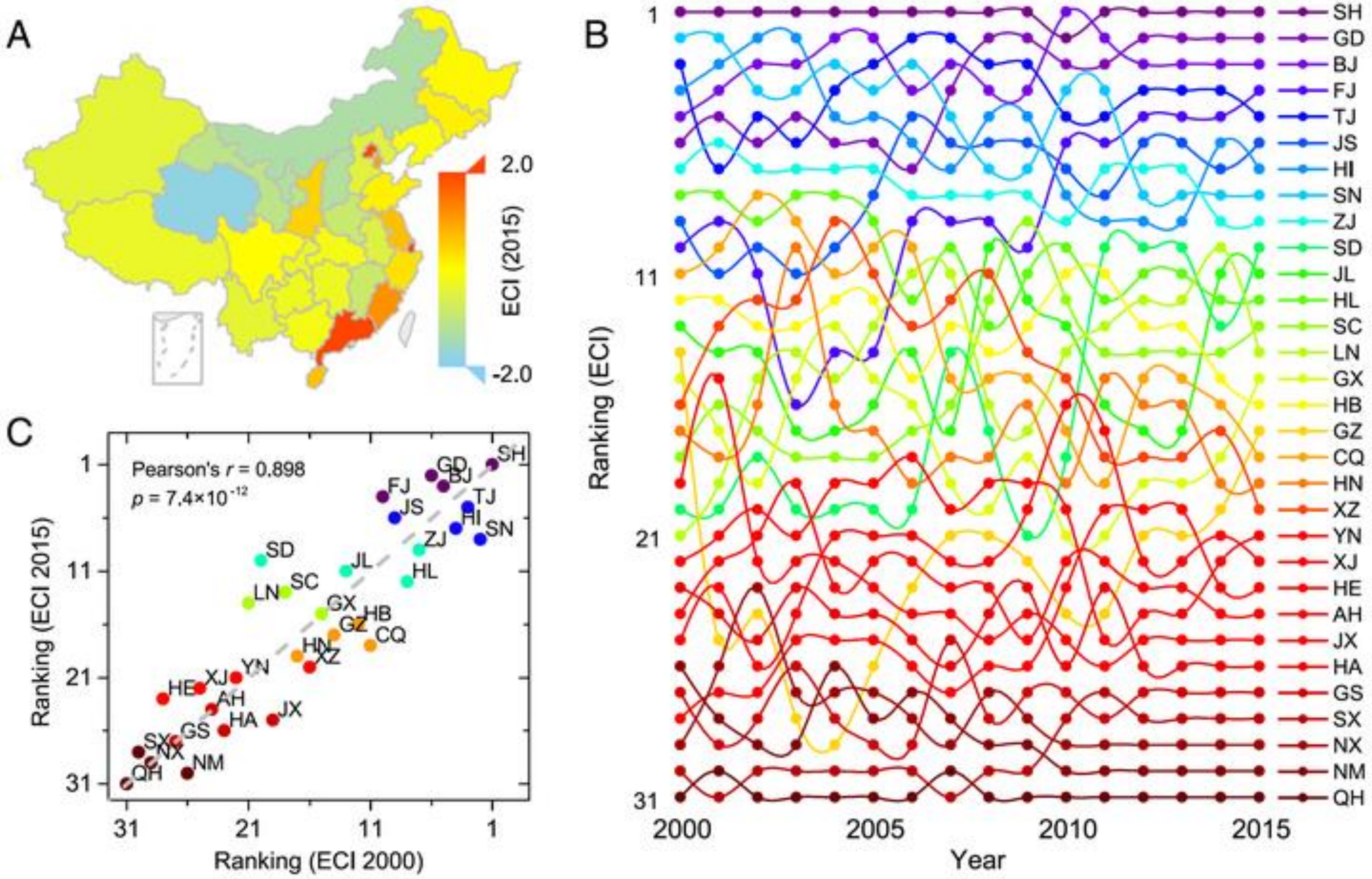
Table 1 | **Rankings of economic complexity**

Rank	**Economic complexity rankings**	
	US metro areas: payroll by industry (2018)	**US metro areas: patents by technology (2018)**
1	San Jose–Sunnyvale–Santa Clara, CA	San Jose–Sunnyvale–Santa Clara, CA
2	San Francisco–Oakland–Hayward, CA	Austin–Round Rock–San Marcos, TX
3	Boston–Cambridge–Newton, MA–NH	San Francisco–Oakland–Fremont, CA
4	Los Angeles–Long Beach–Anaheim, CA	Boise City–Nampa, ID
5	Seattle–Tacoma–Bellevue, WA	Rochester, MN



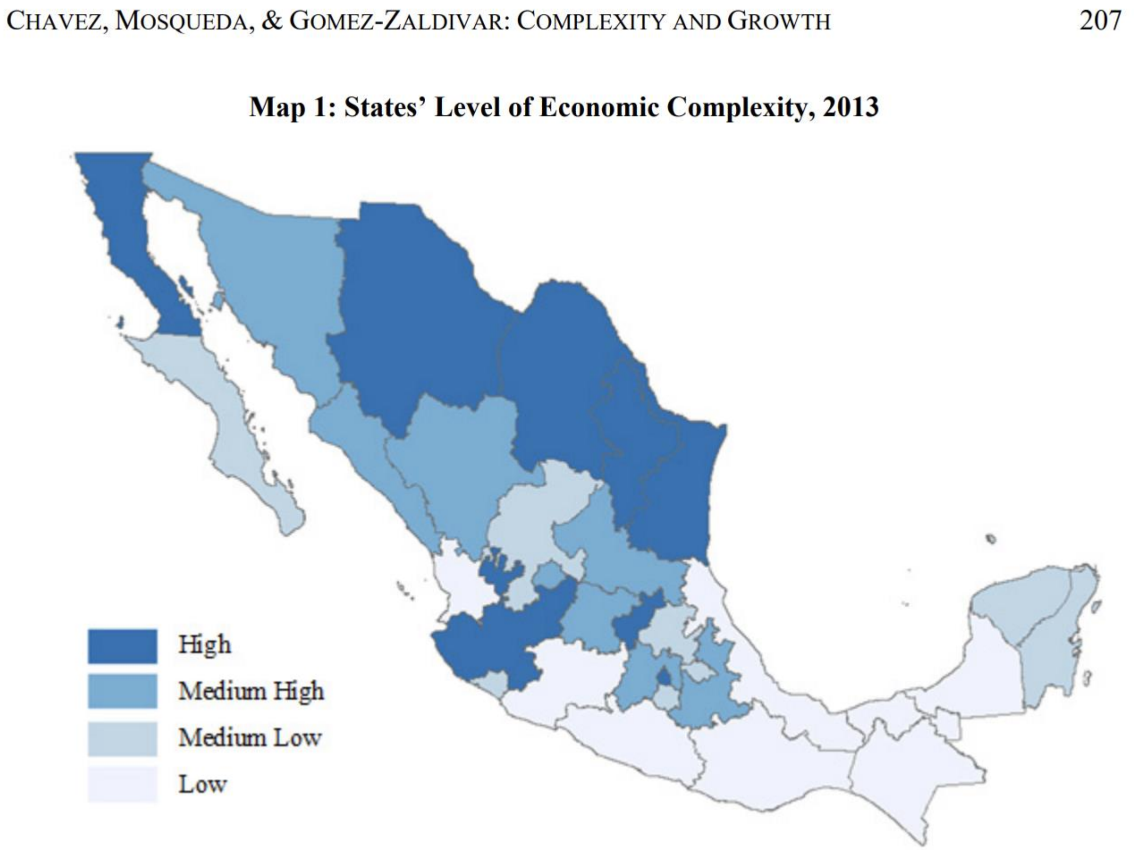
Hidalgo et al. Economic Complexity Theory and Applications. Nature Review Physics (2021)

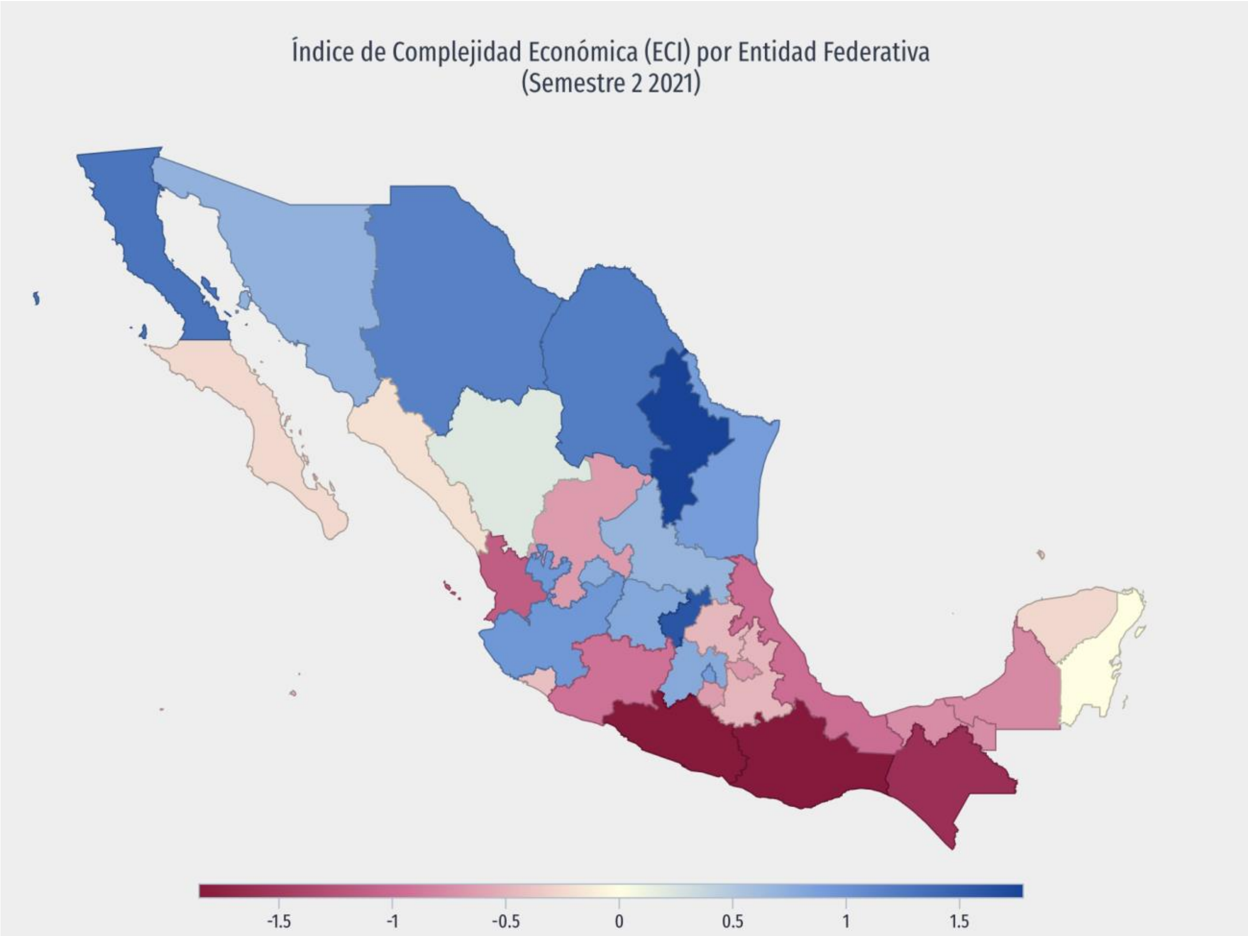
Economic Complexity of Chinese Provinces Using Data on Publicly Listed Firms



Gao, Jian, and Tao Zhou. "Quantifying China’s regional economic complexity." *Physica A: Statistical Mechanics and its Applications* 492 (2018): 1591-1603.

Economic Complexity of Mexican States Using Industry Data





Central Bank of Mexico

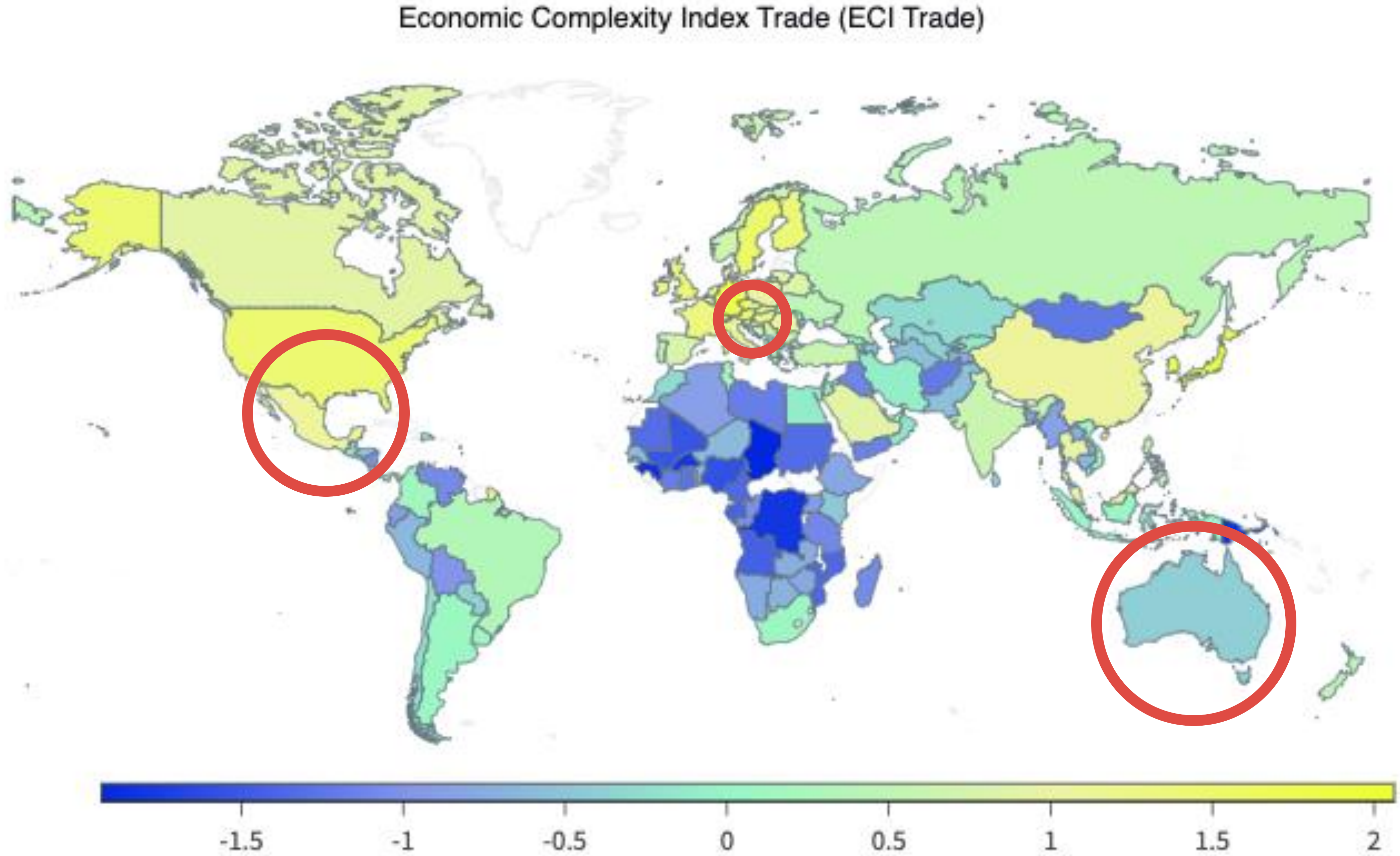
DataMexico.org

Multidimensional Turn

In 2009:
ECI(trade)
Economic Growth

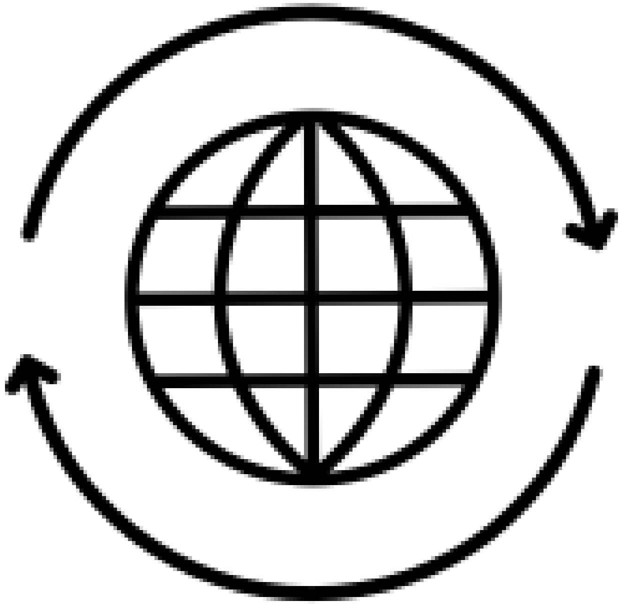
In 2023:
ECI(trade), ECI(tech), ECI(research)
Economic Growth, Inequality, Emissions

Limitations of trade ECI

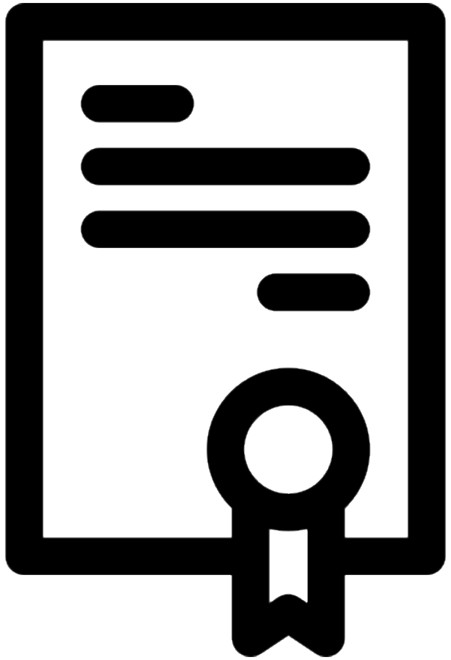


Stojkoski, Viktor, Philipp Koch, and César A. Hidalgo. "Multidimensional economic complexity and inclusive green growth." *Communications Earth & Environment* 4.1 (2023): 130.

Solution: Combine Data from Different Outputs



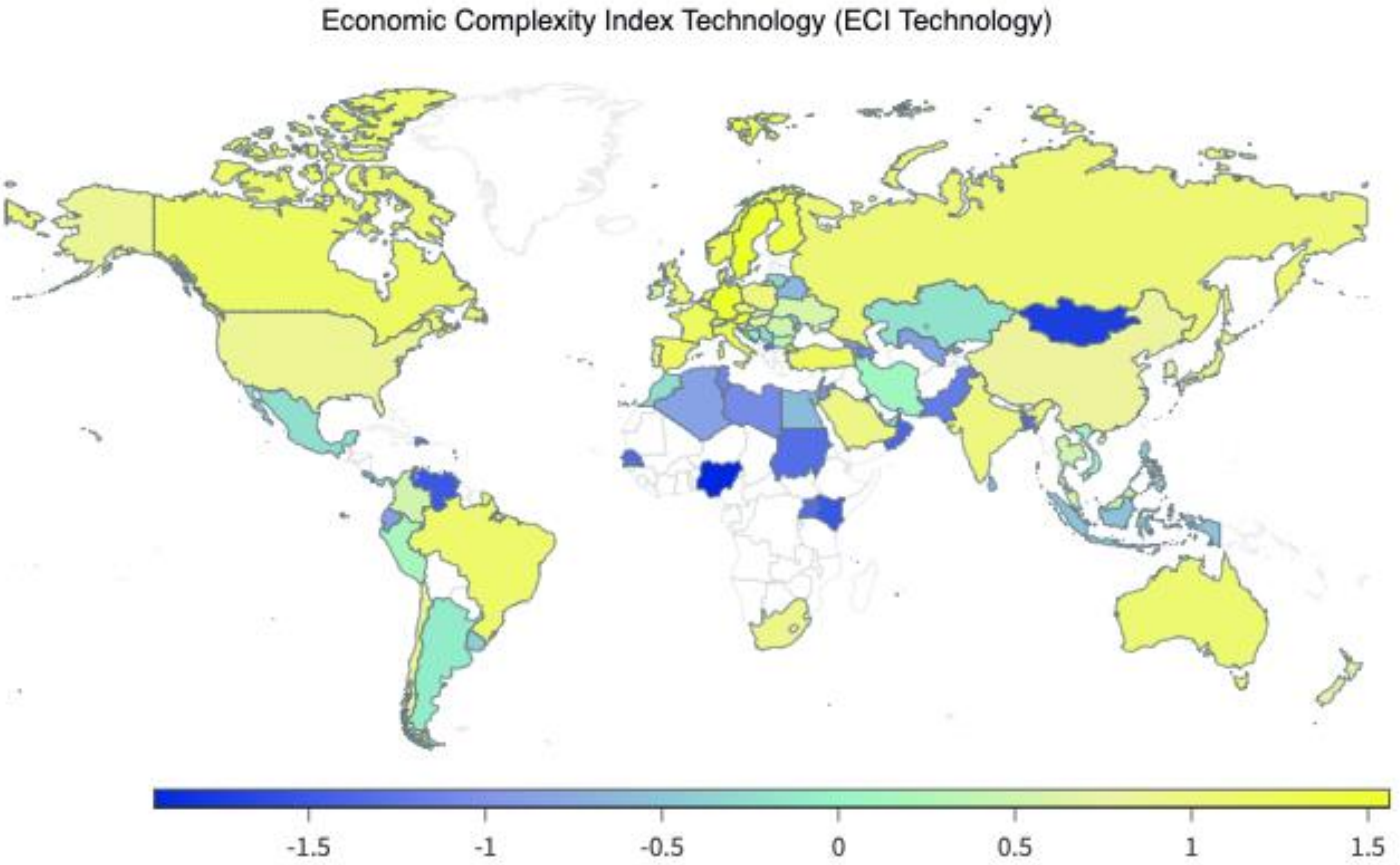
International Trade



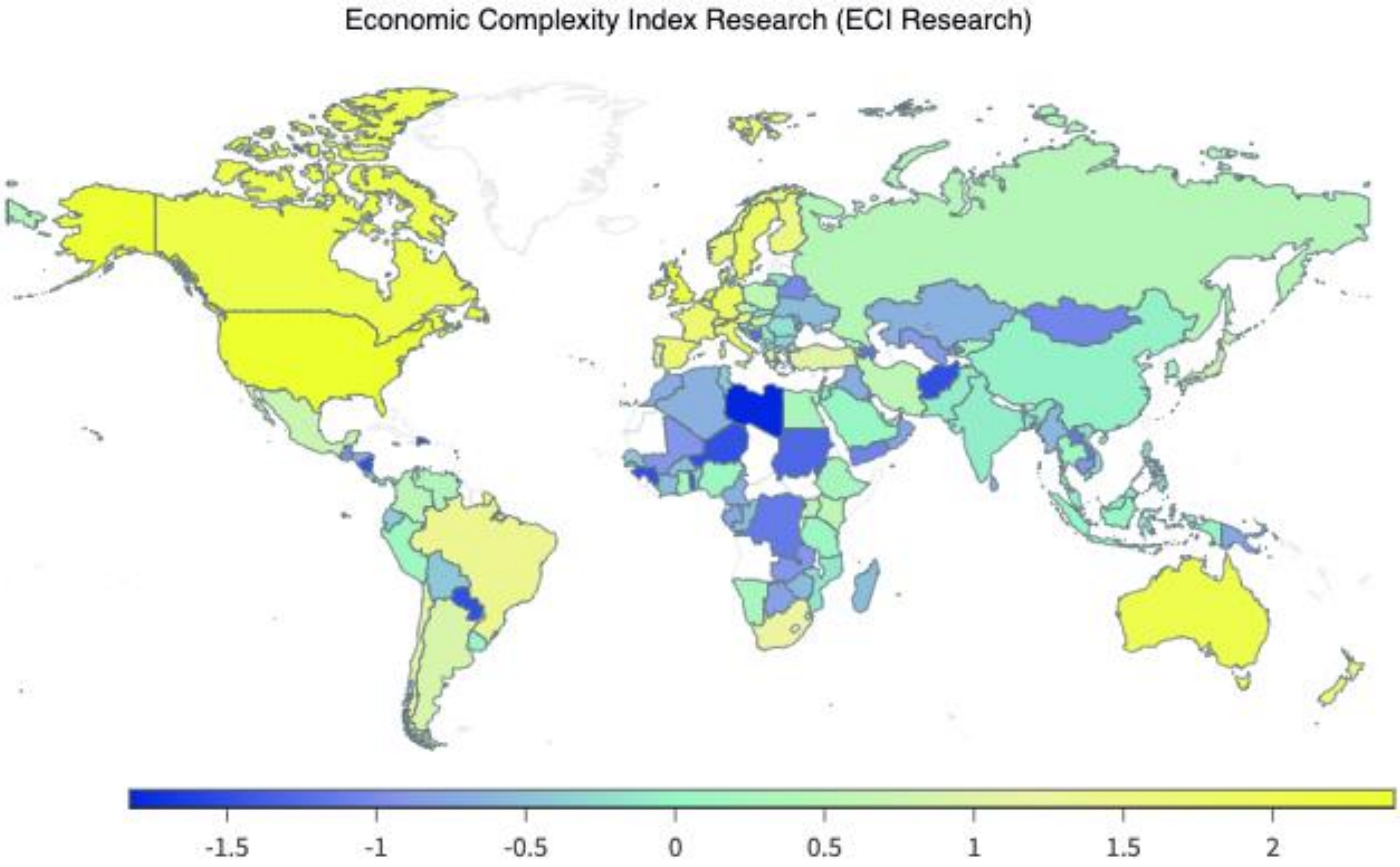
Patents

Research Papers

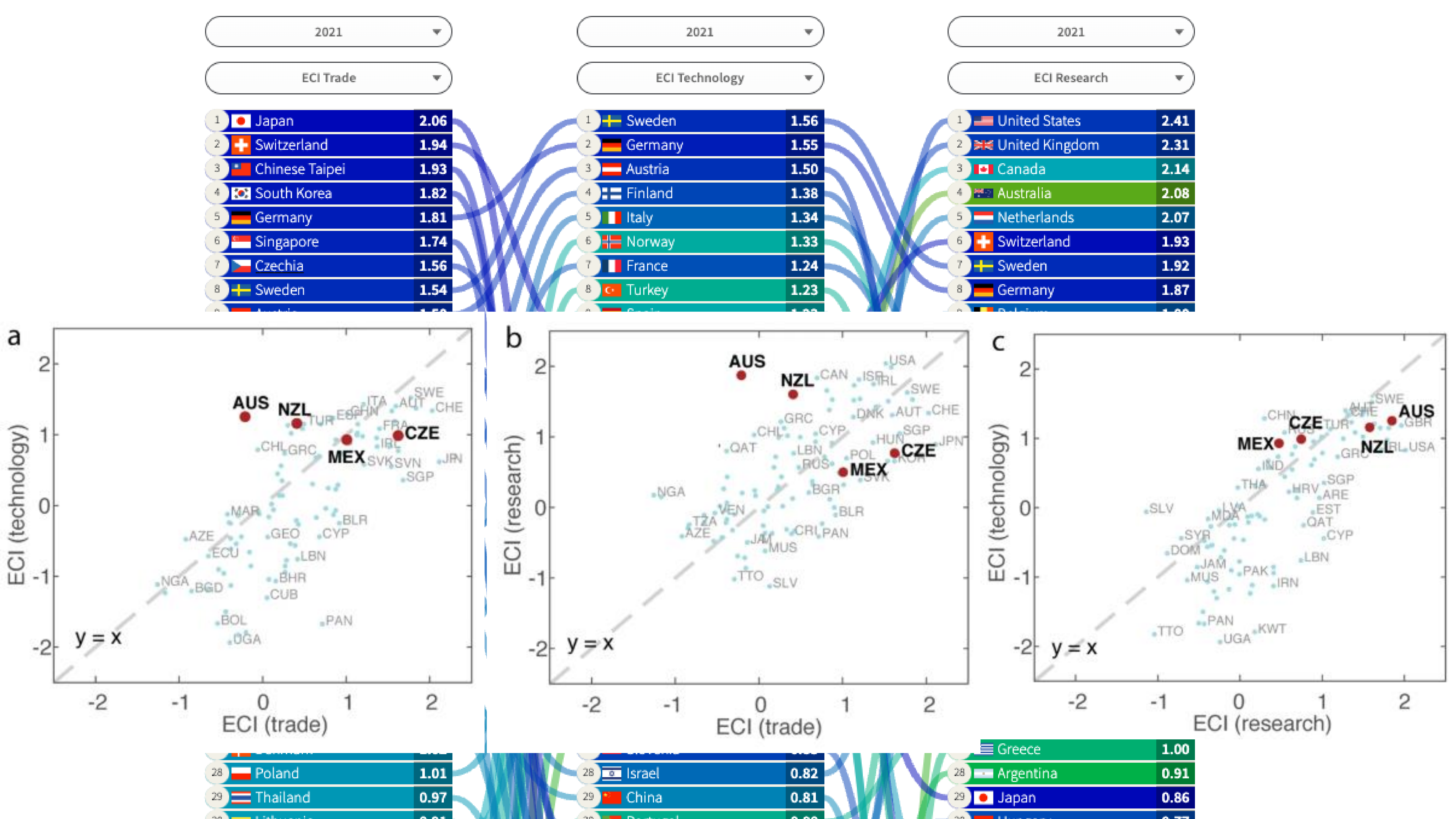
Stojkoski, Viktor, Philipp Koch, and César A. Hidalgo. "Multidimensional economic complexity and inclusive green growth." *Communications Earth & Environment* 4.1 (2023): 130.



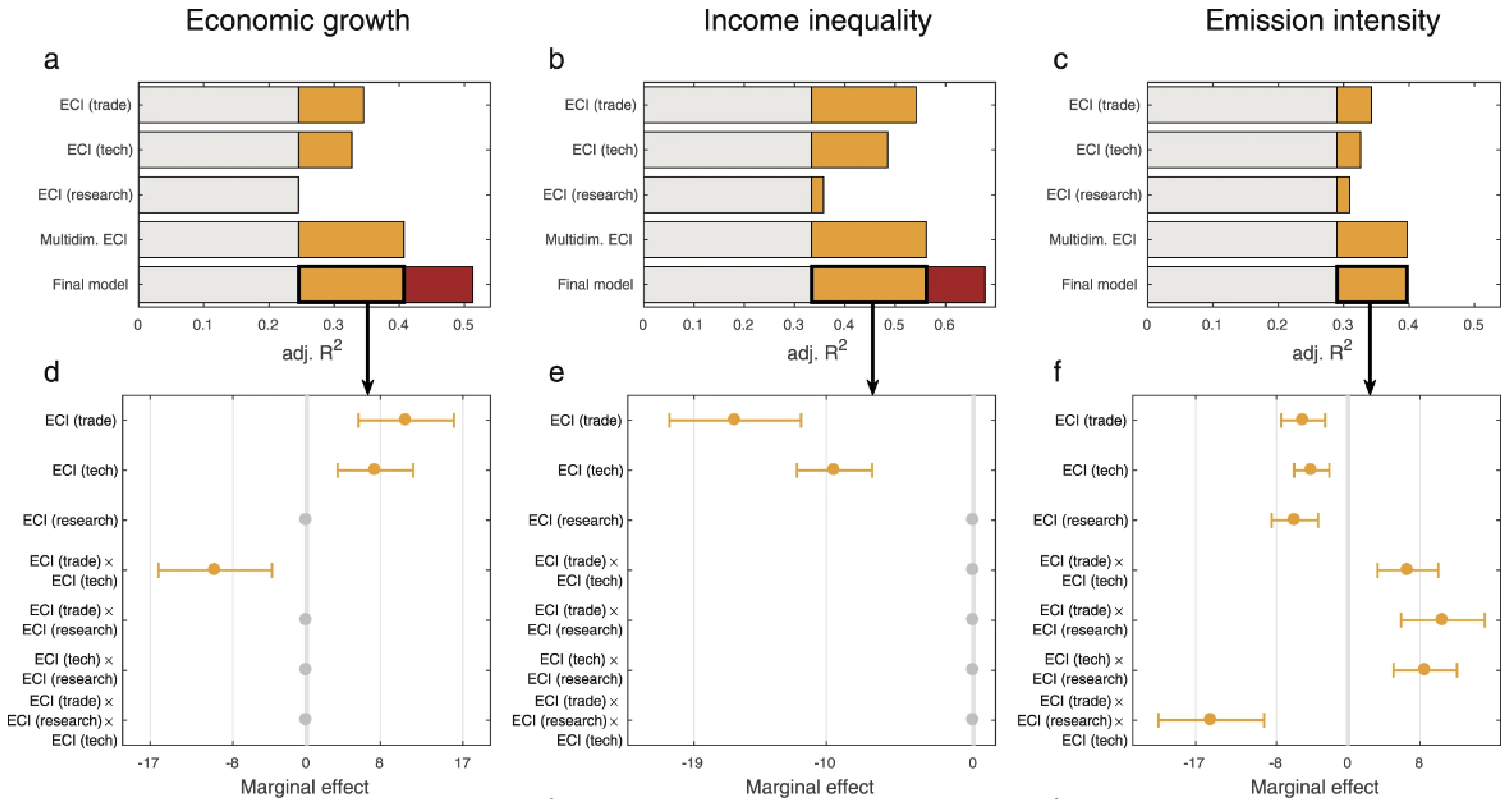
Stojkoski, Viktor, Philipp Koch, and César A. Hidalgo. "Multidimensional economic complexity and inclusive green growth." *Communications Earth & Environment* 4.1 (2023): 130.



Stojkoski, Viktor, Philipp Koch, and César A. Hidalgo. "Multidimensional economic complexity and inclusive green growth." *Communications Earth & Environment* 4.1 (2023): 130.



2021
ECI Trade
1 Japan 2.06
2 Switzerland 1.94
3 Chinese Taipei 1.93
4 South Korea 1.82
5 Germany 1.81
6 Singapore 1.74
7 Czechia 1.56
8 Sweden 1.54
28 Poland 1.01
29 Thailand 0.97
2021
ECI Technology
1 Sweden 1.56
2 Germany 1.55
3 Austria 1.50
4 Finland 1.38
5 Italy 1.34
6 Norway 1.33
7 France 1.24
8 Turkey 1.23
28 Israel 0.82
29 China 0.81
2021
ECI Research
1 United States 2.41
2 United Kingdom 2.31
3 Canada 2.14
4 Australia 2.08
5 Netherlands 2.07
6 Switzerland 1.93
7 Sweden 1.92
8 Germany 1.87
Greece 1.00
28 Argentina 0.91
29 Japan 0.86
a
ECI (technology)
ECI (trade)
y = x
AUS
NZL
MEX
CZE
b
ECI (research)
ECI (trade)
y = x
AUS
NZL
CZE
MEX
c
ECI (technology)
ECI (research)
y = x
AUS
CZE
MEX
NZL



Stojkoski, Viktor, Philipp Koch, and César A. Hidalgo. "Multidimensional economic complexity and inclusive green growth." *Communications Earth & Environment* 4.1 (2023): 130.

6	SINGAPORE	1.68	6	NORWAY	1.33	6	SWEDEN	1.99
7	CZECHIA	1.55	7	FRANCE	1.24	7	SWITZERLAND	1.94
8	AUSTRIA	1.50	8	TURKEY	1.23	8	IRELAND	1.91

ECONOMIC COMPLEXITY (TRADE)

1.38

Rank 14 Of 133

12	UNITED...	1.42
13	FINLAND	1.41
14	HUNGARY	1.38
15	SLOVAKIA	1.32
16	FRANCE	1.30
17	ITALY	1.29
18	IRELAND	1.27
19	BELGIUM	1.25
20	ISRAEL	1.23
21	HONG KONG	1.18
22	CHINA	1.12
23	NETHERLANDS	1.11
24	MALAYSIA	1.09
25	MEXICO	1.08
26	ROMANIA	1.03
27	DENMARK	1.01
28	POLAND	0.99
29	THAILAND	0.96
30	LITHUANIA	0.93
31	CANADA	0.92
32	BELARUS	0.82
33	SPAIN	0.81

ECONOMIC COMPLEXITY (TECHNOLOGY)

0.79

Rank 31 Of 96

12	BRAZIL	1.13
13	AUSTRALIA	1.11
14	UNITED...	1.09
15	BELGIUM	1.08
16	NETHERLANDS	1.08
17	RUSSIA	1.08
18	CZECHIA	1.08
19	DENMARK	1.02
20	POLAND	1.00
21	JAPAN	1.00
22	INDIA	1.00
23	SAUDI ARABIA	0.98
24	SOUTH...	0.93
25	UNITED...	0.87
26	CHILE	0.86
27	SLOVENIA	0.83
28	ISRAEL	0.82
29	CHINA	0.81
30	PORTUGAL	0.80
31	HUNGARY	0.79
32	SOUTH...	0.73
33	NEW...	0.73

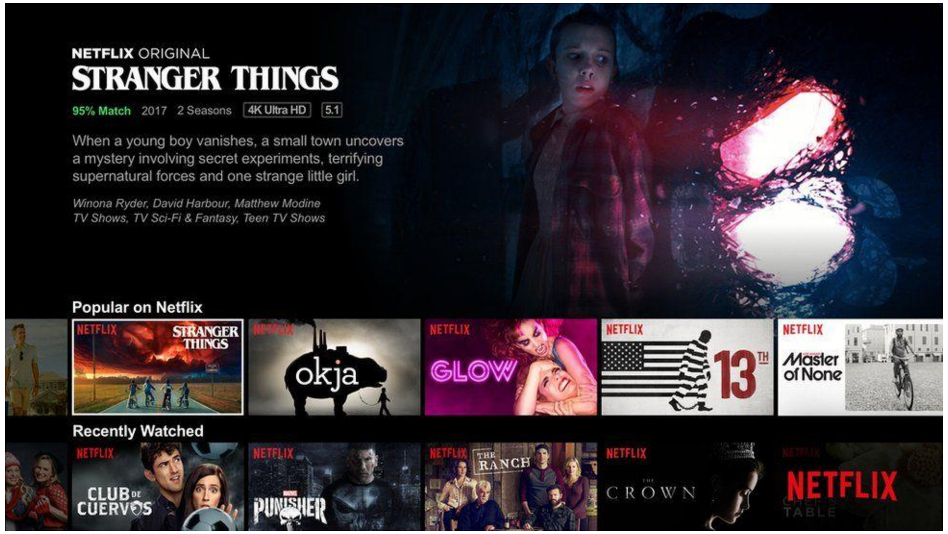
ECONOMIC COMPLEXITY (RESEARCH)

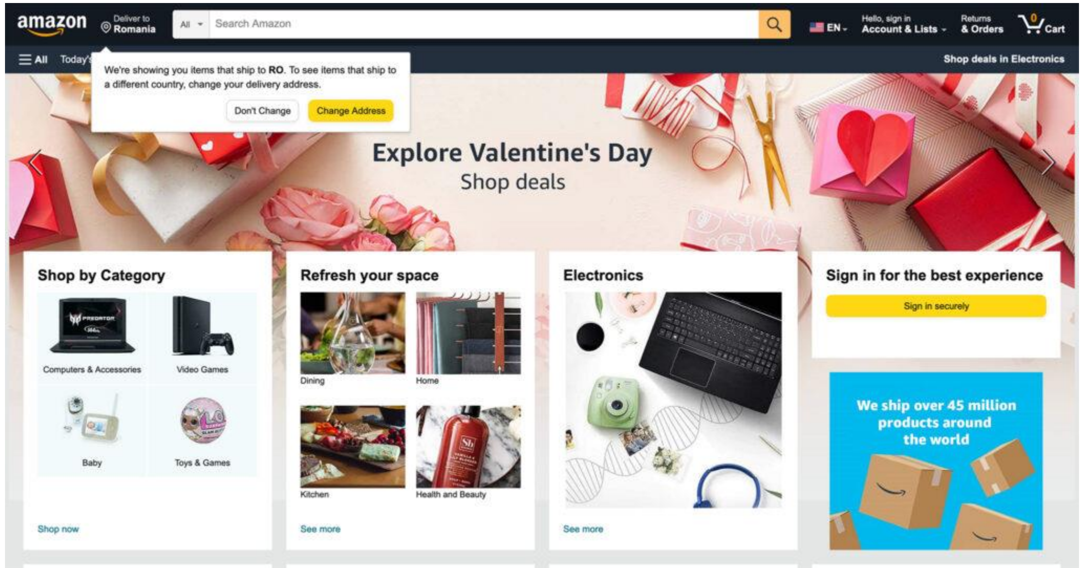
0.66

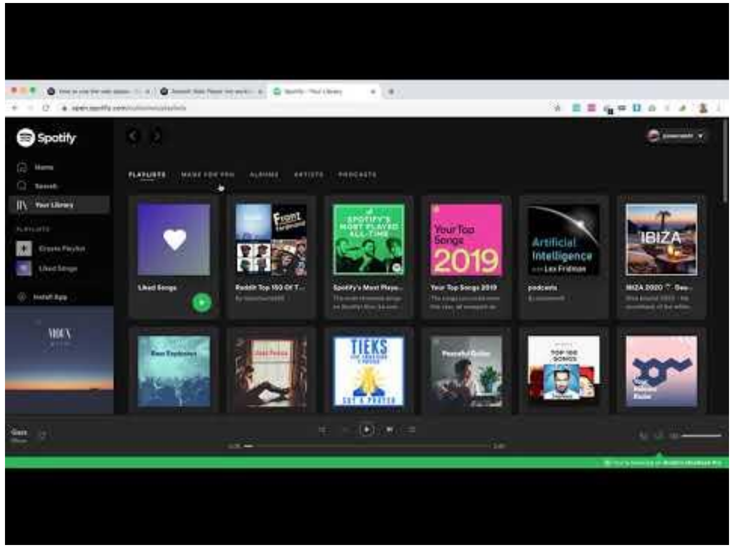
Rank 32 Of 135

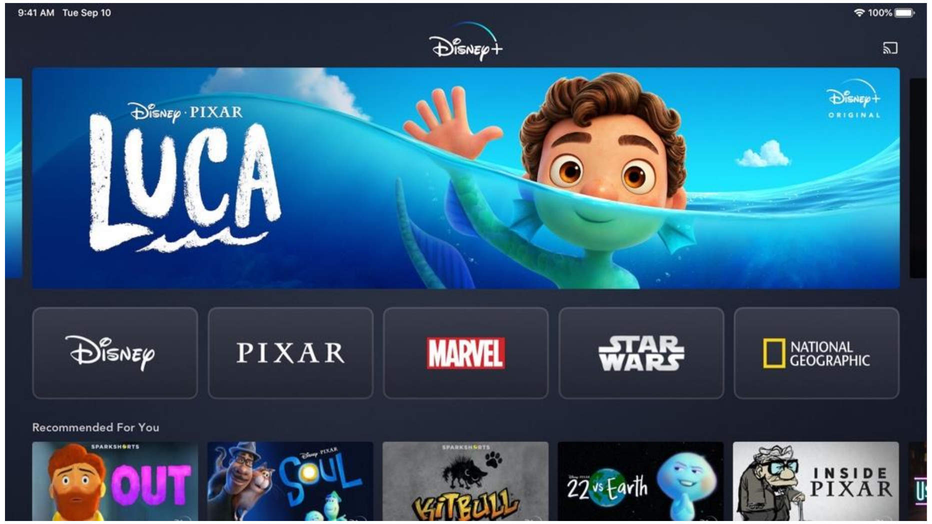
12	DENMARK	1.79
13	ITALY	1.73
14	SPAIN	1.73
15	FRANCE	1.72
16	AUSTRIA	1.69
17	NORWAY	1.66
18	NEW...	1.62
19	FINLAND	1.56
20	CHILE	1.22
21	BRAZIL	1.19
22	ARGENTINA	1.12
23	SOUTH...	1.11
24	TURKEY	0.89
25	PORTUGAL	0.88
26	JAPAN	0.84
27	HONG KONG	0.80
28	GEORGIA	0.77
29	GREECE	0.75
30	PARAGUAY	0.70
31	JAMAICA	0.67
32	HUNGARY	0.66
33	LEBANON	0.60

Have you ever made a purchase in any of these websites?



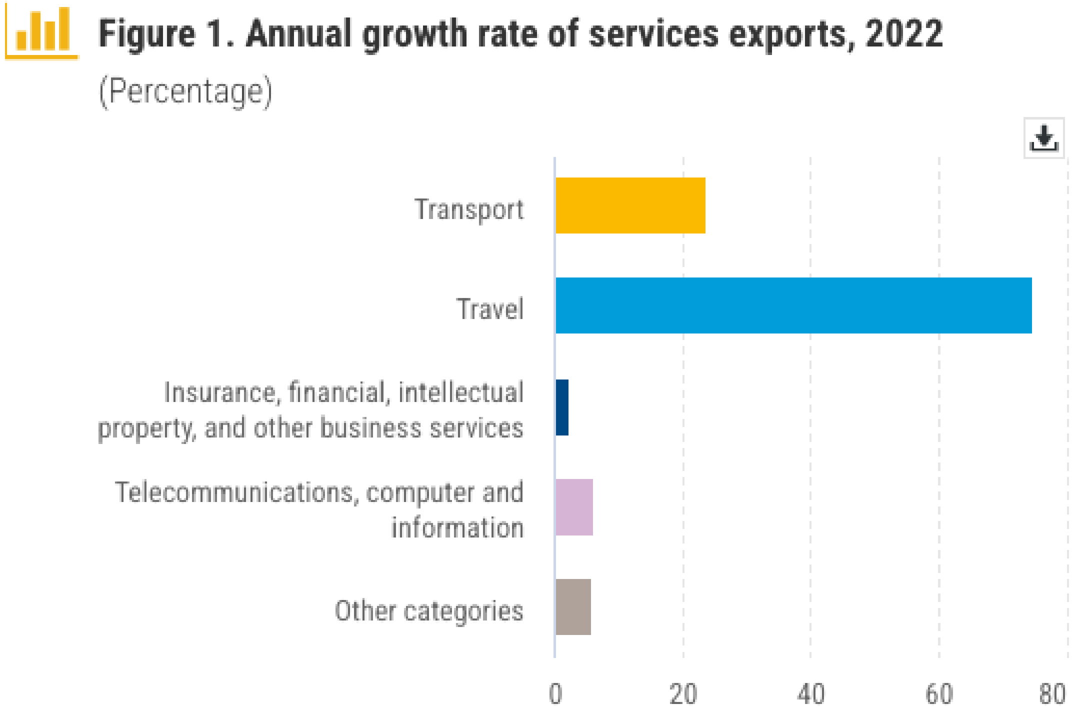






Service Trade Data is Not Very Detailed

UNCTAD HANDBOOK OF STATISTICS 2023



Data Estimation Procedure

2,502 Internet companies with revenues of USD 1B+

orbis

statista

Decompose revenue of each company into 29 sectors

13,013 firns and app developers

Google Play

APPMAGIC

App Store

Digital consumption Patterns for 2 sectors: mobile apps & games

Gradient-boosted regression tree.

(leave 20% out-of-sample cross-validation. MSE 0.048)

(Extends data to 129 additional countries and to 29 sectors)

Optimal Transport Reallocation

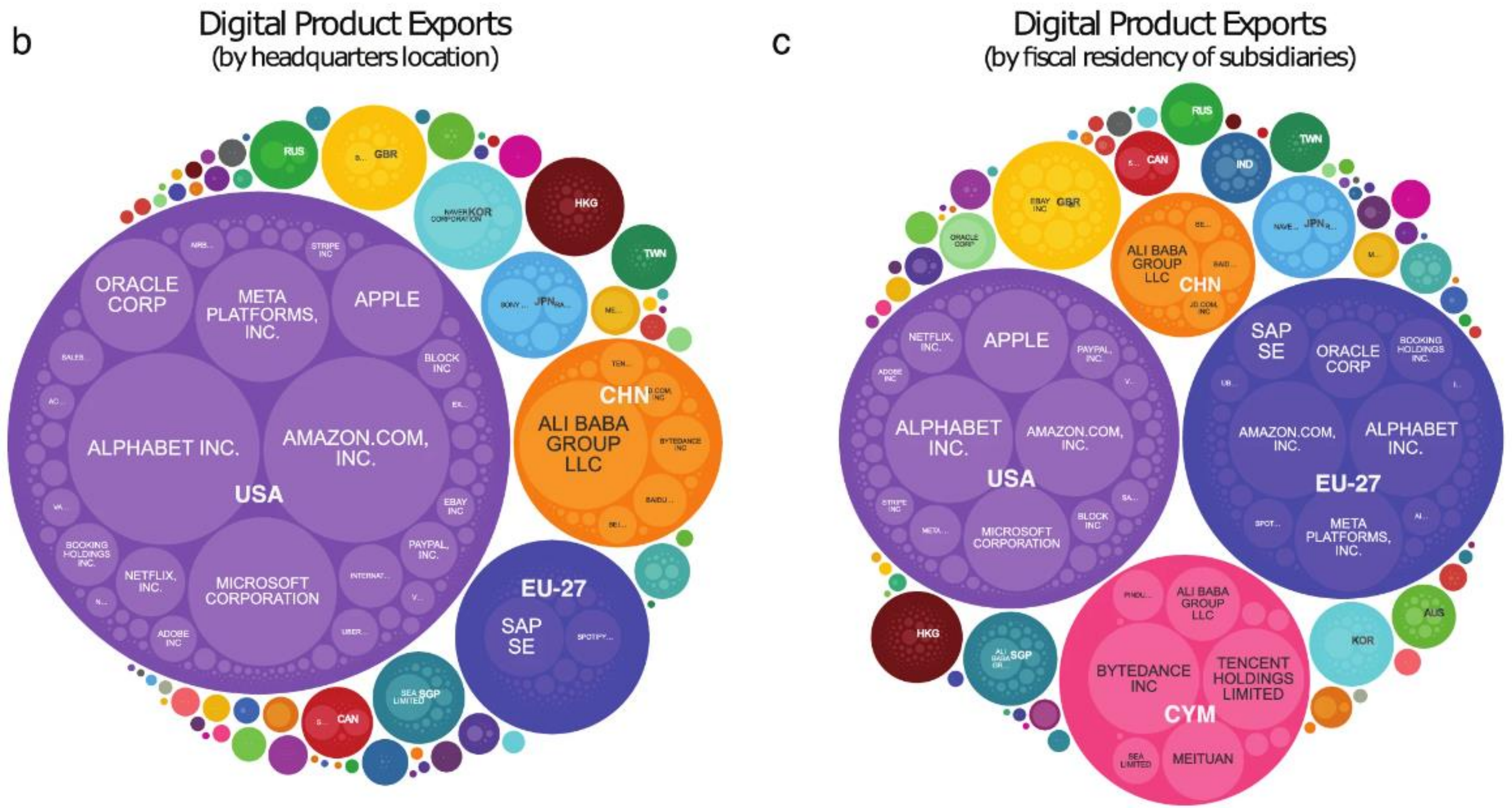
(Assigns trade to closest subsidiary constrained by total revenue amounts)

Resulting Dataset

Bilateral digital product trade estimates for 189 countries, 31 sectors, and 15,255 firms

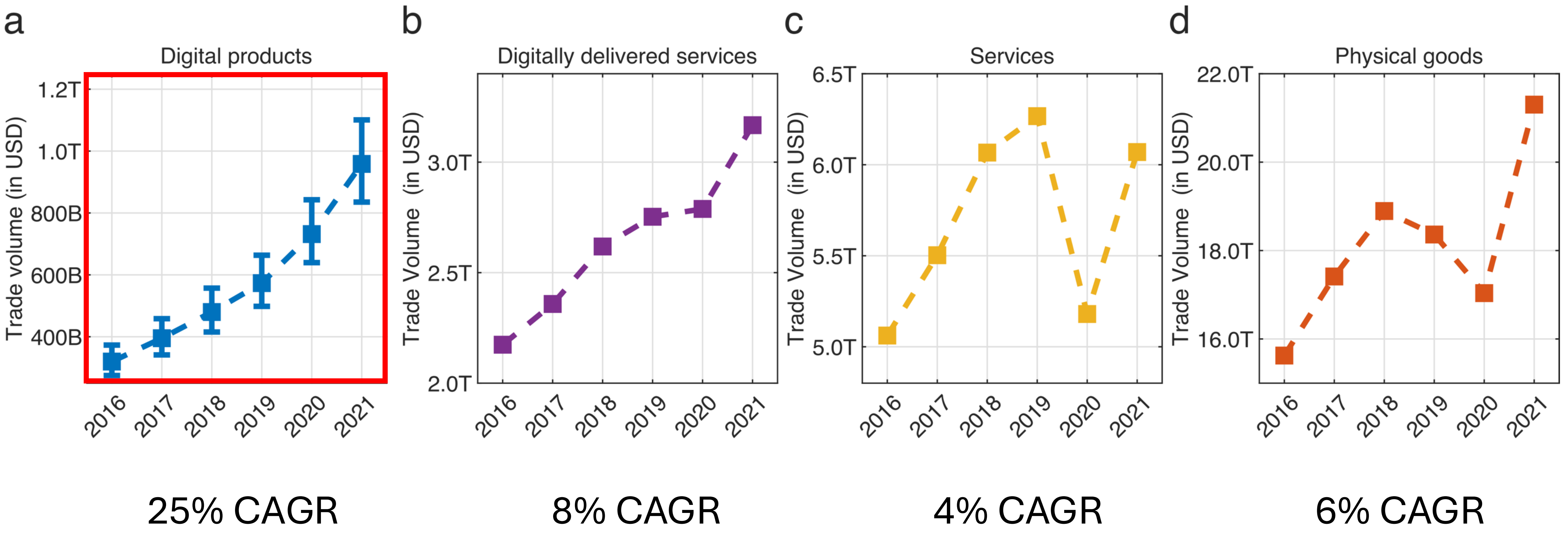
Stojkoski et al. Nature Communications (2024)

Digital Product Trade

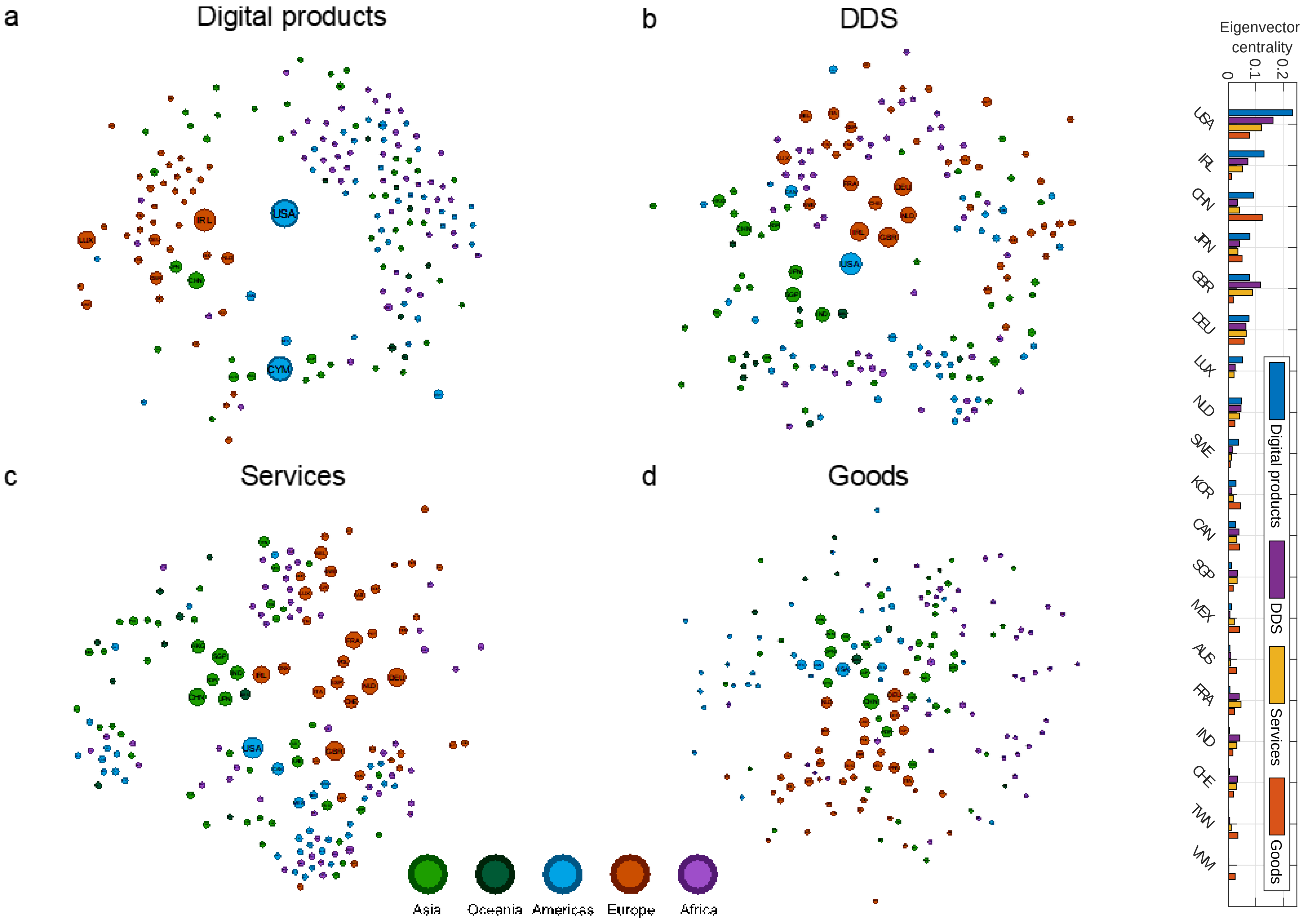


Stojkoski, Viktor, et al. "Estimating digital product trade through corporate revenue data." *Nature Communications* 15.1 (2024): 5262.

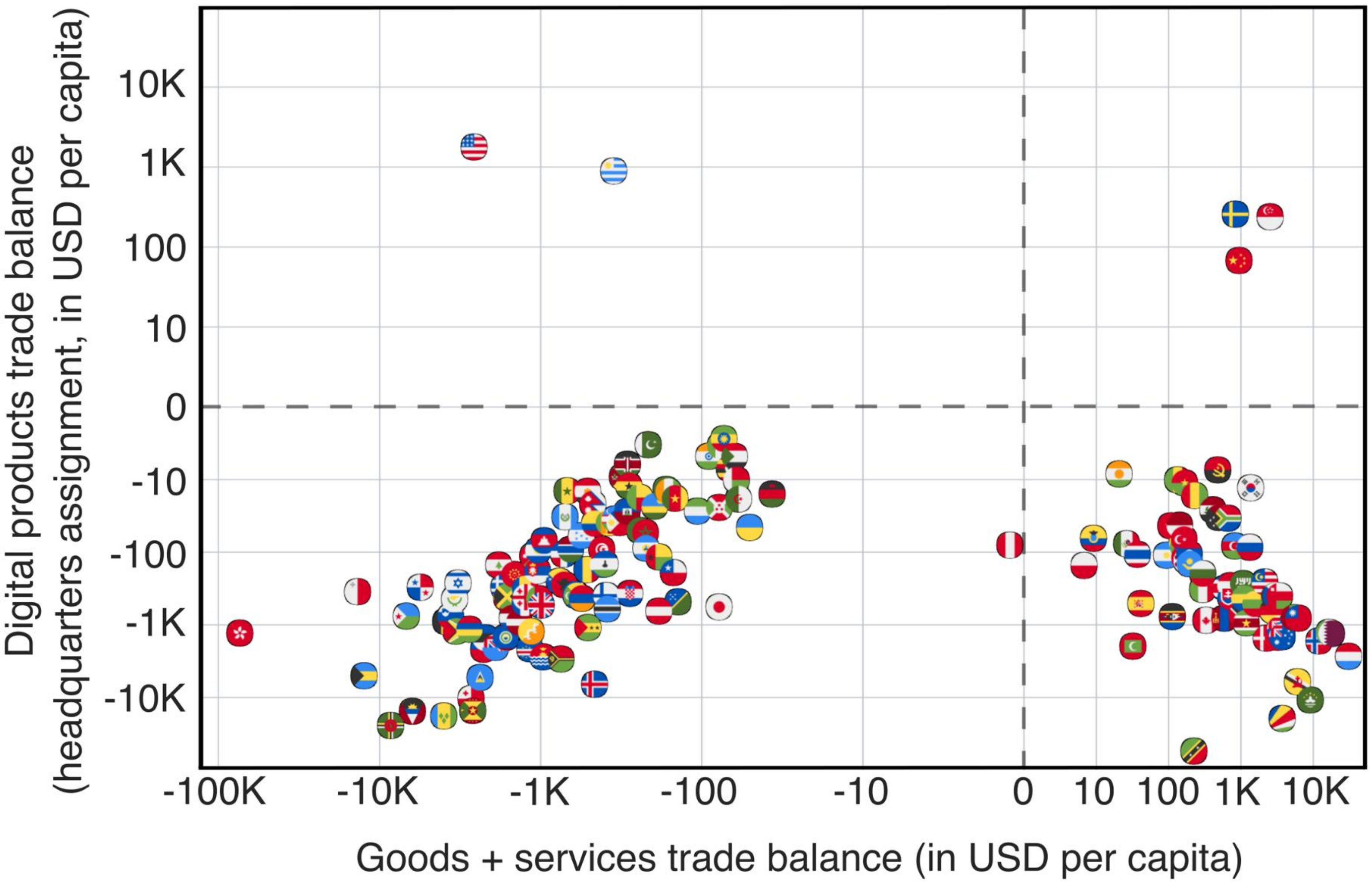
Digital Trade is Growing Fast



Stojkoski, Viktor, et al. "Estimating digital product trade through corporate revenue data." *Nature Communications* 15.1 (2024): 5262.

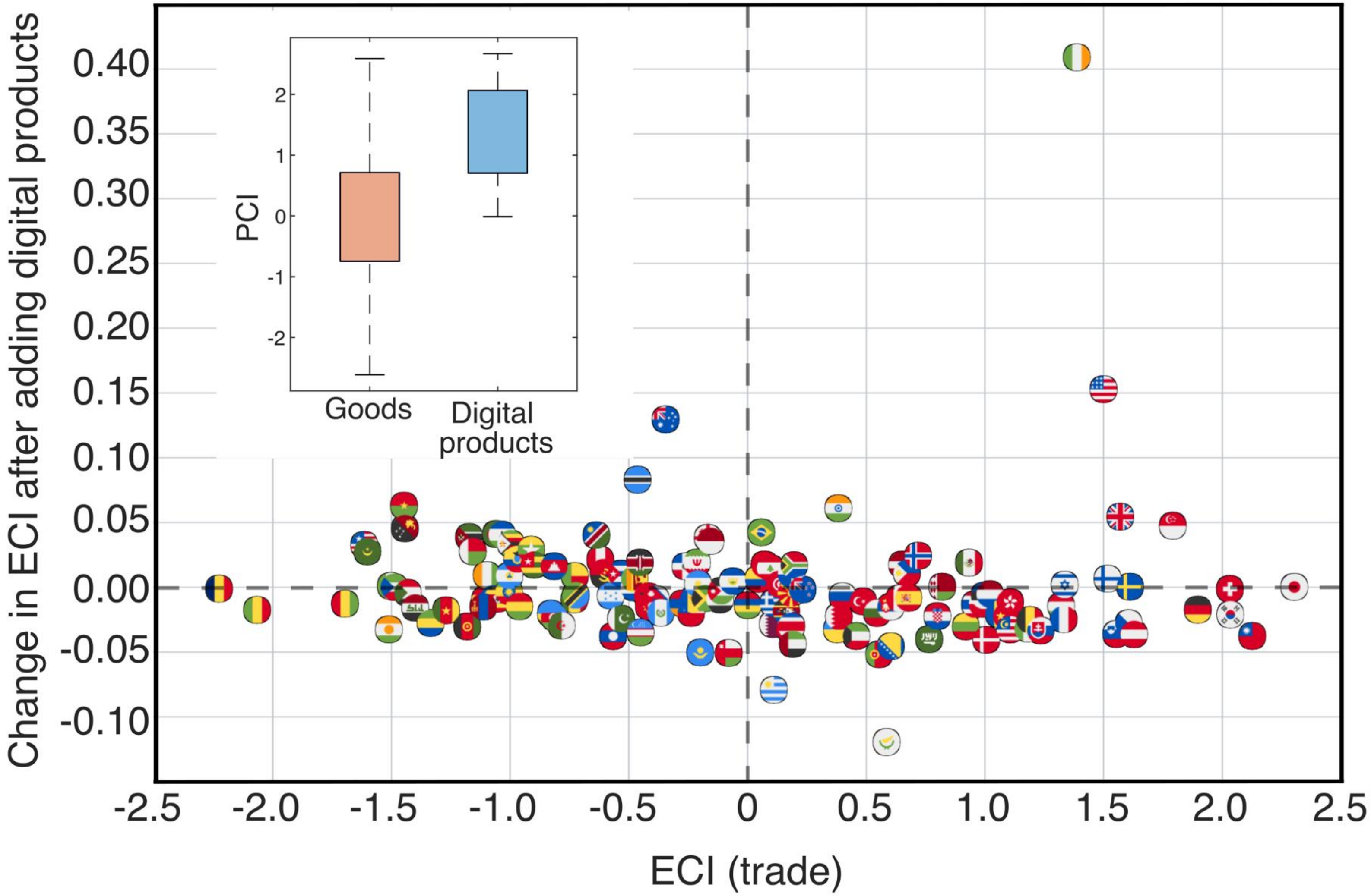

a
Digital products
USA
IRL
LUX
CYM
b
DDS
USA
c
Services
USA
d
Goods
Eigenvector centrality
0
0.1
0.2
USA
IRL
CHN
JPN
GBR
DEU
LUX
NLD
SWE
KOR
CAN
SGP
MEX
AUS
FRA
IND
CHE
TWN
VNM
Digital products
DDS
Services
Goods
Asia
Oceania
Americas
Europe
Africa

It help us rethink trade balances



Stojkoski, Viktor, et al. "Estimating digital product trade through corporate revenue data." *Nature Communications* 15.1 (2024): 5262.

Digital Trade is High Complexity



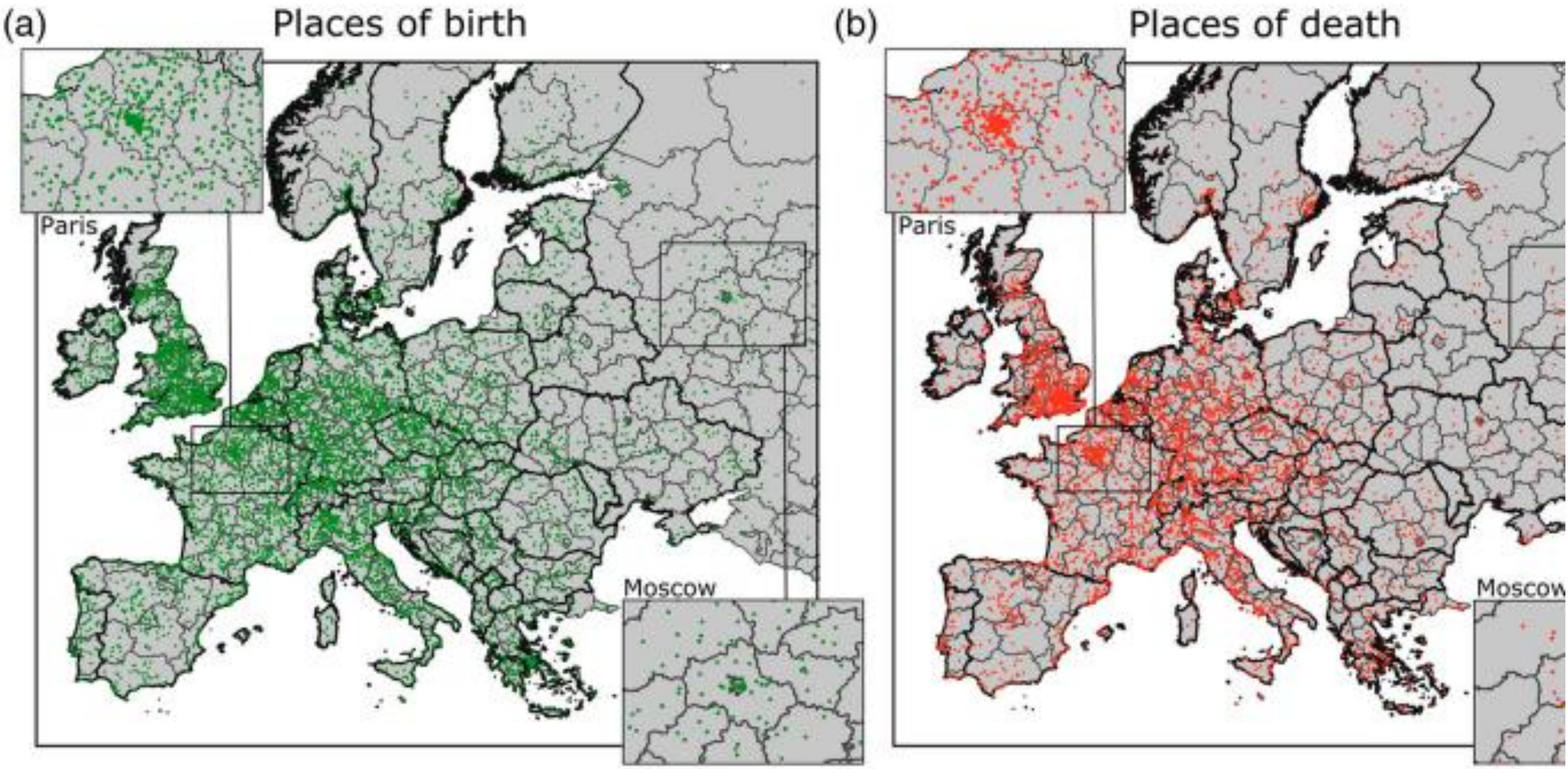
Economic complexity:

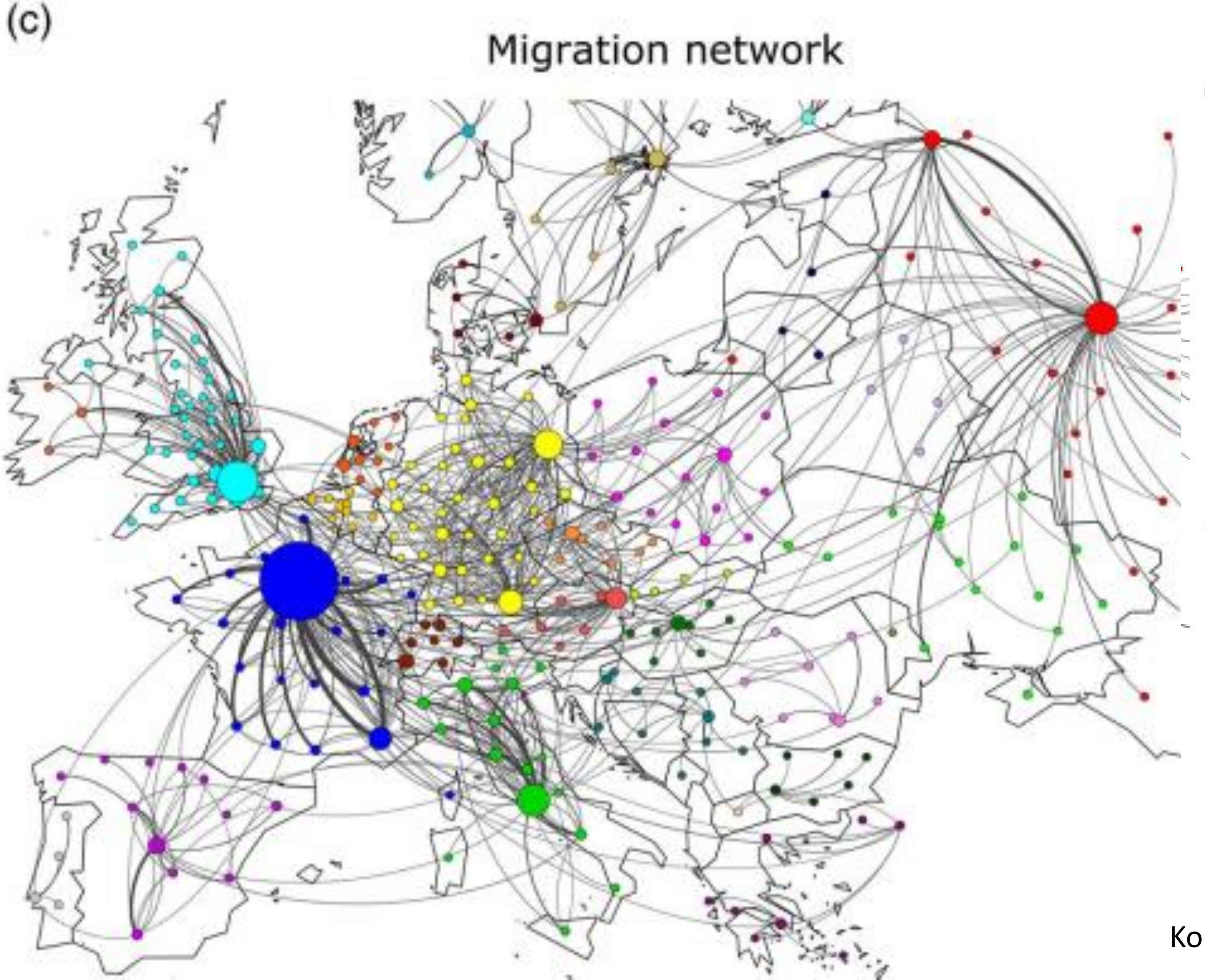
A telescope to understand the past & the future

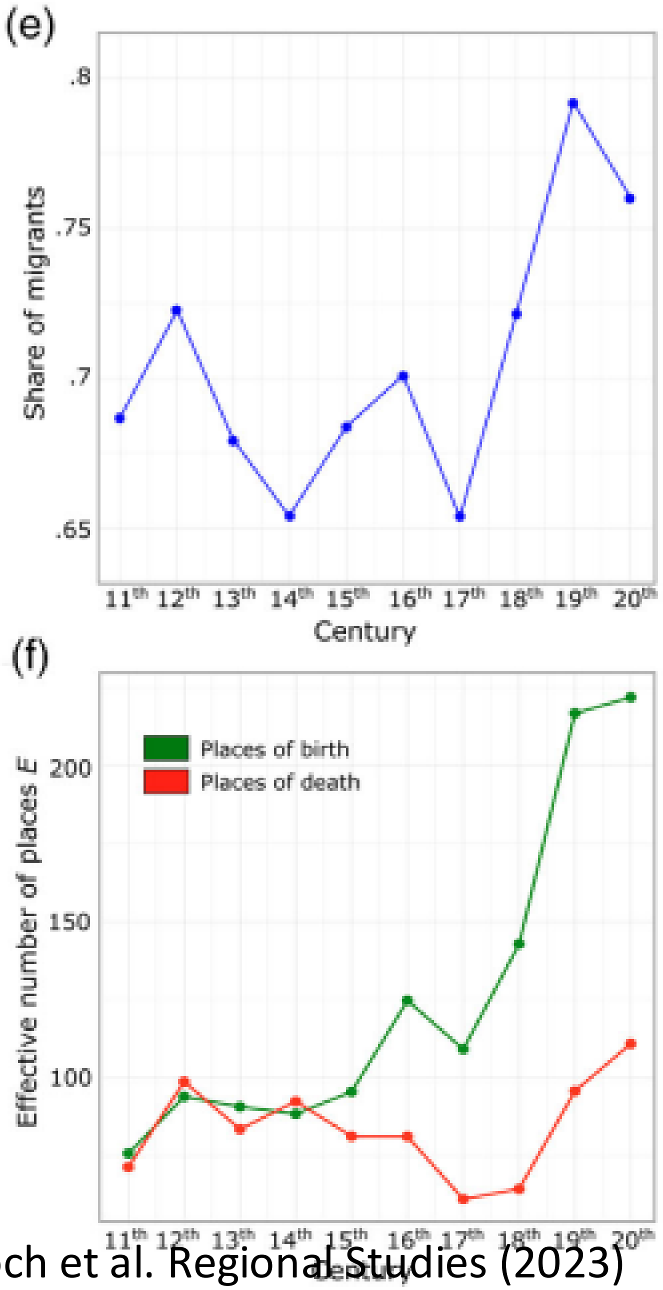
Migrants are known vectors of knowledge flows

- Agrawal, A., Cockburn, I. & McHale, J. Gone but not forgotten: knowledge flows, labor mobility, and enduring social relationships. *J Econ Geogr* **6**, 571–591 (2006).
- Balconi, M., Breschi, S. & Lissoni, F. Networks of inventors and the role of academia: an exploration of Italian patent data. *Research Policy* **33**, 127–145 (2004).
- Breschi, S. & Lissoni, F. Knowledge networks from patent data. in *Handbook of quantitative science and technology research* 613–643 (Springer, 2004).
- Singh, J. Collaborative Networks as Determinants of Knowledge Diffusion Patterns. *Management Science* **51**, 756–770 (2005).
- Moser, P., Voena, A. & Waldinger, F. German Jewish émigrés and US invention. *American Economic Review* **104**, 3222–3255 (2014).
- Waldinger, F. Peer Effects in Science: Evidence from the Dismissal of Scientists in Nazi Germany. *The Review of Economic Studies* **79**, 838–861 (2012).
- Waldinger, F. Quality matters: The expulsion of professors and the consequences for PhD student outcomes in Nazi Germany. *Journal of political economy* **118**, 787–831 (2010).
- Ganguli, I. Immigration and ideas: what did Russian scientists "bring" to the United States? *Journal of Labor Economics* **33**, S257–S288 (2015).
- Borjas, G. J. & Doran, K. B. Cognitive mobility: labor market responses to supply shocks in the space of ideas. *Journal of Labor Economics* **33**, S109–S145 (2015).
- Azoulay, P., Liu, C. C. & Stuart, T. E. Social influence given (partially) deliberate matching: Career imprints in the creation of academic entrepreneurs. *American Journal of Sociology* **122**, 1223–1271 (2017).
- Collins, H. M. The TEA Set: Tacit Knowledge and Scientific Networks. *Science Studies* **4**, 165–185 (1974).
- Parsons, C. & Vézina, P.-L. Migrant Networks and Trade: The Vietnamese Boat People as a Natural Experiment. *Economic Journal* (2017).
- Scoville, W. C. The Huguenots and the diffusion of technology. I. *Journal of Political Economy* **60**, 294–311 (1952).
- Scoville, W. C. The Huguenots and the diffusion of technology. II. *Journal of Political Economy* **60**, 392–411 (1952).
- Kerr, W. R. *The Gift of Global Talent: How Migration Shapes Business, Economy & Society*. (Stanford University Press, 2018).
- Lissoni, F. & Miguelez, E. Migration and Innovation: Learning from Patent and Inventor Data. *Journal of Economic Perspectives* **38**, 27–54 (2024).
- Autant-Bernard, C. Science and knowledge flows: evidence from the French case. *Research policy* **30**, 1069–1078 (2001).

But most research focus on a recent context and on knowledge flows within an activity (MAR spillovers)


(a)
Places of birth
Paris
Moscow
(b)
Places of death
Paris
Moscow





Koch et al. Regional Studies (2023)

Knowledge spillovers can occur within and across both locations and activities:

- Spillovers across locations within the same activity:

  Immigrant mathematicians -> P(region begets mathematicians)

- Spillovers across both locations and activities:

  Immigrant physicists -> P(region begets mathematicians)

- Spillovers within locations across activities:

  Local physicists -> P(region begets mathematicians)

Estimate direct and relatedness effects

E..g. Do famous immigrant mathematicians contribute to the birth of new famous mathematicians and physicists?

Step 1: Estimate "excess" migration

$$R_{ik,t}^{immi} = \frac{N_{ik,t}^{immi}}{\hat{N}_{ik,t}^{immi}}$$

$$R_{ik,t}^{emi} = \frac{N_{ik,t}^{emi}}{\hat{N}_{ik,t}^{emi}},$$

$$\hat{N}_{ik,t} = f(\alpha_0 + \alpha_1 N_{ik,t-1} + \alpha_2 S_{ik,t-1}^{births} + \theta_{it} + \vartheta_{kt}),$$

$$M_{ik,t}^{immi} = \begin{cases} 1 & if\ R_{ik,t}^{immi} \geq 1 \\ 0 & otherwise \end{cases}$$

$$M_{ik,t}^{emi} = \begin{cases} 1 & if\ R_{ik,t}^{emi} \geq 1 \\ 0 & otherwise. \end{cases}$$

Step 2: Estimate separate relatedness for immigrants, emigrants, and locals

$$\omega_{ik,t}^{immi} = \frac{\sum_{k'} M_{ik',t}^{immi} \varphi_{kk',t}^{immi}}{\sum_{k'} \varphi_{kk',t}^{immi}},$$

$$\omega_{ik,t}^{emi} = \frac{\sum_{k'} M_{ik',t}^{emi} \varphi_{kk',t}^{emi}}{\sum_{k'} \varphi_{kk',t}^{emi}}, \qquad (9)$$

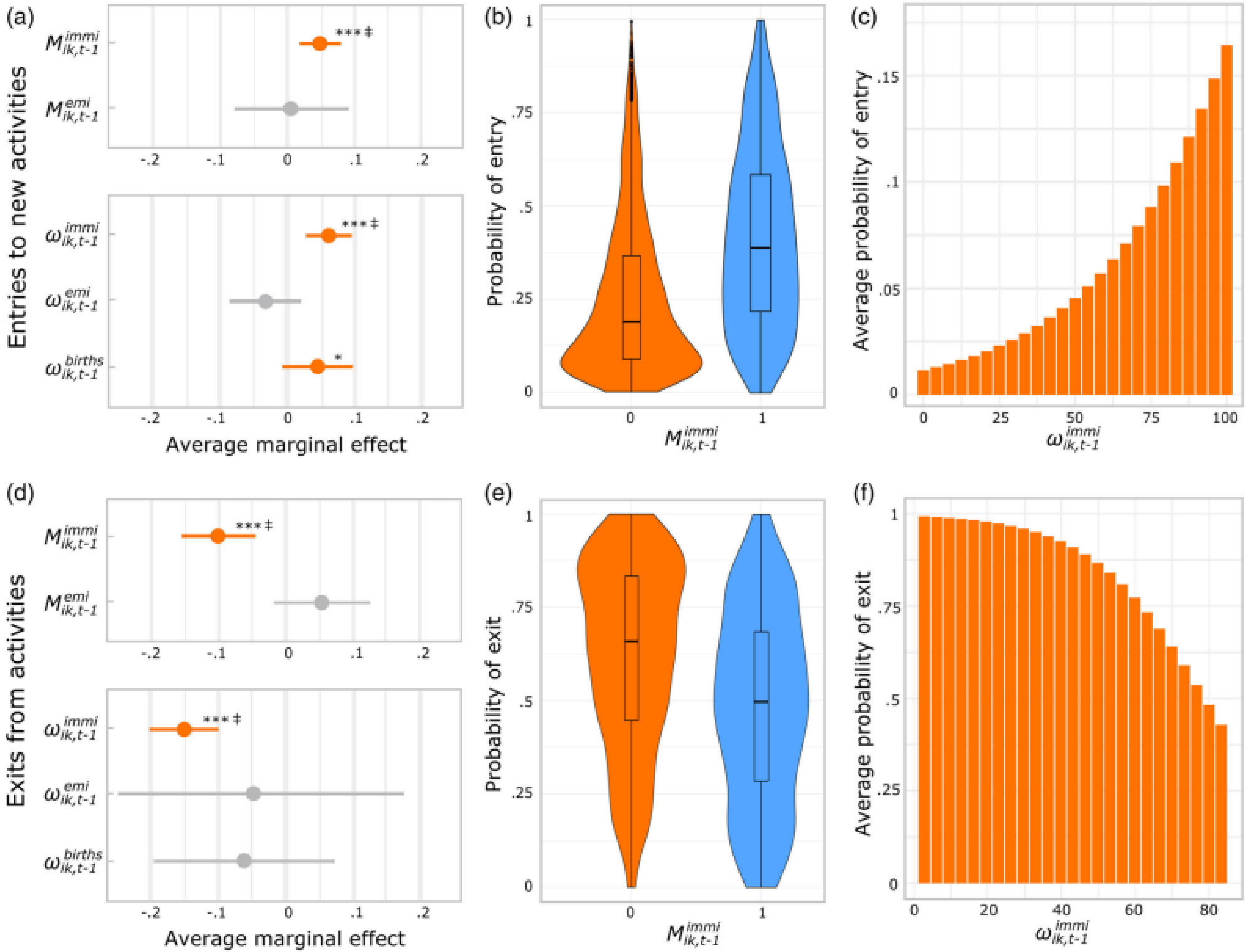
$$\omega_{ik,t}^{births} = \frac{\sum_{k'} M_{ik',t}^{births} \varphi_{kk',t}^{births}}{\sum_{k'} \varphi_{kk',t}^{births}}.$$

Koch et al. Regional Studies (2023)

Table 1. Main results of logistic regression models explaining entries and exits of activities.

	Dependent variable: $Entry_{ik,t}$					Dependent variable: $Exit_{ik,t}$				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$M^{immi}_{ik,t-1}$	0.334***	0.303***	0.336***	0.331***	0.300***	−0.603***	−0.584***	−0.591***	−0.587***	−0.571***
	(0.080)	(0.075)	(0.086)	(0.080)	(0.076)	(0.127)	(0.134)	(0.120)	(0.126)	(0.126)
$M^{emi}_{ik,t-1}$	0.115	0.045	0.106	0.121	0.018	0.310	0.330	0.233	0.306	0.291
	(0.261)	(0.278)	(0.261)	(0.255)	(0.270)	(0.240)	(0.232)	(0.216)	(0.222)	(0.203)
$\omega^{immi}_{ik,t-1}$		0.027***			0.028***		−0.067***			−0.064***
		(0.006)			(0.007)		(0.016)			(0.011)
$\omega^{emi}_{ik,t-1}$			−0.006		−0.024			−0.048		−0.025
			(0.012)		(0.019)			(0.038)		(0.063)
$\omega^{births}_{ik,t-1}$				0.011	0.027*				−0.059***	−0.034
				(0.008)	(0.015)				(0.018)	(0.041)
Further controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Fixed effects										
Broad category–region–century	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Category–century	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	3944	3944	3944	3944	3944	1051	1051	1051	1051	1051
Pseudo-R^2	0.213	0.214	0.213	0.213	0.215	0.224	0.230	0.226	0.226	0.232
BIC	9537.0	9539.4	9545.0	9544.5	9553.1	3619.6	3618.0	3623.4	3623.3	3628.8

Note: The fixed effects in these models are highly restrictive, amounting to more than 700 parameters in columns (1) to (5) and more than 350 parameters in columns (6) to (10). All regions included in the regression model exhibit a minimum number of births and migrants such that measures of specialisation and relatedness are defined (see Section 2.2 in the supplemental data online). Standard errors are clustered by region and period. For the full regression tables with all control variables, see Sections 3.2 and 3.3 online. BIC, Bayesian information criterion. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.



(a)
(b)
(c)
(d)
(e)
(f)
Entries to new activities
Exits from activities
$M^{immi}_{ik,t-1}$
$M^{emi}_{ik,t-1}$
$\omega^{immi}_{ik,t-1}$
$\omega^{emi}_{ik,t-1}$
$\omega^{births}_{ik,t-1}$
*** ‡
*
Average marginal effect
Probability of entry
Average probability of entry
Probability of exit
Average probability of exit

PNAS

RESEARCH ARTICLE | ECONOMIC SCIENCES

OPEN ACCESS

Augmenting the availability of historical GDP per capita estimates through machine learning

Philipp Koch[a,b,1], Viktor Stojkoski[a,c], and César A. Hidalgo[a,d,e,1]

Affiliations are included on p. 10.

Edited by Marshall Burke, Stanford University, Stanford, CA; received January 31, 2024; accepted August 9, 2024 by Editorial Board Member Ronald D. Lee

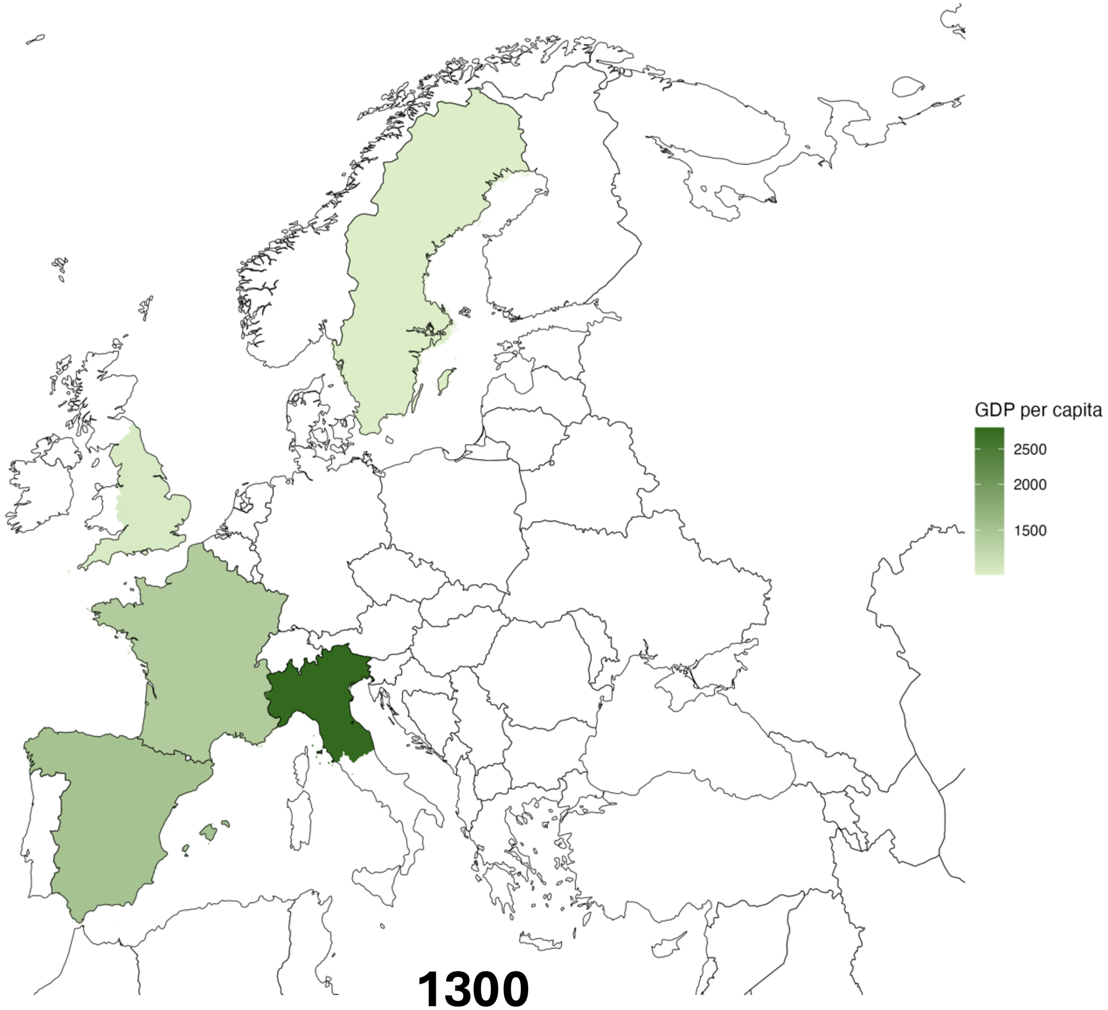
Can we use data on the biographies of historical figures to estimate the GDP per capita of countries and regions? Here, we introduce a machine learning method to estimate the GDP per capita of dozens of countries and hundreds of regions in Europe and North America for the past seven centuries starting from data on the places of birth, death, and occupations of hundreds of thousands of historical figures. We build an elastic net regression model to perform feature selection and generate out-of-sample estimates that explain 90% of the variance in known historical income levels. We use this model to generate GDP per capita estimates for countries, regions, and time periods for which these data are not available and externally validate our estimates by comparing them with four proxies of economic output: urbanization rates in the past 500 y, body height in the 18th century, well-being in 1850, and church building activity in the 14th and 15th century. Additionally, we show our estimates reproduce the well-known reversal of fortune between southwestern and northwestern Europe between 1300 and 1800 and find this is largely driven by countries and regions engaged in Atlantic trade. These findings validate the use of fine-grained biographical data as a method to augment historical GDP per capita estimates. We publish our estimates with CI together with all collected source data in a comprehensive dataset.

economic history | machine learning | economic development

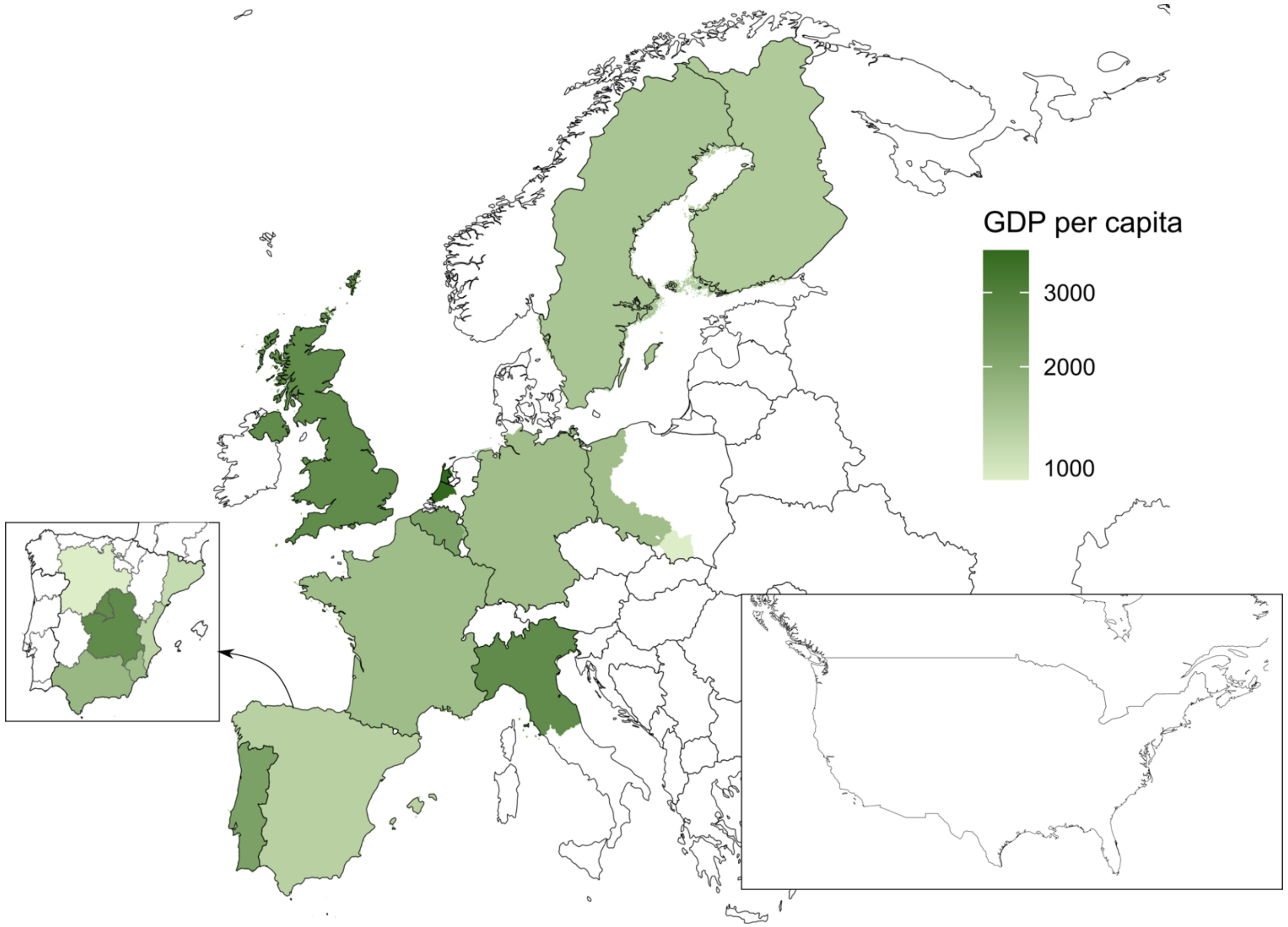
Significance

The scarcity of historical GDP per capita data limits our ability to explore questions of long-term economic development. Here, we introduce a machine learning method using detailed data on famous biographies to estimate the historical GDP per capita of hundreds of regions in Europe and North America. Our model generates accurate out-of-sample estimates (R^2 = 90%) that quadruple the availability of historical GDP per capita data and correlate positively with

Estimating Historical GDPpc

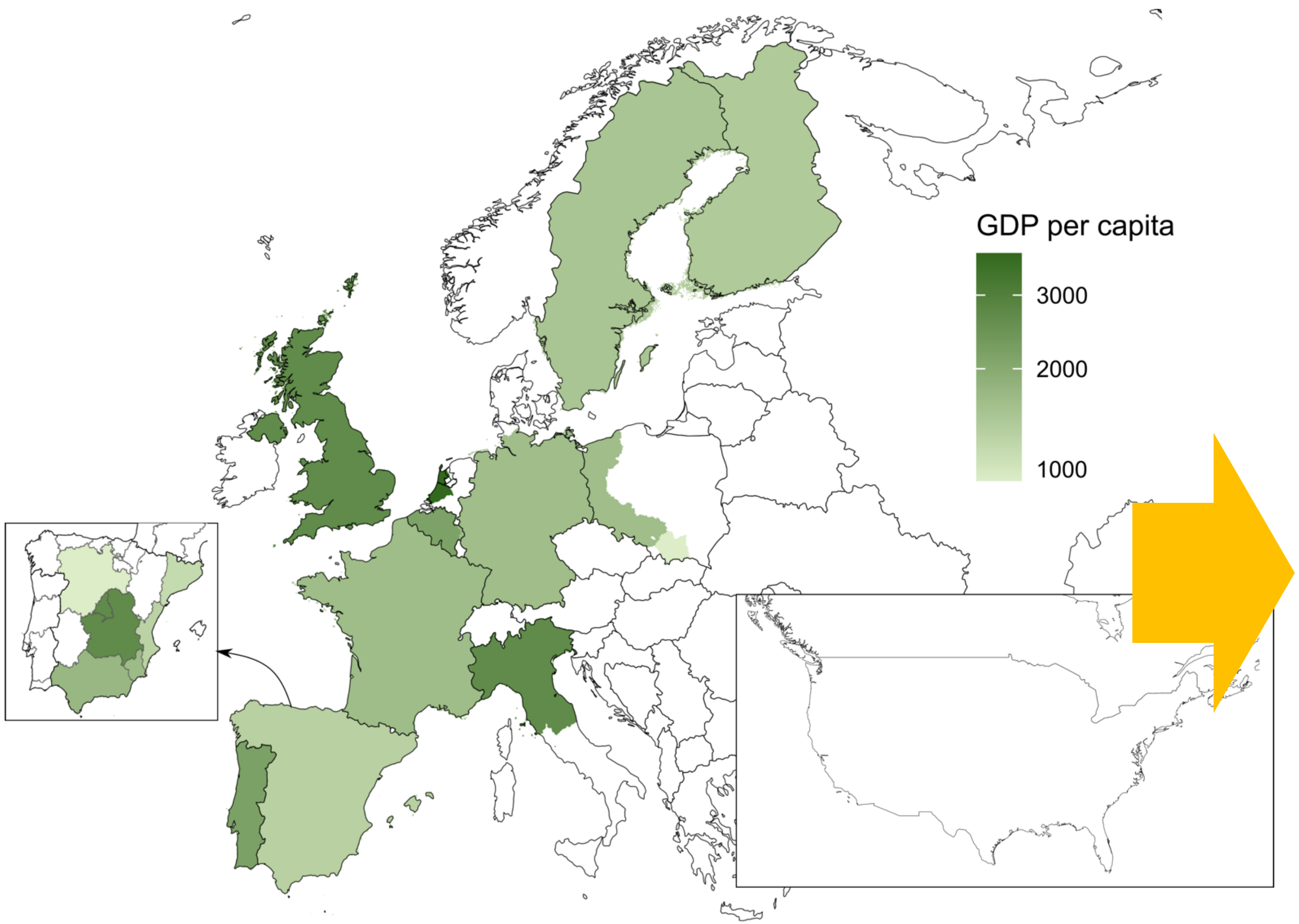


1300
Maddison project

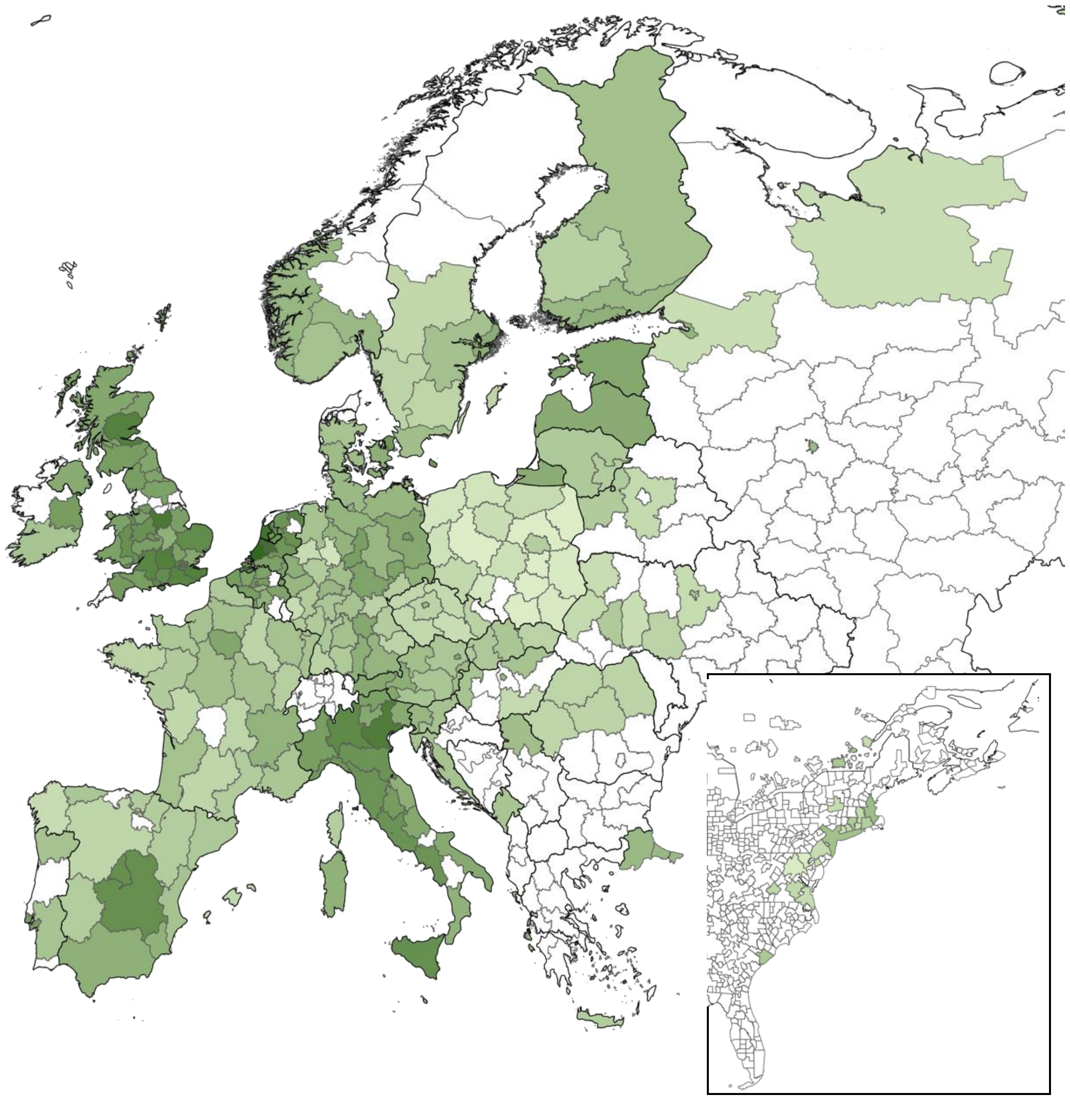


1750
Maddison project

Getting from here… … to here



1750
Maddison project



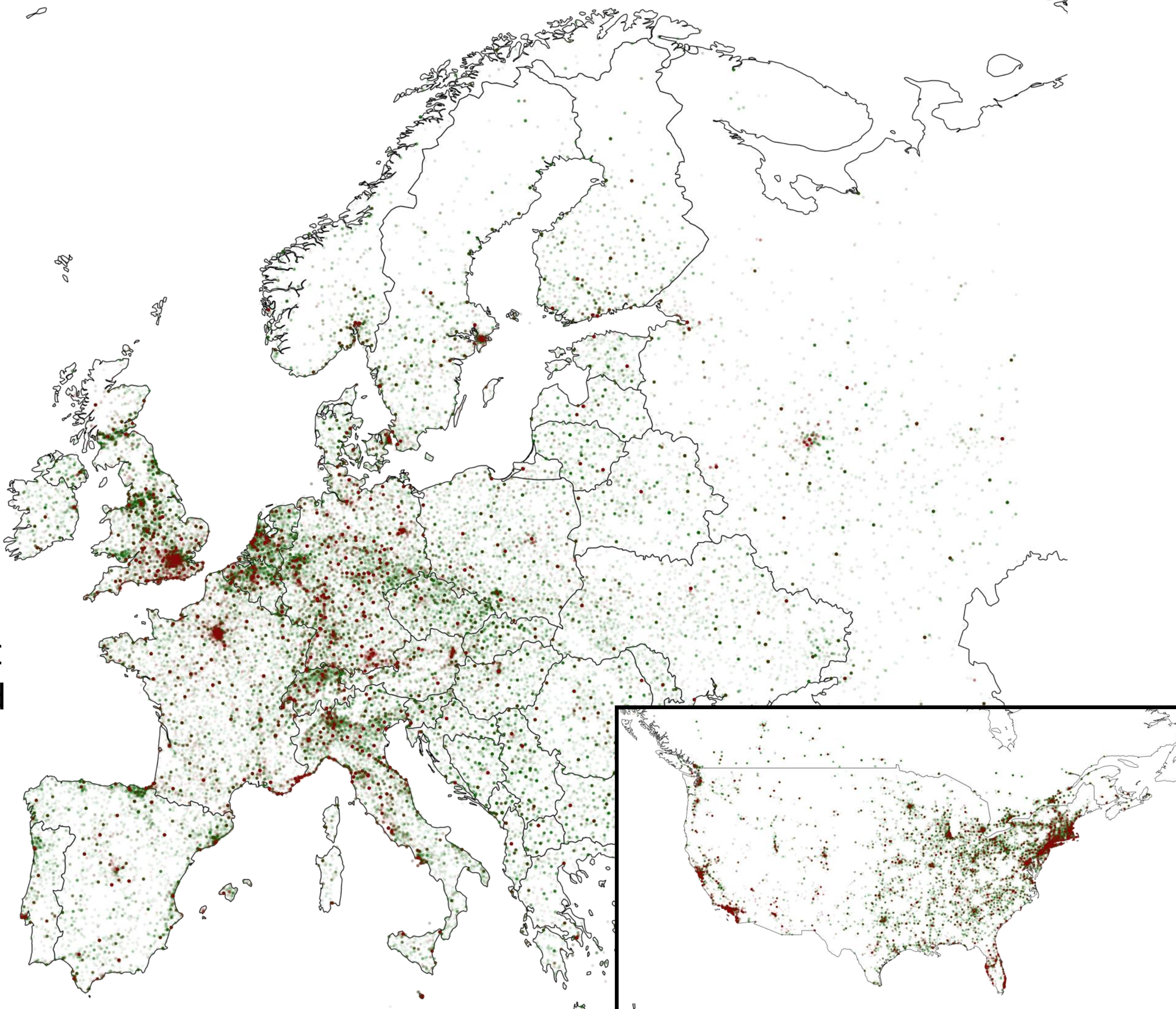
1750
Koch et al. PNAS 2024

Using This…

Place of birth, death, and occupation data of famous individuals from Wikipedia[9].

~561k famous individuals assigned to one of 49 occupations with a birth or death in Europe and North America between 1300 and 2000 (only individuals with at least 2 language editions and an identifiable occupation).

Koch, Stojkoski, Hidalgo, PNAS (2024)

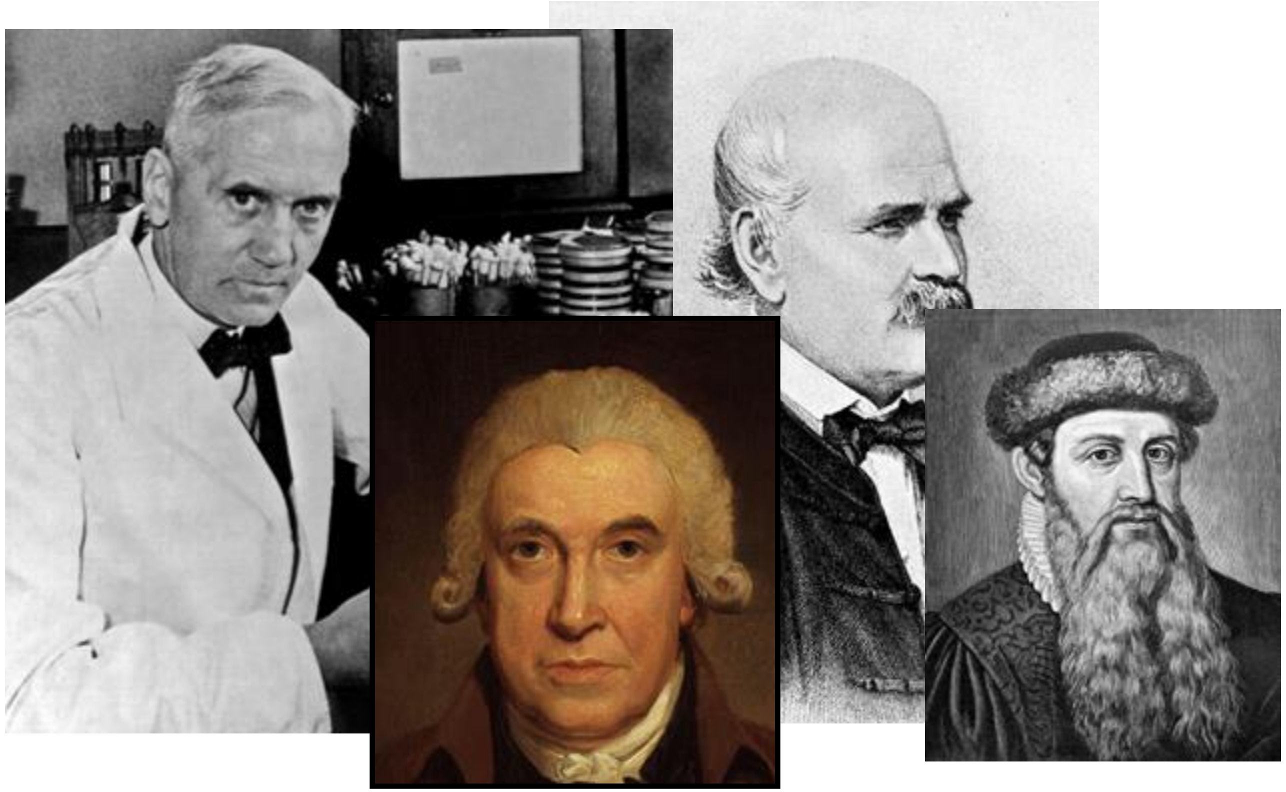


Why biographies?

Our collective memory on famous individuals is likely one of the most comprehensive representation of the historical geography of knowledge.

The famous individuals that were born at, have died at, immigrated to or emigrated from a specific place tell us something about the level of economic development.

Direct



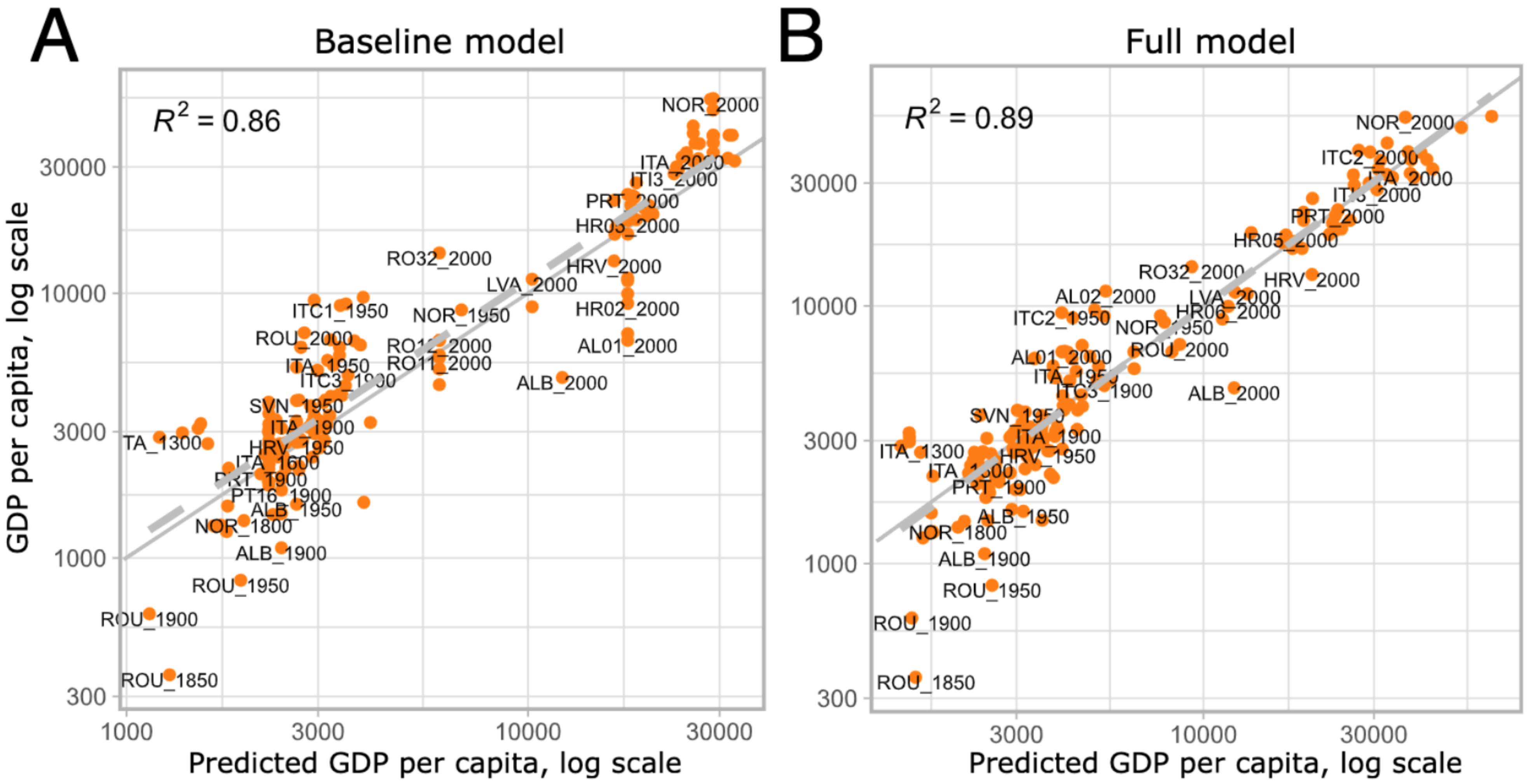
Indirect

Model

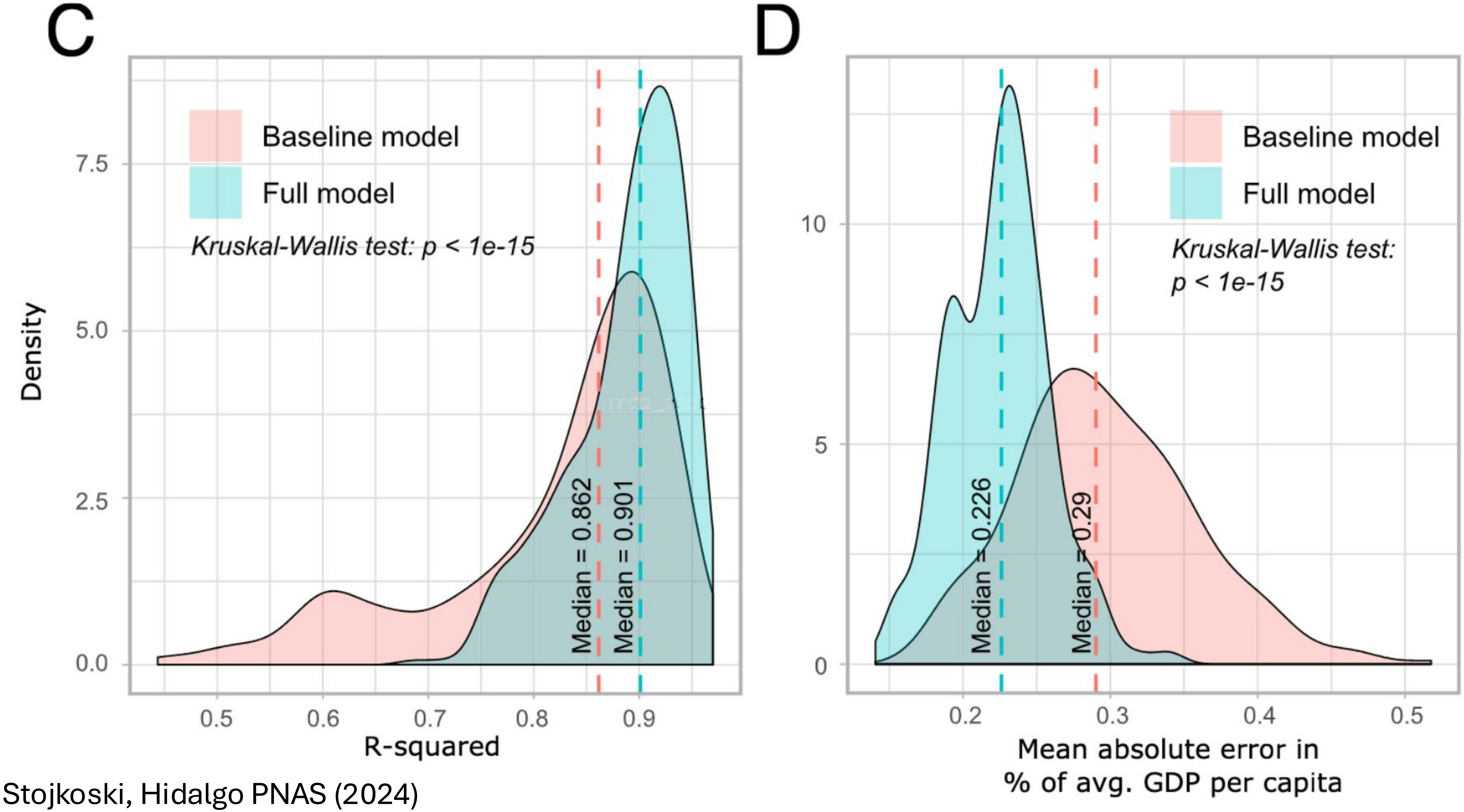
Regularized Elastic Net

Leave 20% out-of-sample cross validation

$$\widehat{\beta_{EN}} = min_{\beta}(\|y - X\beta\|^2 + \lambda[(1-\alpha)\|\beta\|_2^2 + \alpha\|\beta\|_1])$$



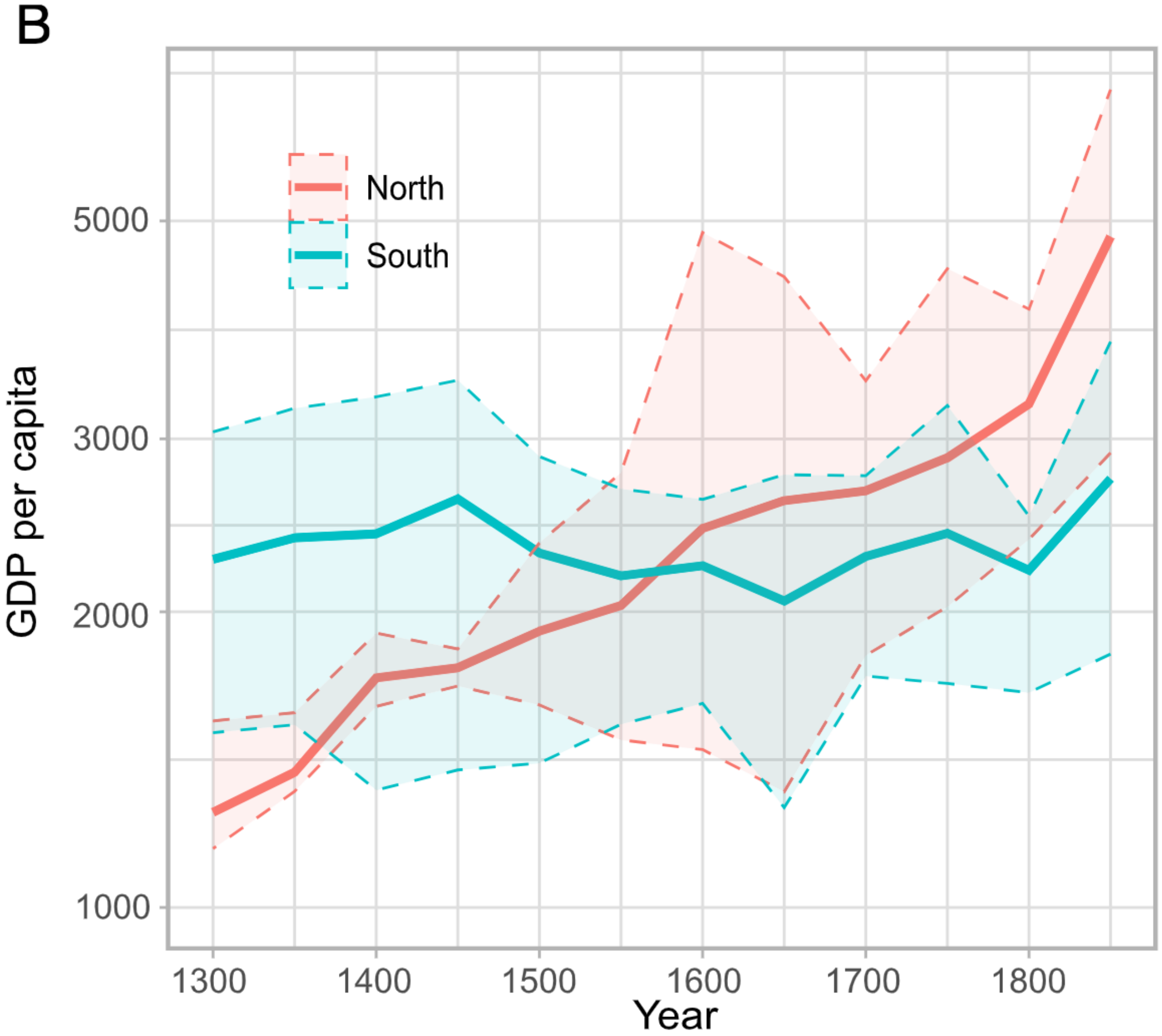
Model Performance



Koch, Stojkoski, Hidalgo PNAS (2024)

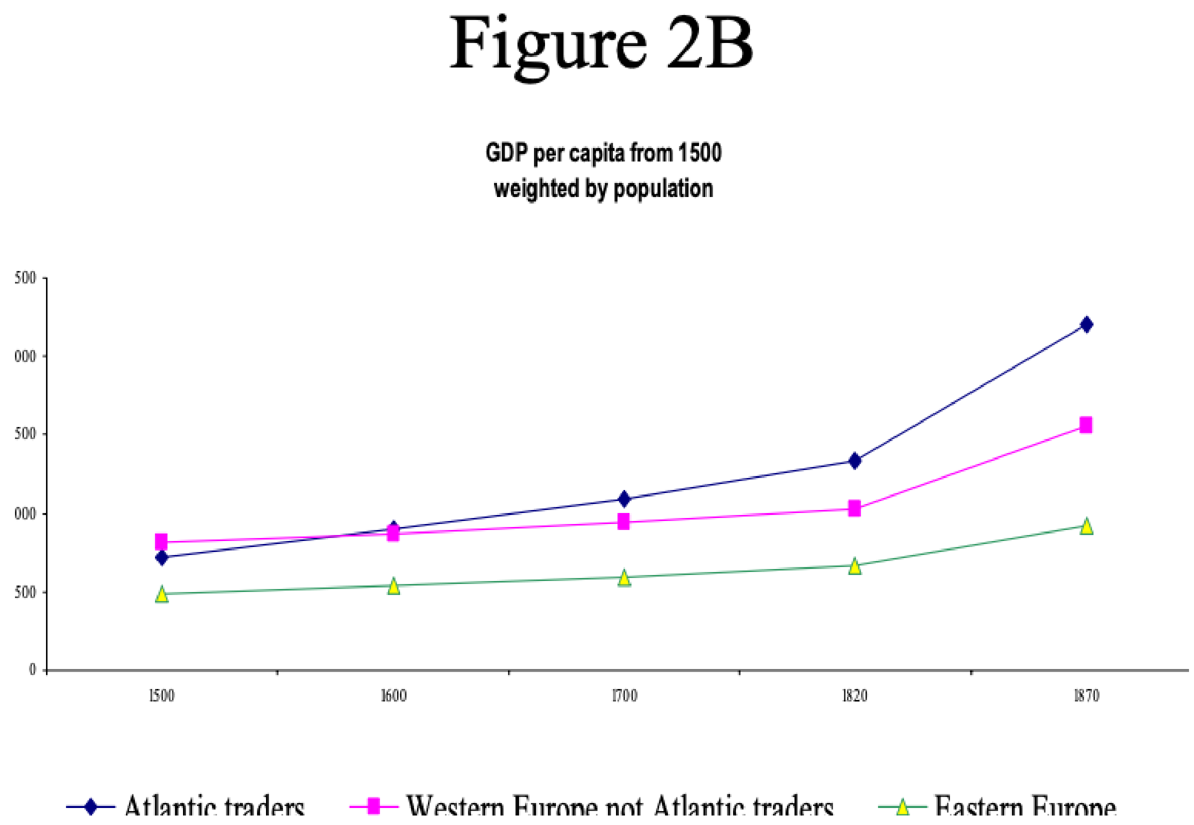
Validation - Little Divergence

In 1300, the bottom 10th percentile of the South has been as rich as the top 90th percentile of the North.
In 1800, the opposite holds: The bottom 10th percentile of the North exhibits a similar income level as the 90th percentile of the South.

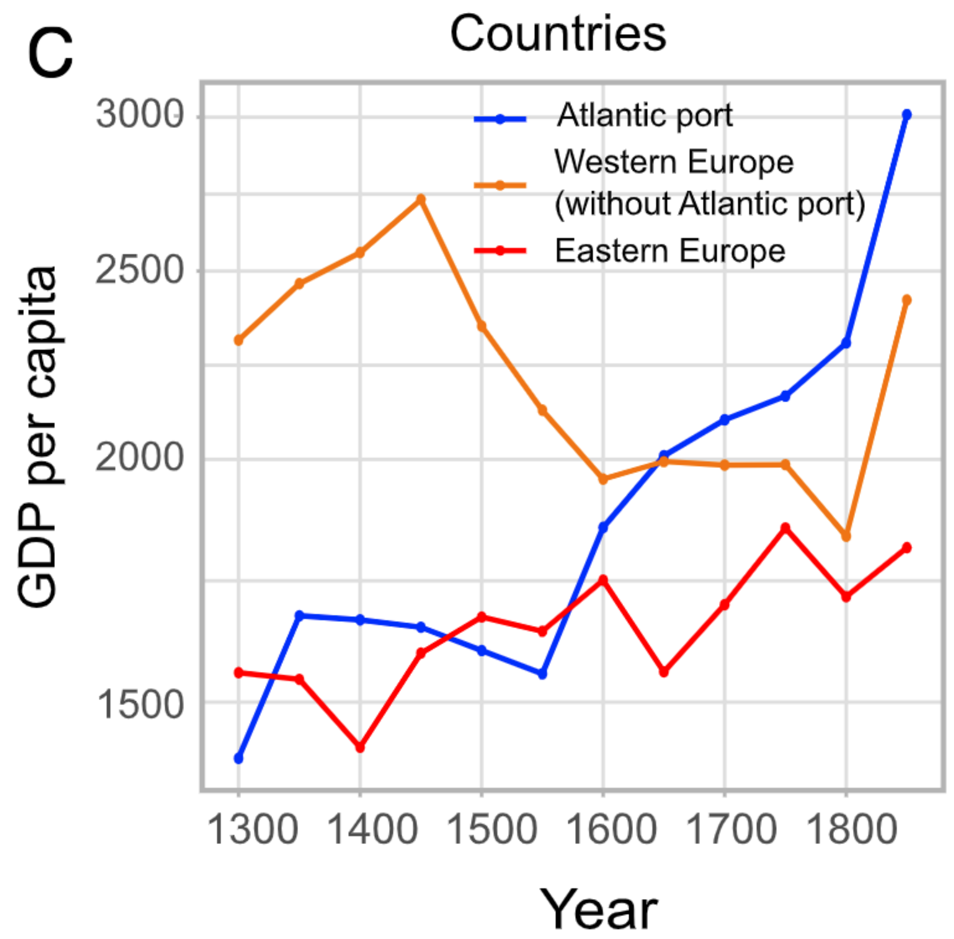


Koch, Stojkoski, Hidalgo PNAS (2024)

Validation – Acemoglu, Johnson, Robinson Atlantic Trade

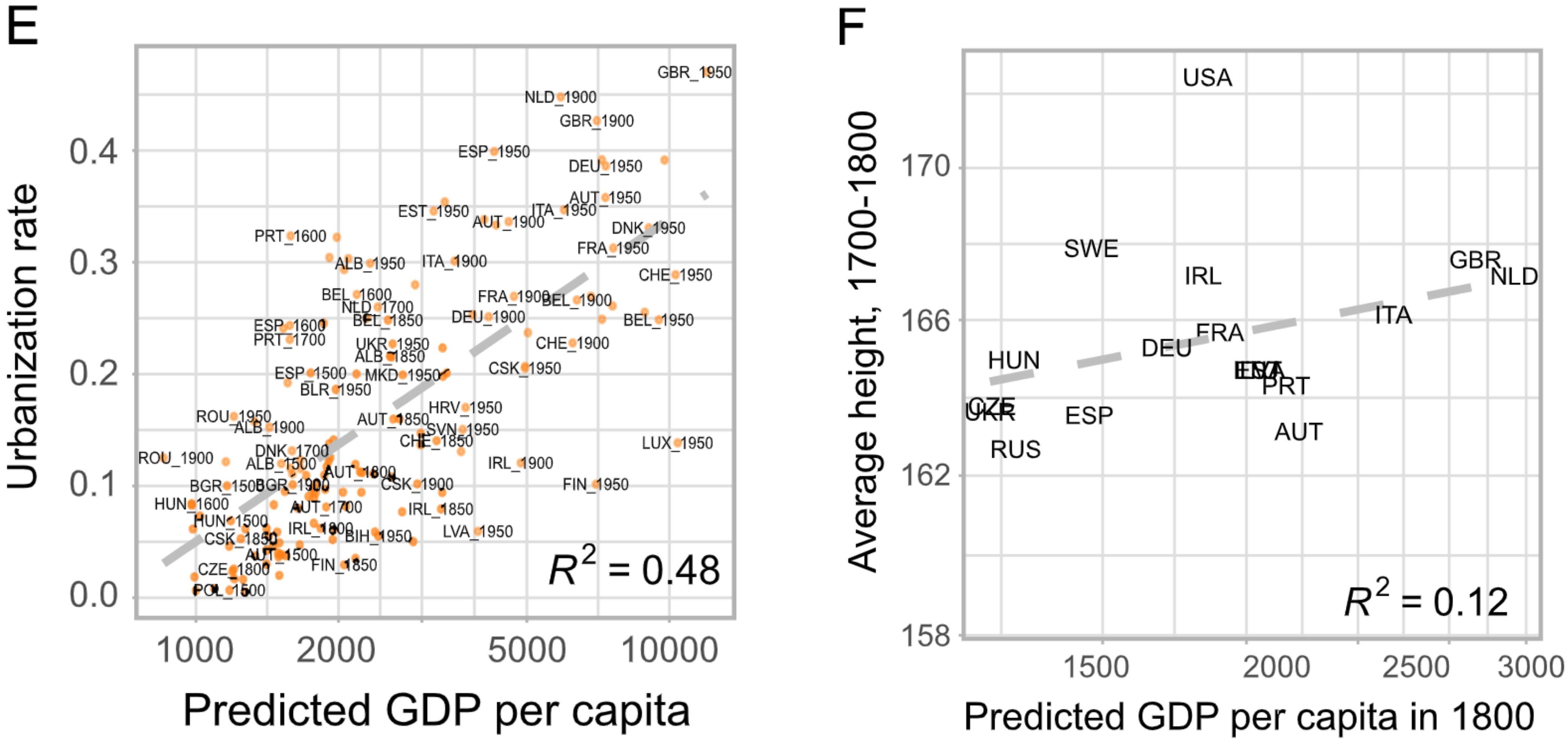


Acemoglu et al. 2005



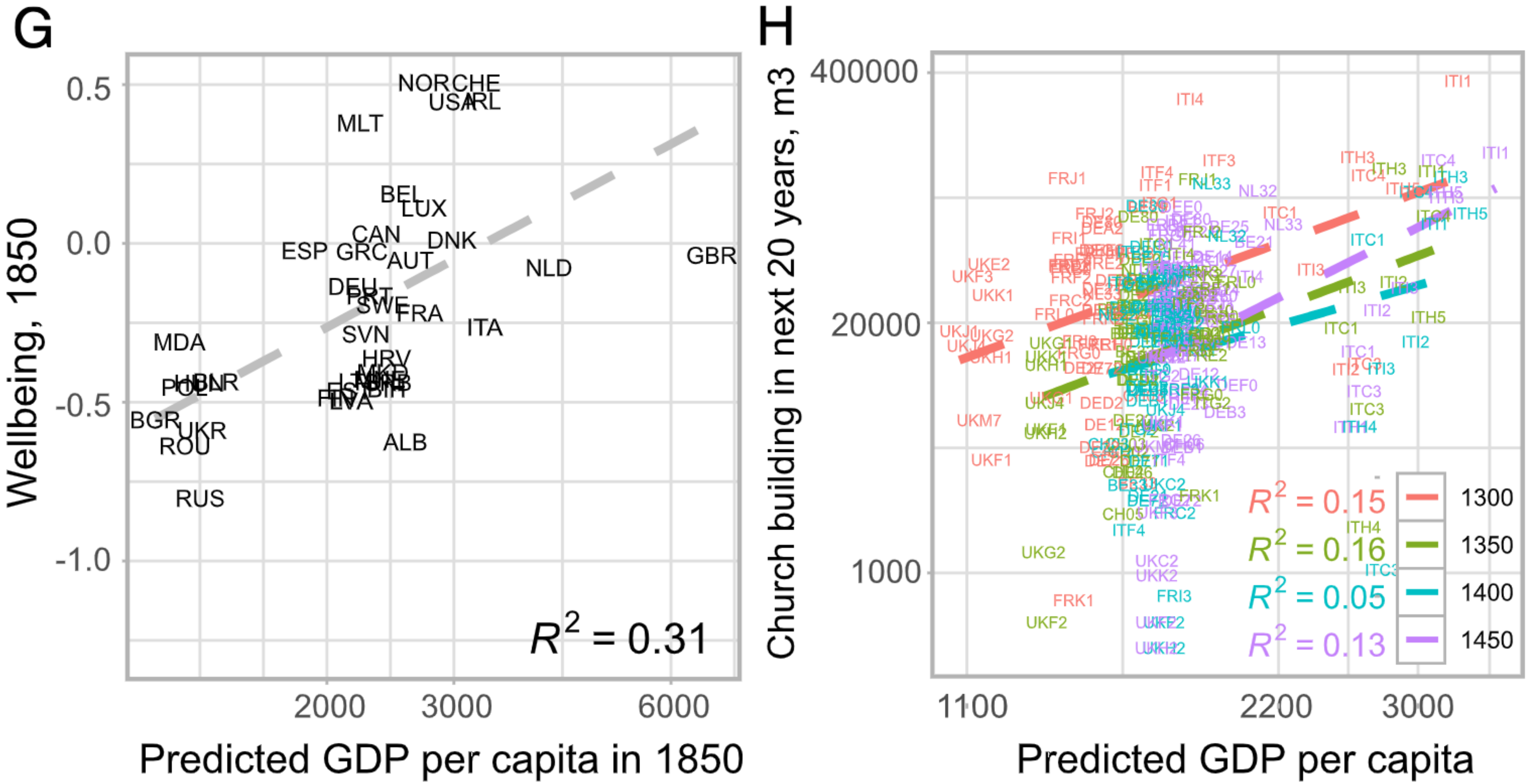
Koch et al. 2024

Validation - proxies of economic development



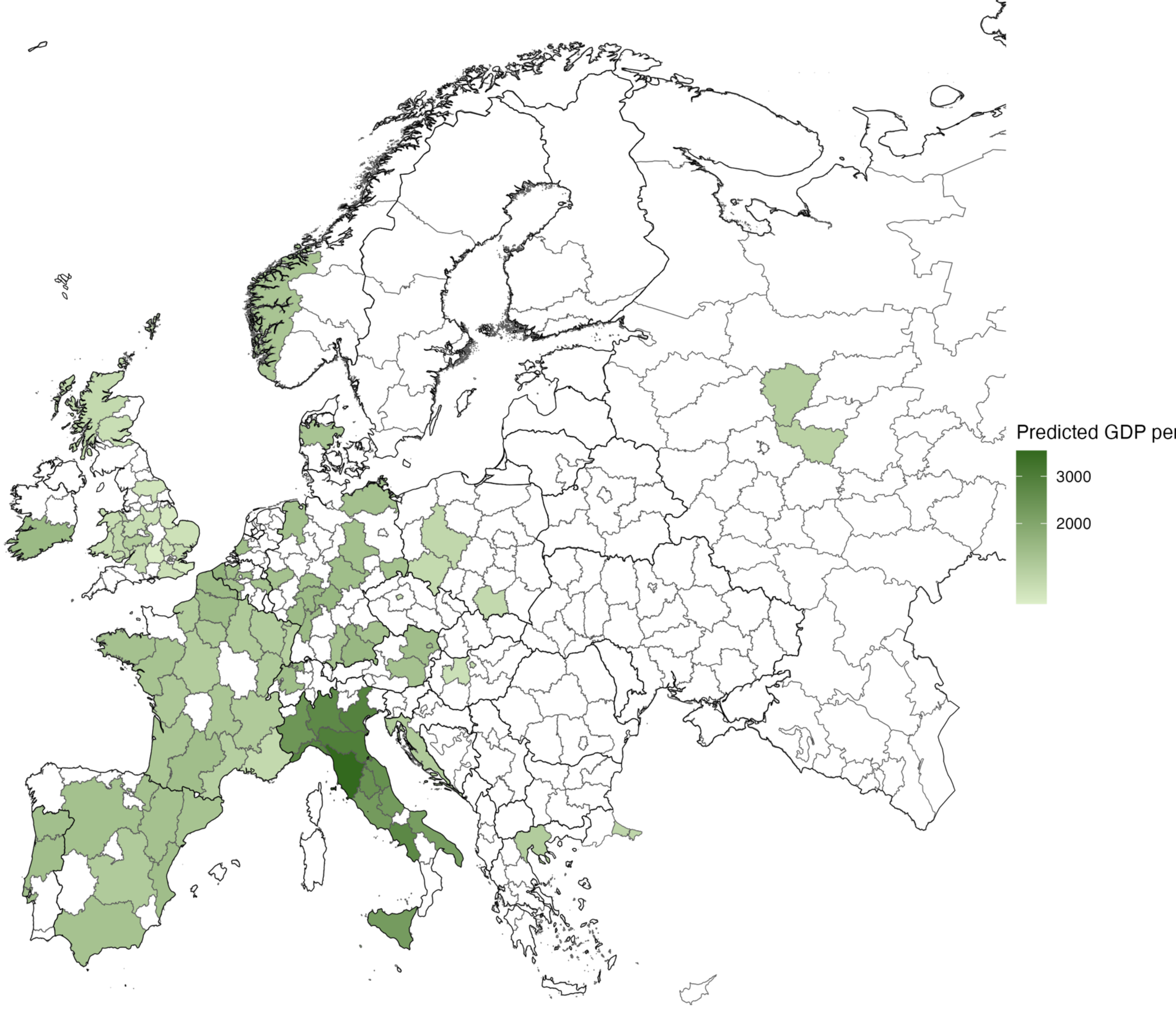
Koch, Stojkoski, Hidalgo PNAS (2024)

Validation - proxies of economic development

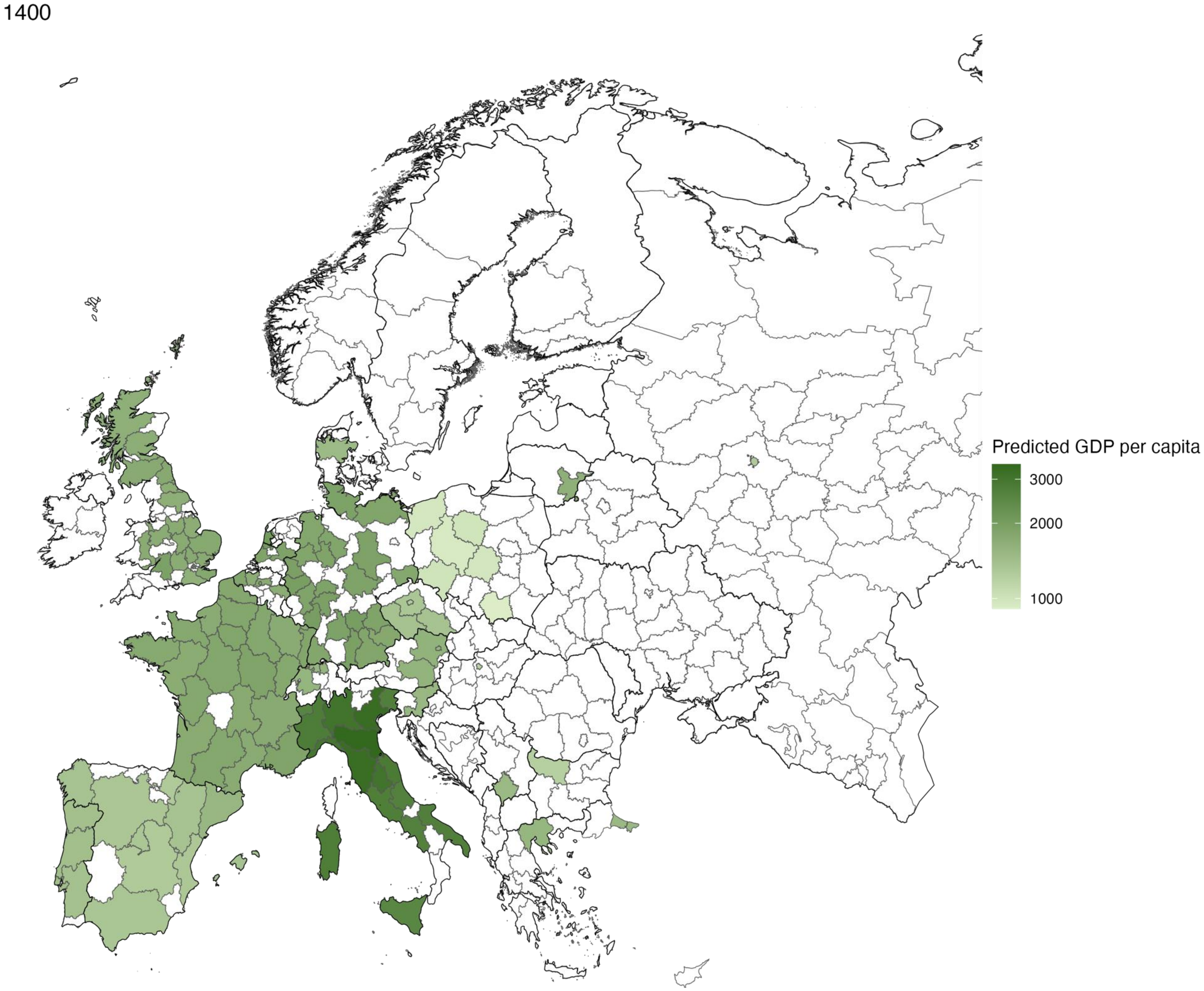


Koch, Stojkoski, Hidalgo PNAS (2024)

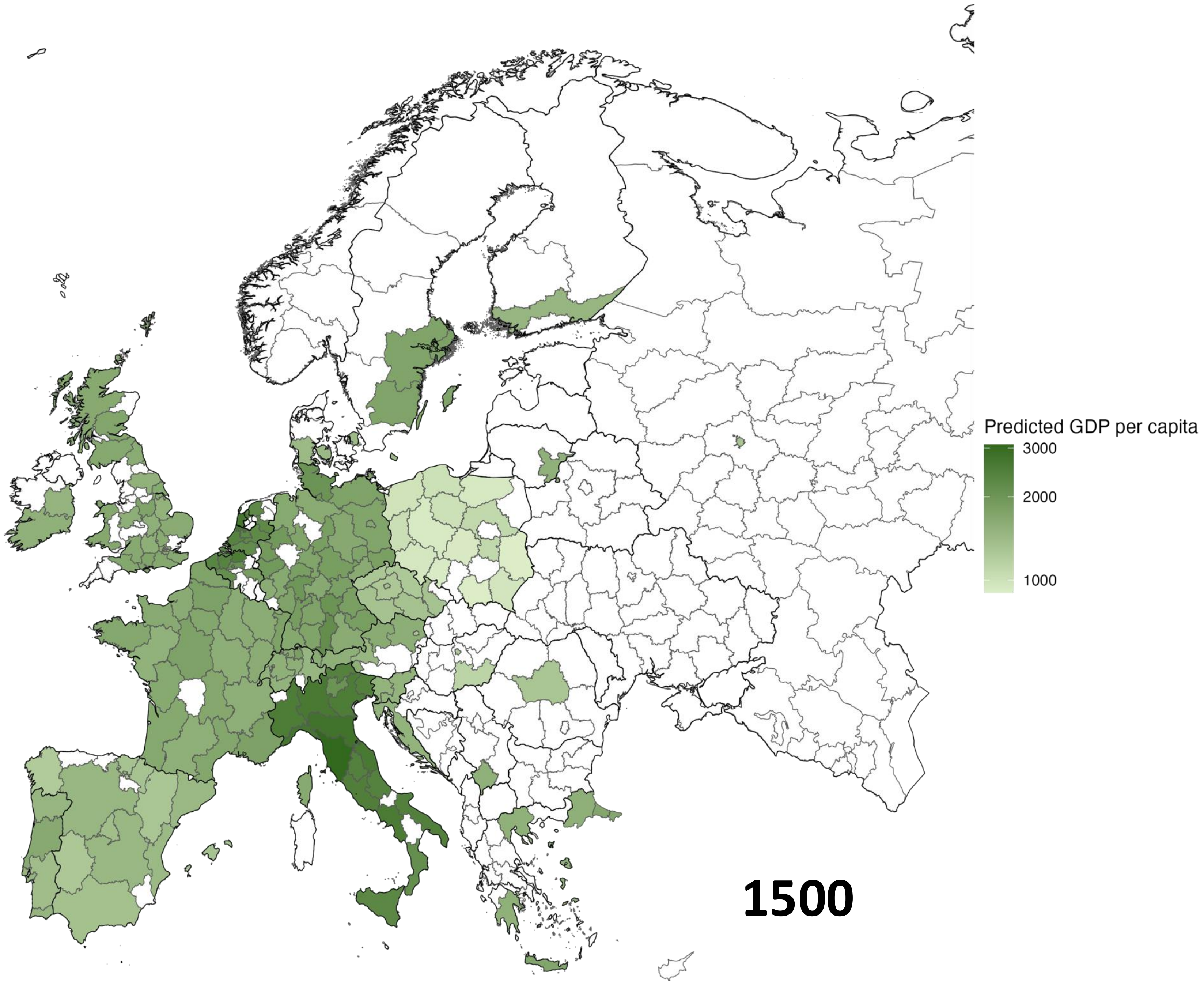
1300



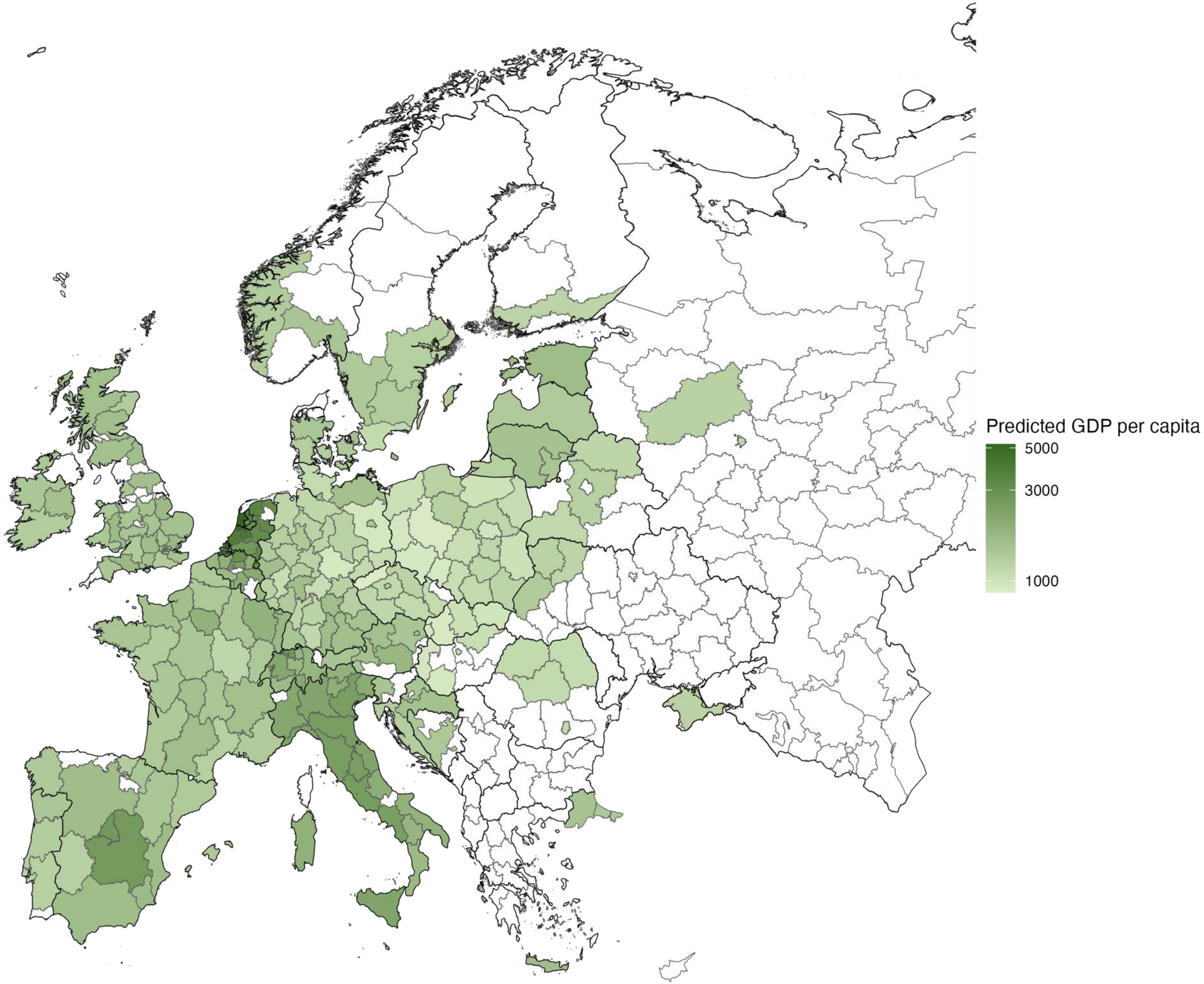
1400



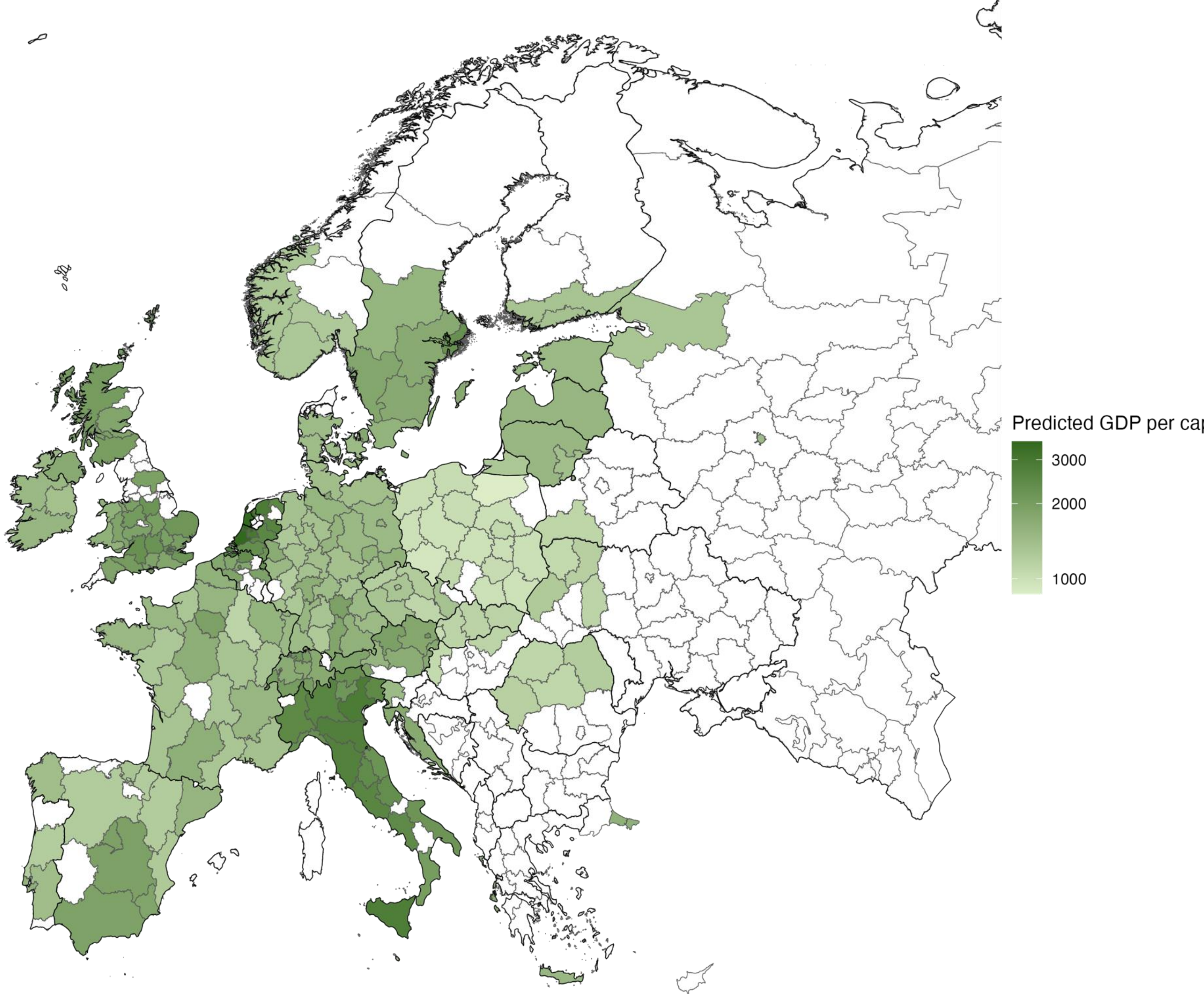
1500



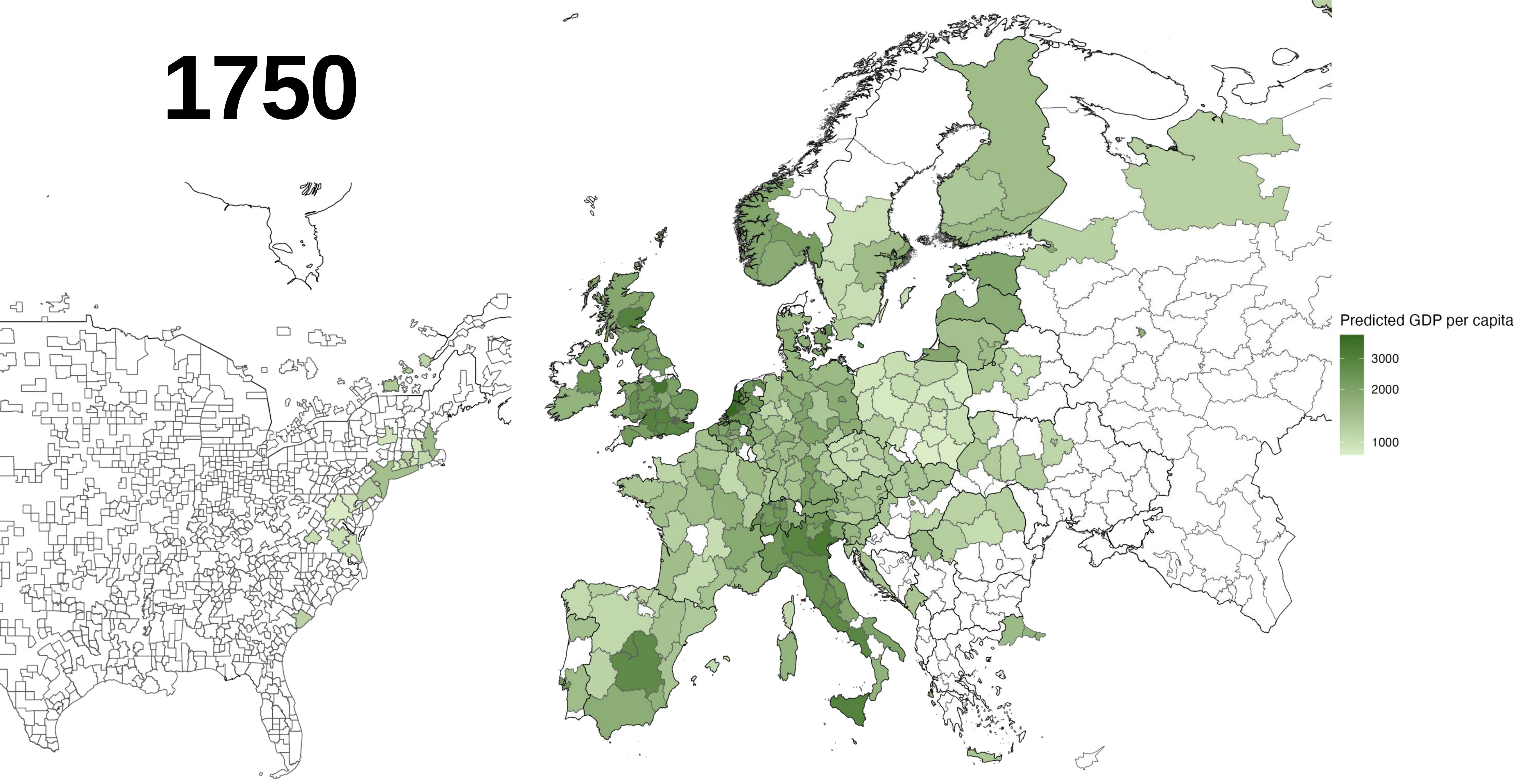
1600



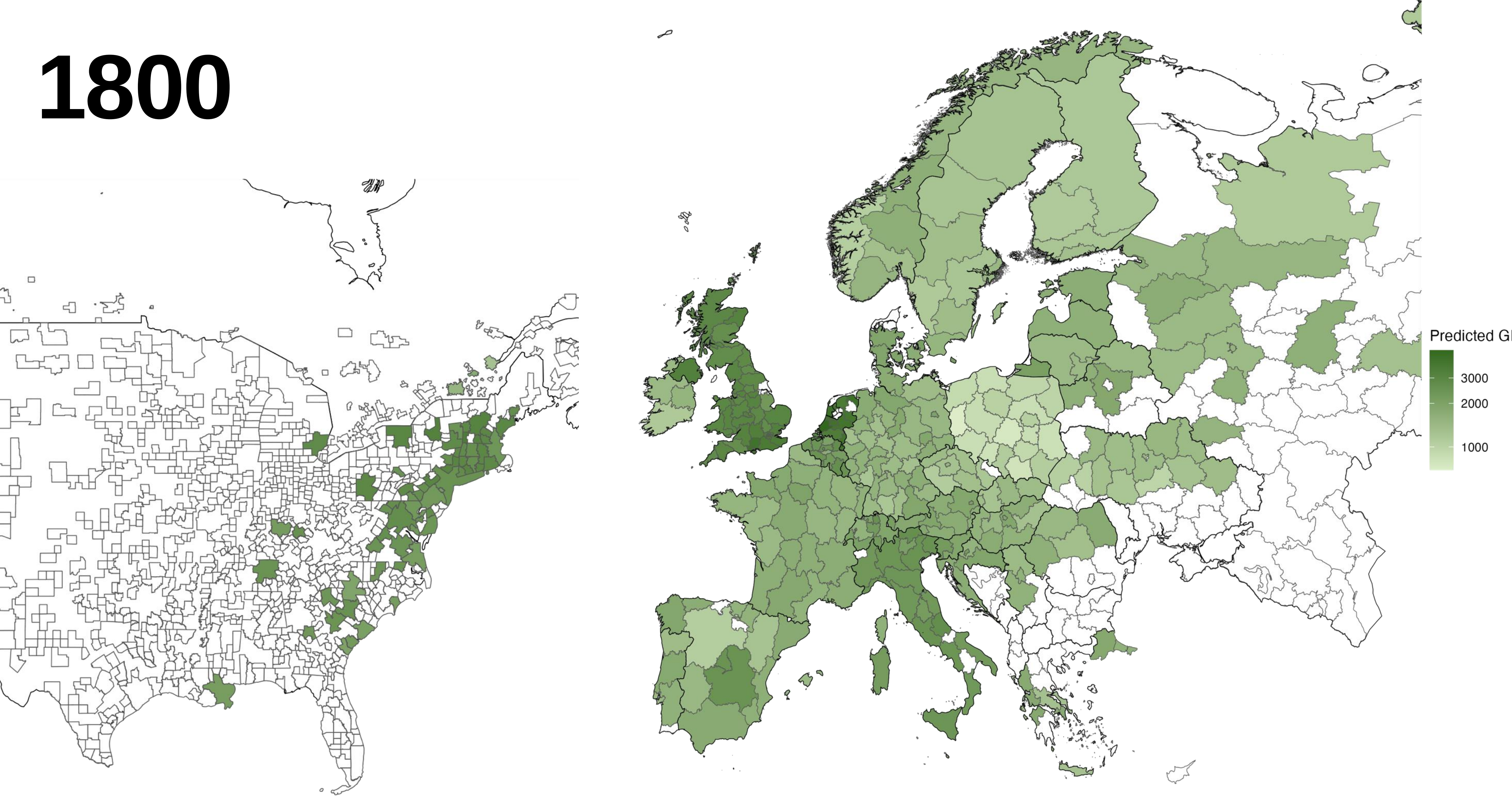
1700



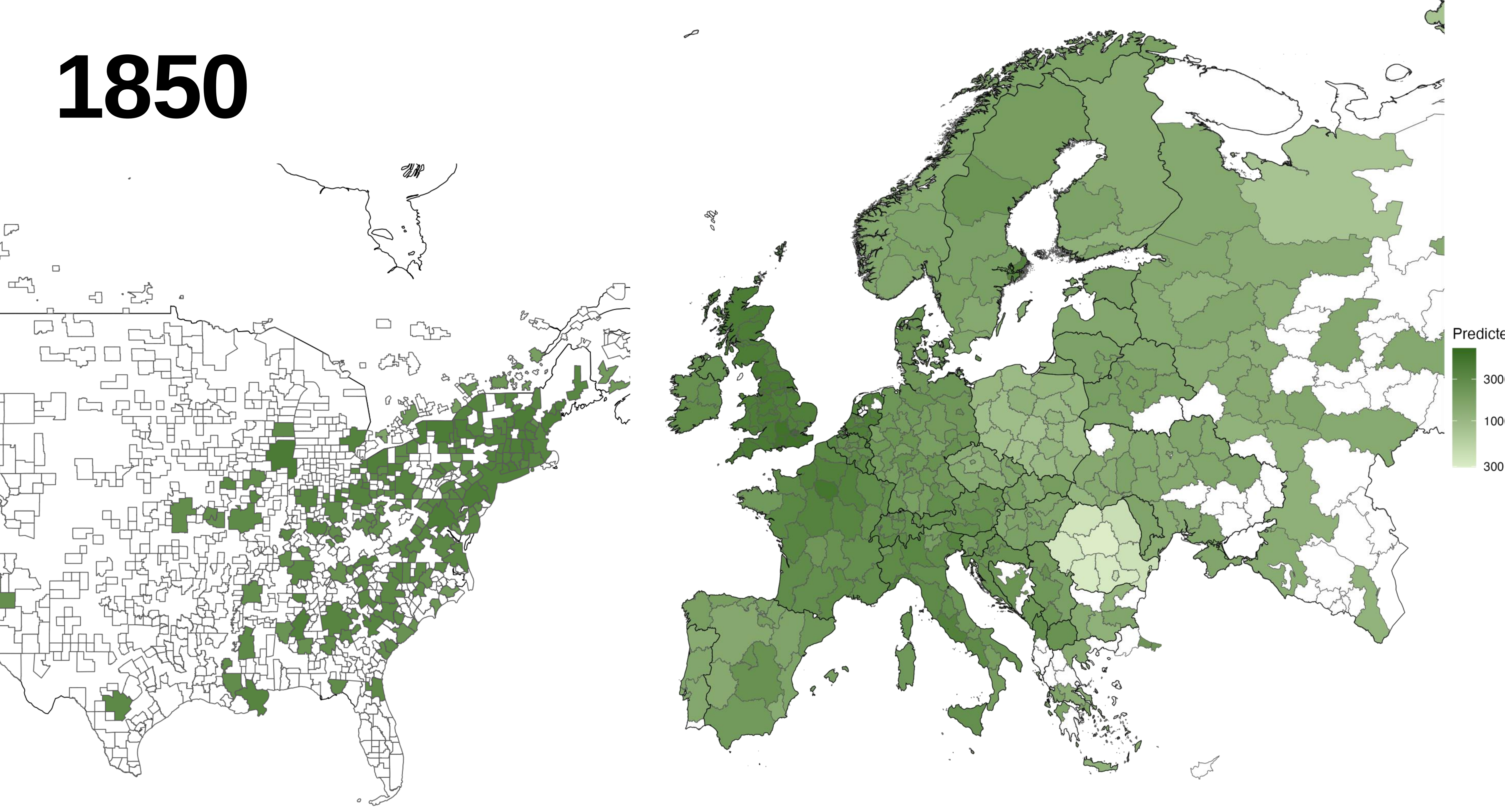
1750



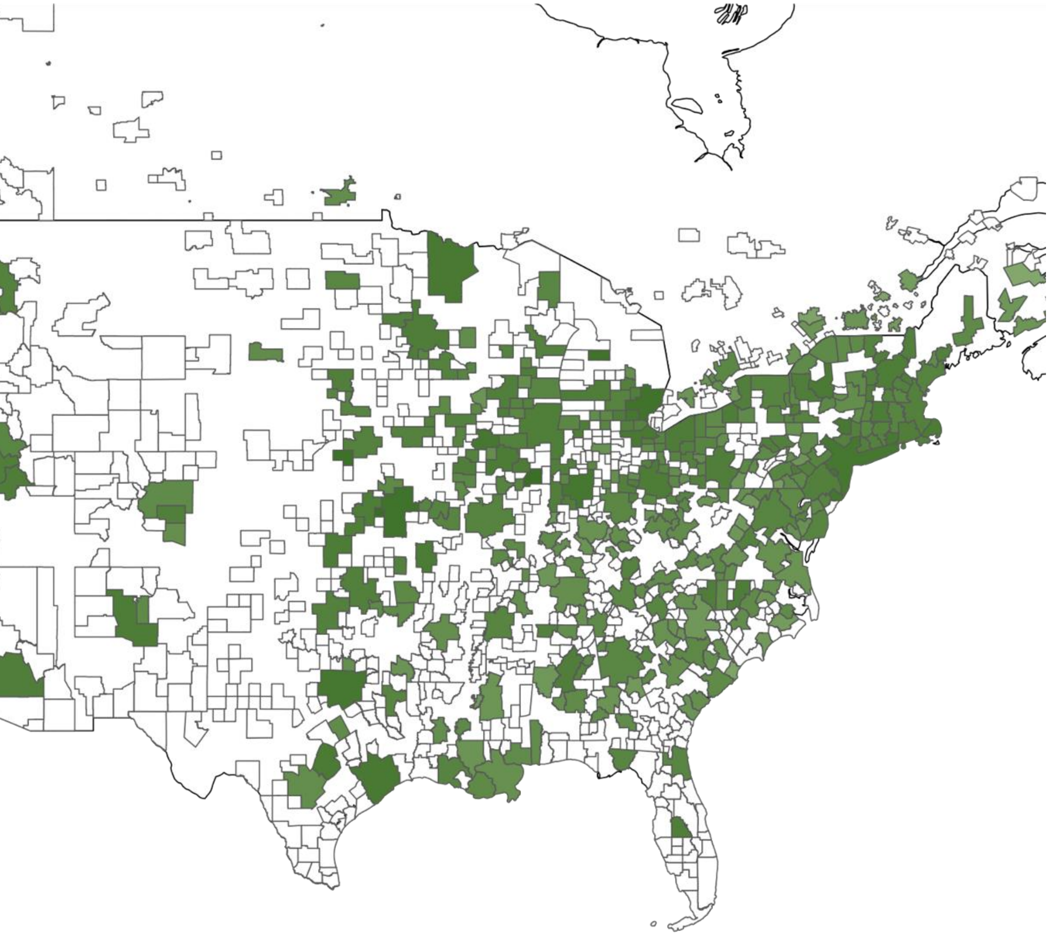
1800

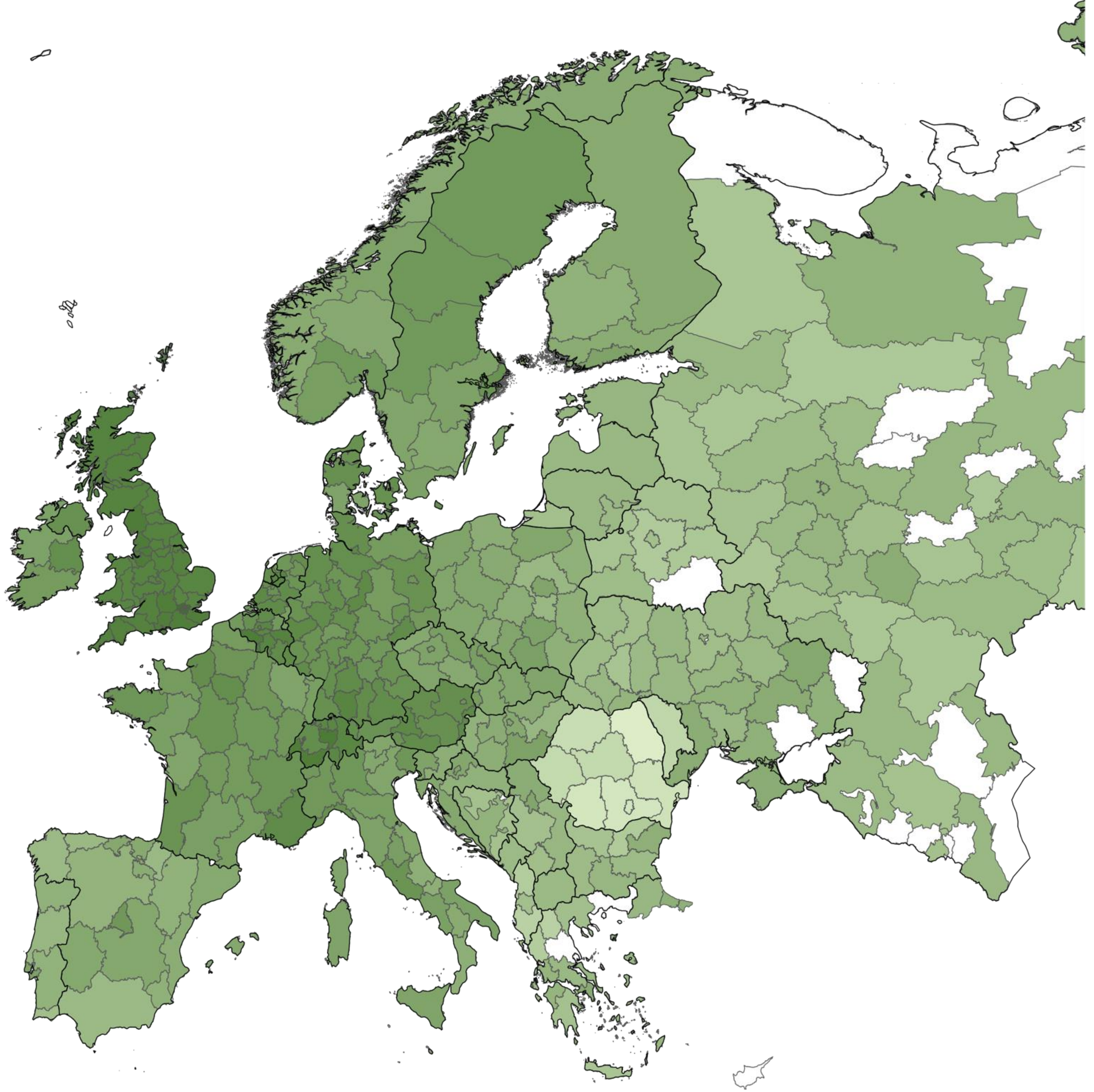


1850



1900





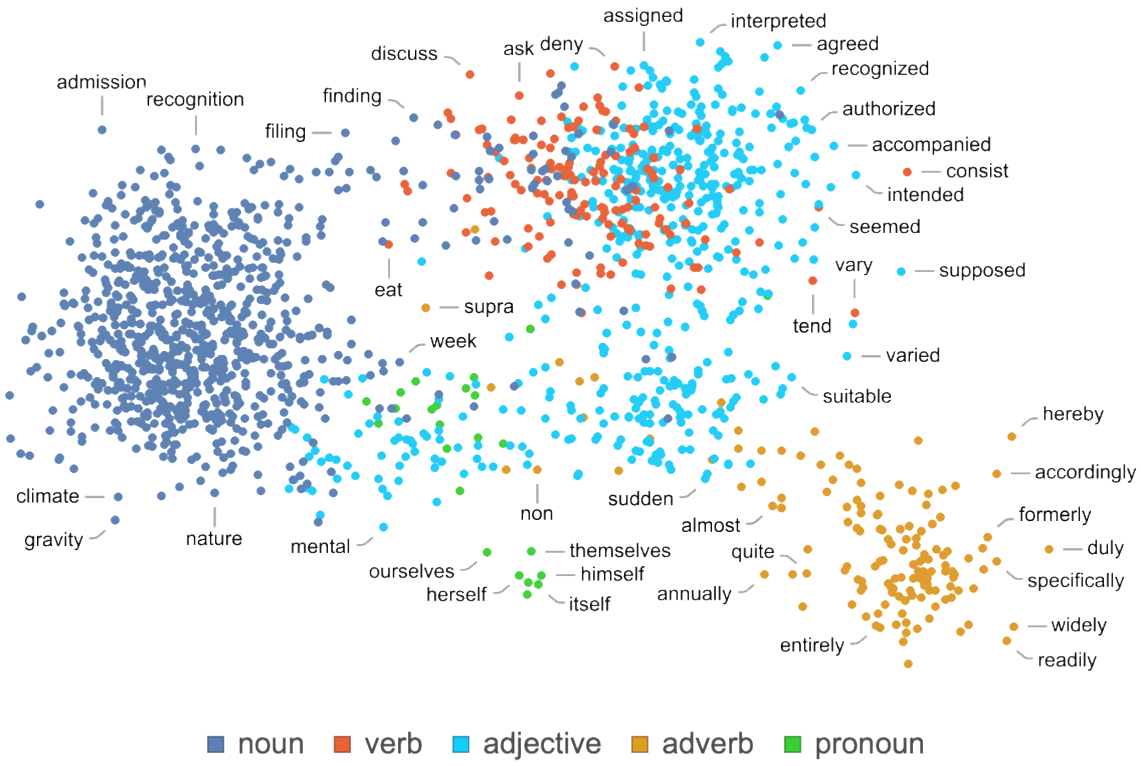
...to Conclude

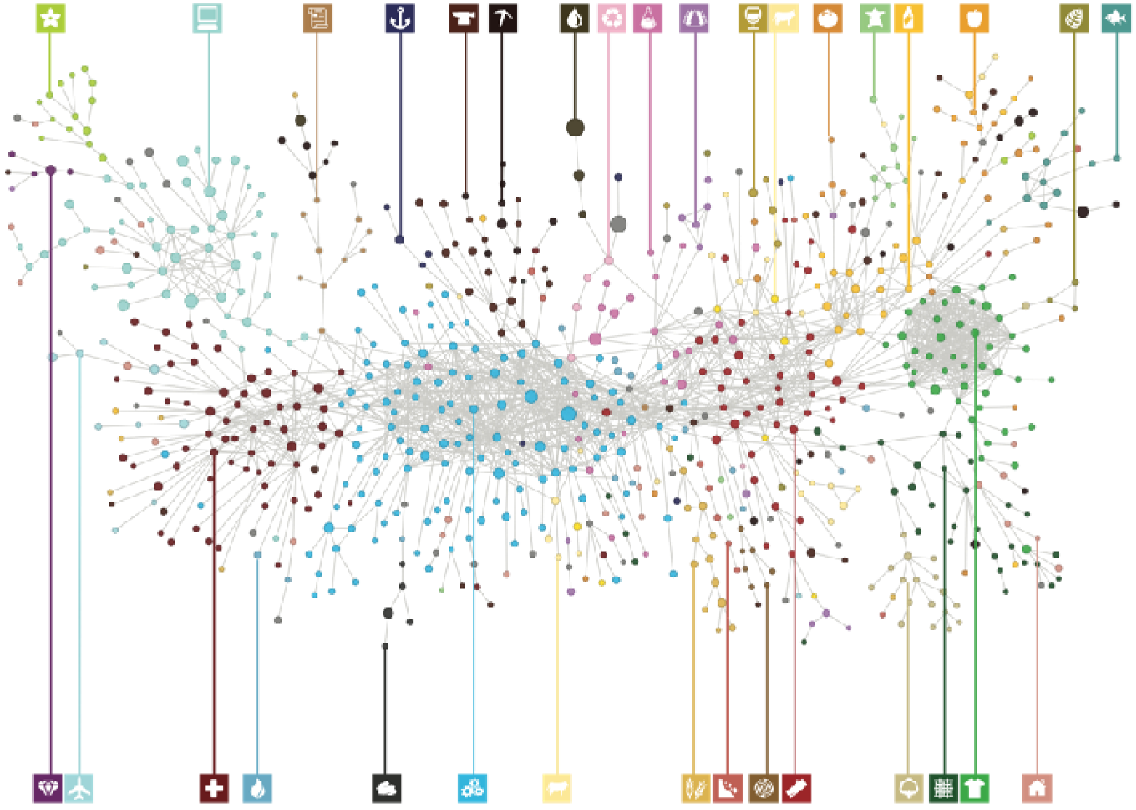
Center for Collective Learning

The world is complex

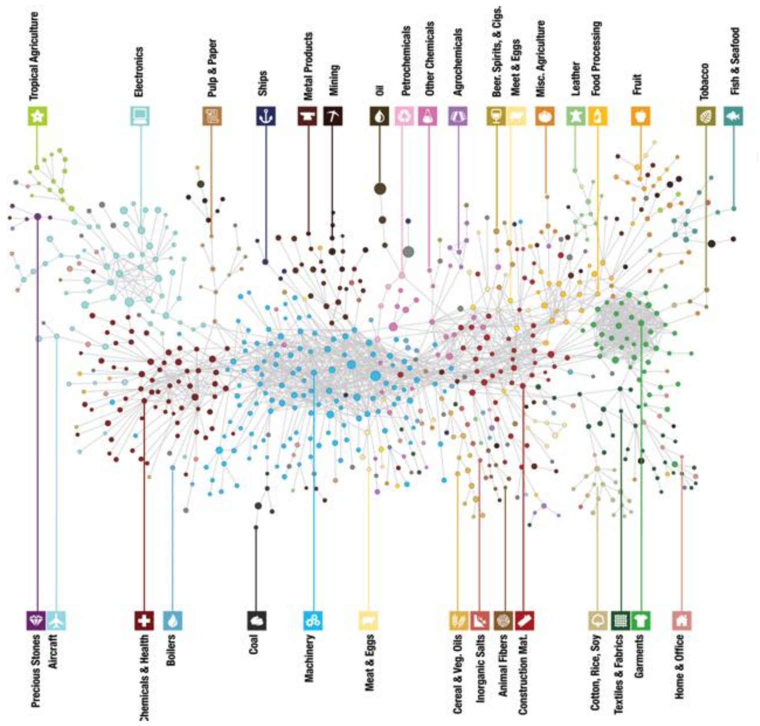
made of highly-specific
and non-fungible
knowledge

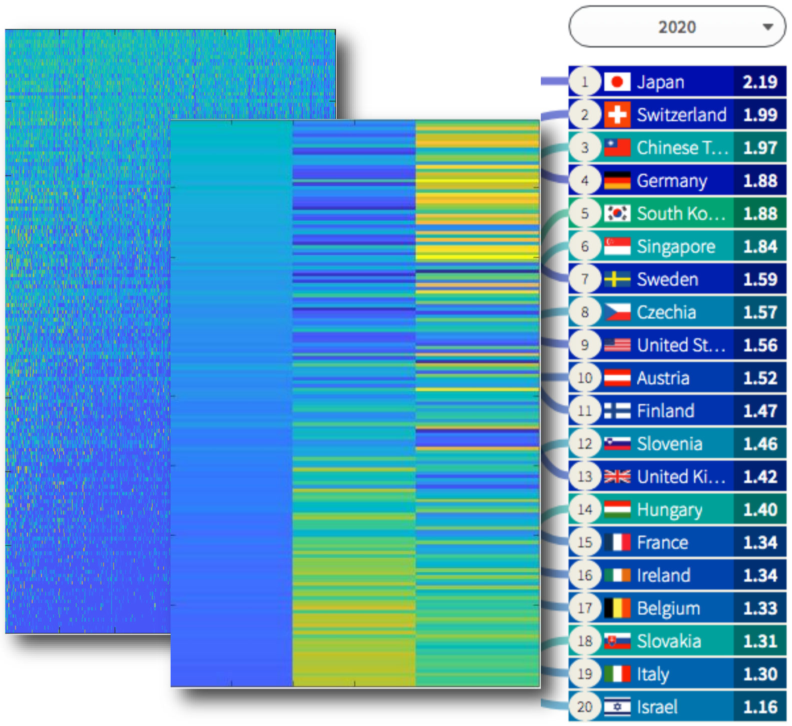

Knowledge is
non-fungible !


assigned
interpreted
agreed
recognized
authorized
accompanied
consist
intended
seemed
supposed
vary
tend
varied
suitable
discuss
ask
deny
finding
filing
admission
recognition
eat
supra
week
climate
gravity
nature
mental
non
sudden
almost
quite
annually
entirely
themselves
himself
itself
ourselves
herself
hereby
accordingly
formerly
duly
specifically
widely
readily
noun
verb
adjective
adverb
pronoun



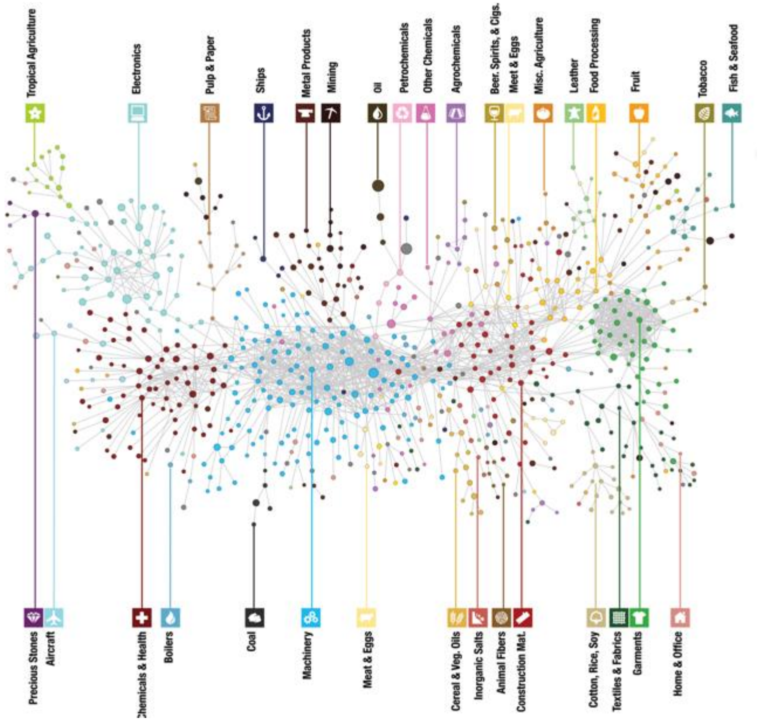
Economic complexity methods allow us to make granular representations of economies to understand where they stand and where they are going.

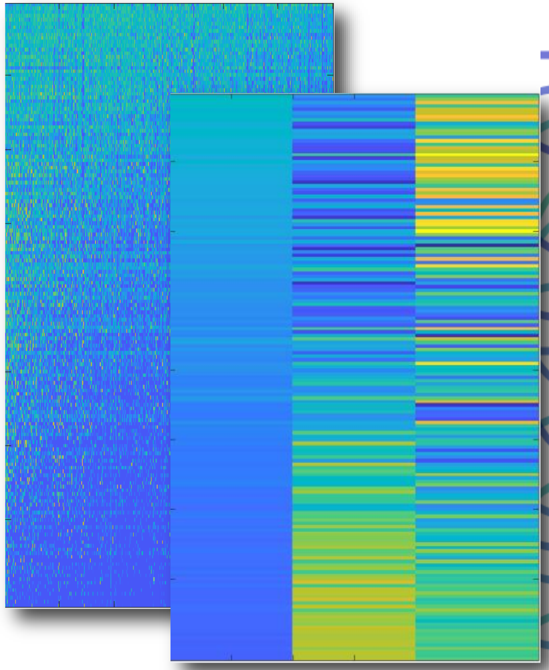


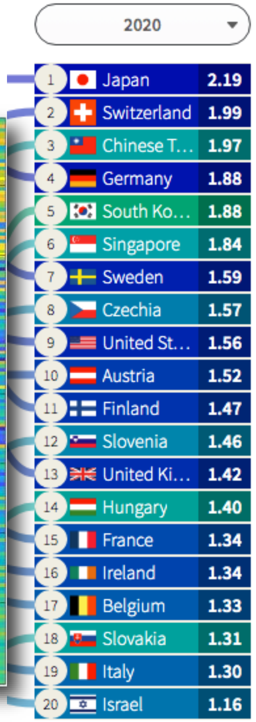


Research: centerforcollectivelearning.org
Platforms: datawheel.us oec.world

Thanks!


Tropical Agriculture
Electronics
Pulp & Paper
Ships
Metal Products
Mining
Oil
Petrochemicals
Other Chemicals
Agrochemicals
Beer, Spirits, & Cigs.
Meet & Eggs
Misc. Agriculture
Leather
Food Processing
Fruit
Tobacco
Fish & Seafood
Precious Stones
Aircraft
hemicals & Health
Boilers
Coal
Machinery
Meat & Eggs
Cereal & Veg. Oils
Inorganic Salts
Animal Fibers
Construction Mat.
Cotton, Rice, Soy
Textiles & Fabrics
Garments
Home & Office




2020
1 Japan 2.19
2 Switzerland 1.99
3 Chinese T... 1.97
4 Germany 1.88
5 South Ko... 1.88
6 Singapore 1.84
7 Sweden 1.59
8 Czechia 1.57
9 United St... 1.56
10 Austria 1.52
11 Finland 1.47
12 Slovenia 1.46
13 United Ki... 1.42
14 Hungary 1.40
15 France 1.34
16 Ireland 1.34
17 Belgium 1.33
18 Slovakia 1.31
19 Italy 1.30
20 Israel 1.16