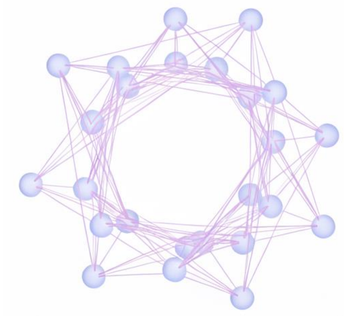
Industry Market Capitalization and Technology Complexity

Yuya Wang (yuyawang@brandeis.edu) https://yuya-wang.github.io/
PhD Candidate, School of Business and Economics, Brandeis University, Massachusetts, US

INTRODUCTION



Technological innovation accumulates over time by combining fundamental knowledge components.[1] As new features and processes build on earlier discoveries, technologies become increasingly complex.

This paper constructs a novel technological complexity score for all NAICS industries using U.S. patent data. By building annual CPC-based knowledge networks and characterizing their structural heterogeneity, the measure captures how diverse and interconnected a technology's underlying components are. The study then examines how complexity evolves over time and how it relates to product-market competition, industry profitability, and market capitalization.

METHODS

Construct technology networks[2]

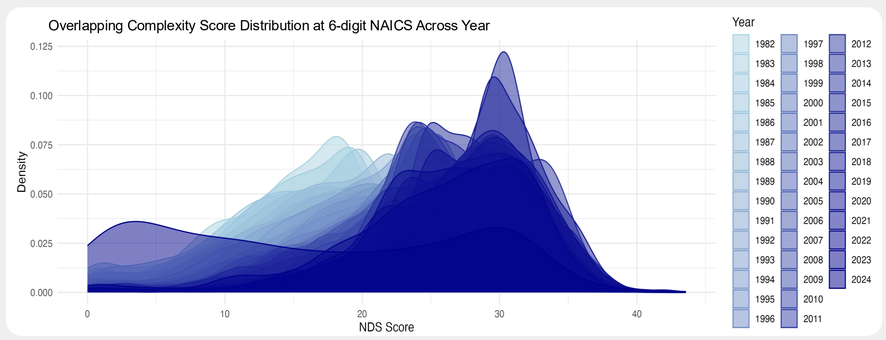
Node: 10-digit CPC subclasses; "knowledge components"
Edge: cooccurrence in one patent

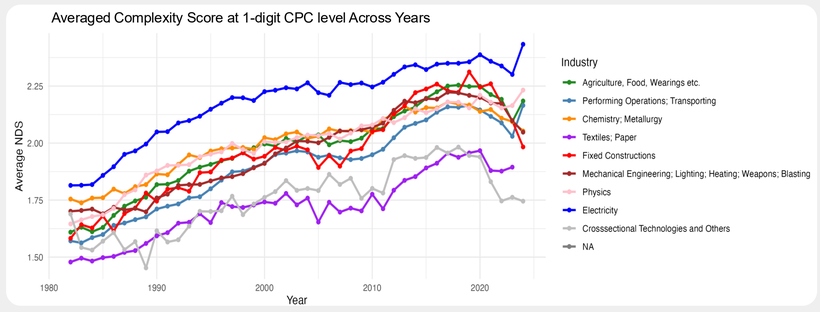
For each 4-digit CPC technology class, yearly networks are built and the proposed Network Diversity Score (NDS) summarizes structural heterogeneity using four network-structure indicators, outperforming single-metric approaches

Convert CPC network data to NAICS industries

CPC classes are mapped to NAICS using a two-step method:
1) ALP probabilistic links for initial NAICS vintages
2) Census concordance files to chain all seven NAICS versions

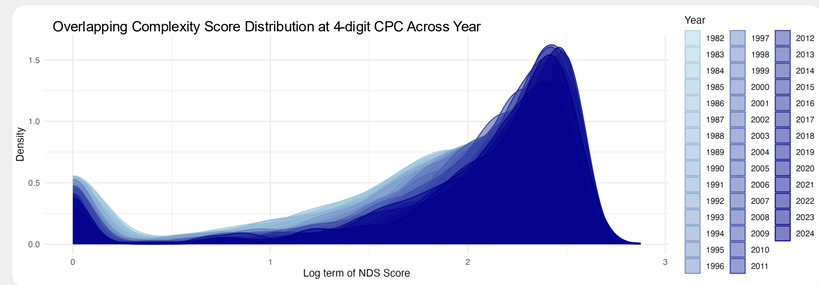
Patent counts, citations, and complexity scores are scaled and reweighted to remain comparable across conversions

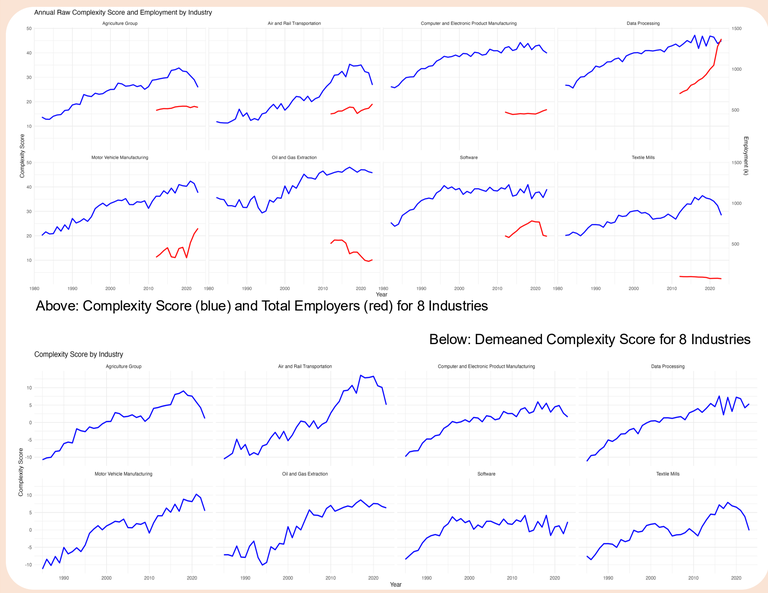




There shows a steady increase all areas until 2020, followed by a post-2020 pullback which is most pronounced in fixed construction.

Overall complexity rises, with more technologies scoring high and fewer scoring zero, and this fits the cumulative nature of innovation.





Above: Complexity Score (blue) and Total Employers (red) for 8 Industries

Below: Demeaned Complexity Score for 8 Industries

Complexity score varies across industries, making direct comparison challenging.
High-tech industries' technological complexity levels from forty years ago are estimated to be similar to those of labor-intensive/more competitive industries today.

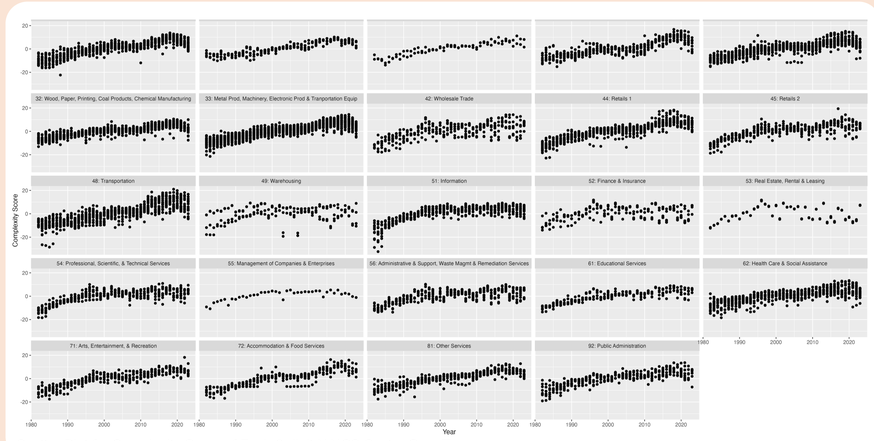
Industry-level benchmark: the averaged score across the time period for each industry
Excess score: initial Score – industry benchmark

All industries exhibit climbing score before 2020.
After the covid-19 pandemic struck, labor-intensive industries suffered reductions, while high-tech ones remained stable.

RESULTS

Comparisons of NDS with other network-based Metrics in Environment-related Areas (left) and AI-related Areas (right)

NDS has a sustained performance relative to the other methods



Scatter Plot for Complexity Score of Each 2-digit NAICS Sector Over Time

Sectors with fewer usable observations have limited innovation opportunities and thus generate less patents.
Two possible explanations:
1) technological improvement isn't central to their processes;
2) they've largely reached an innovation plateau.

Regression Model:

$$C_{i,t} = \beta_{i,t}^{0} + \beta_{i,t}^{A} \cdot Asset_{i,t-1} + \beta_{i,t}^{NI} \cdot NI_{i,t-1} + \beta_{i,t}^{L} \cdot Leverage_{i,t-1} + \beta_{i,t}^{g} \cdot g(patent)_{i,t-1}^{1} + \beta_{i,t}^{C} \cdot Citation_{i,t-1}^{1} + \gamma_i + \alpha_t + u_{i,t}$$

Total asset: size Indicator

Net income: revenue generated

Leverage ratio: debt reliance

Patent growth rate: innovation and technological advancement in the industry

Number of citations: influence of this area's on future ones

	Industrial Complexity Level					
	Full Sample		*HHI<p50*		*HHI≥p50*	
	(1)	(2)	(3)	(4)	(5)	(6)
Accounting Variables:						
Total Assets	-0.029**	-0.035**	-0.034	-0.059	-0.019	-0.014
	(0.013)	(0.016)	(0.032)	(0.040)	(0.016)	(0.019)
Net Income	0.006**	0.006*	-0.002	-0.001	0.020***	0.020***
	(0.003)	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)
Leverage Ratio	-0.263**	-0.219*	-0.033	0.071	-0.266**	-0.210
	(0.107)	(0.123)	(0.249)	(0.282)	(0.120)	(0.137)
Patent Activity Indicators:						
1-year Patent Growth Rate	0.242***		0.264***		0.127	
	(0.073)		(0.089)		(0.097)	
3-year Patent Growth Rate		0.612***		0.605***		0.475***
		(0.074)		(0.093)		(0.121)
1-year Citation Count	-0.114***		-0.199***		0.020	
	(0.023)		(0.031)		(0.032)	
3-year Citation Count		-0.239***		-0.394***		0.068
		(0.045)		(0.062)		(0.064)
Absorb Industry Effect	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs	6,840	5,858	3,476	2,971	3,364	2,887
Adj. R^2	48.90%	43.98%	49.14%	44.67%	55.44%	51.48%

Regression Results for Determinants of Industrial Complexity

Competitive industries: only patent activity indicators are significant
Complexity is driven both by innovation growth and the need to protect one's market share.

Concentrated industries: only net income remain significant
Complexity is purely driven by firm's profitability, increasing potential barriers to entry.

Selected Reference:
1 Arthur,W. Brian, "The Nature of Technology : What It Is and How It Evolves," Free Press, 2009.
2 Broekel, Tom, "Using structural diversity to measure the complexity of technologies," PLOS ONE, 05 2019, 14 (5), 1–23.