From Shipments to Supply Chains: Mining Input-Output Links from Firm-level Trade Flows

Karsten Mau[*] Antonio Vicencio[*] Mingzhi Xu[◇] Yawen Zheng[+]

[*]SBE, Maastricht University
[◇]INSE, Peking University
[+]SAFTI, Shenzhen University

5th Conference on "Trade, value chains and financial linkages in the global economy"
Rome, 29-30 October 2026

Introduction

Introduction

Dominant trade in intermediate inputs and Global Value Chains (GVCs)
Johnson (2014); Antràs & Chor (2021)

- Non-trivial economic interdependencies, source of concern (Kleinman et al., 2024)
- Shocks propagate along supply chains and across borders

Challenge: limited ability to trace trade or industrial policy shocks

- Big gap between granularity of shocks and supply-chain linkages in data
- Product-level linkages rarely observed

This paper: infer product-level cross-border supply-chain linkages; analyze two questions:

1/ Input-output linkages as a source of export-import co-movement

2/ Propagation of export shocks to third country input suppliers

Approach and main findings

ARM-based "Bidirectional Frequency Association" (BFA)

- Detect recurring patterns in firm-level customs data, co-occurrence of exported and imported goods *ij*
- China in 2004: large, emerging economy that (still) relied on imported inputs; informed sample selection to minimize false-positives/negatives

Recover very sparse bipartite network of (first-tier) input-output linkages

- Heterogeneity in product "complexity" and input "ubiquity"
- Aggregated relationships correlated with conventional IO tables

Main findings: China's WTO entry as early export-shock to GVCs

- Im-ex co-movement: 10% increase in exports $\rightarrow$ 3.4% more imported inputs
- Cross-border propagation: demand shock for input suppliers; esp. if RCA$>$1

Contributions to the Literature

Big-data and network inference in economics

- **Product-relatedness and the firms' product space**: Macedoni et al. (2024)
- **Firm (and establishment) networks**: Inoue & Todo (2025), Boehm et al. (2025)

  Flexible, easy-to-implement, and reproducible method; reveals unseen detail of firms' sourcing behavior and operational configurations that go beyond input-output linkages studied here

Empirical evaluations of shock propagation along the supply-chain

- **Supply-chain propagation of US-China trade war**: Flaaen & Pierce (2024) use US w $\approx$ 200 manufacturing industries $\rightarrow$ *better than WIOD/ICIO but US only*
  Handley et al. (2020) assume that inputs and outputs belong to same (higher-tier) product category $\rightarrow$ *very strong assumption; published version infers linkages differently, partly similar to us* (Handley et al., 2025)
- **Shock propagation via firm-to-firm linkages** (e.g. Barrot & Sauvagnat, 2016; Boehm et al., 2019; Carvalho et al., 2020; Borin et al., 2025) $\rightarrow$ *bypasses matching of inputs to outputs, but limited scope for application in other settings*

  Empirically quantify the relationship between exports and imports of a GVC-integrated economy, incl. how export expansions generates input demand in third countries

Measurement and inference of granular I-O linkages

- **Inputs purchased by single-product firms**: Vandenbussche & Viegelahn (2018) [raw-material inputs, India] Ma et al. (2025) [imported inputs for processing trade, China]
- **Case-studies**: Cajal-Grossi et al. (2023) [garment producers, Bangladesh], Cox (2025) [tariff-exposed steel inputs, US]
- **Expert interviews**: Pichler et al. (2020) [focus on essential inputs]
- **AI/LLM-based**: Fetzer et al. (2024) [*AIPNET*, ask ChatGTP based on harmonized product-code descriptions]
- **Data-driven approaches**: Karbevska & Hidalgo (2025) [sub-national trade statistics]; O'Clery et al. (2025) [firm-level transactions]

  Extend on data-driven approaches, taking into account trade theory and firm behavior to minimize false-positives/negatives

Background and method

The problem

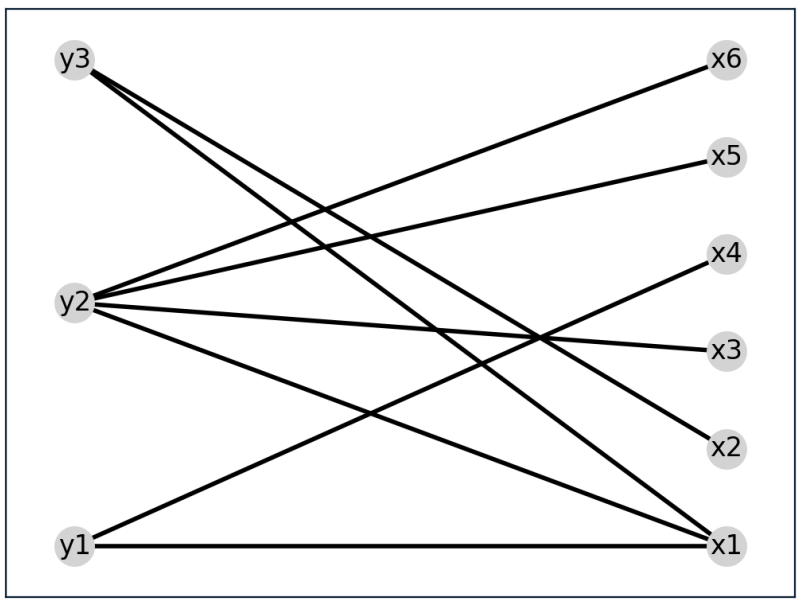
A very general description of a production function

$$y_i = f(X_i, \cdot) \tag{1}$$

Q: What are these inputs in $X_i = (x_{i1}, ..., x_{ij}, ..., x_{iJ})$ that are used to produce y_i?

A: Difficult to say; production of single goods typically not directly observed

Firm as the subset of a bipartite network



Note: Output y_i (left) requires inputs x_j (right)

Challenges

- Edges cannot be observed
- Prevalence of multi-product firms
- Vertical integration, CAT, economies of scope, etc.
- ▸ More

Our approach

- Look at many "similar" firms
- Infer edges based on systematic co-occurrences

Association Rule Mining (ARM)

Extracts associations among items in transaction databases
→ Agrawal et al. (1993); Zhao & Bhowmick (2003); Solanki & Patel (2015); Telikani et al. (2020); Kaushik et al. (2021)

- Applied in many settings, e.g. marketing research, computer networks, recommendation systems, healthcare

Key statistics

(i) "Support" $(\mathcal{S}_k) \rightarrow$ measures prevalence of attribute k in the data

$$\mathcal{S}_k \equiv \frac{N_k}{N}; \qquad \text{fraction of firms } N \text{ where } k \text{ observed is observed}$$

(ii) "Lift" $(\mathcal{L}_{ij}) \rightarrow$ expresses the relative co-occurrence of two items ij

$$\mathcal{L}_{ij} \equiv \frac{\mathcal{S}_{ij}}{\mathcal{S}_i \times \mathcal{S}_j} = \frac{N_{ij}}{N_i \times N_j} N; \qquad \mathcal{L}_{ij} > 1 \rightarrow \text{ more than random co-occurrence}$$

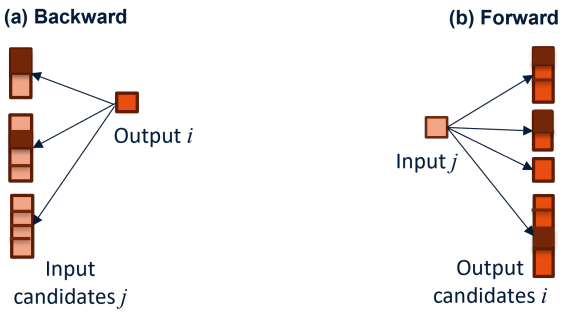
Directional linkages: $\mathcal{L}_{j \rightarrow i} \rightarrow$ non-random co-occurrence of input j and output i

- Caveat: ARM less effective in very sparse networks (large N and small N_k)

ARM extension

"Bidirectional Frequency Association" (BFA)

- Backward: input-candidates for each produced good $\rightarrow \mathcal{S}^{\text{back}}_{i \leftarrow j} = N_i(j)/N_i$
- Forward: output-candidates for each purchased input $\rightarrow \mathcal{S}^{\text{fwd}}_{j \rightarrow i} = N_j(i)/N_j$



Combine to obtain: $Score^{BFA}_{j \rightarrow i} = (\mathcal{S}^{\text{back}}_{i \leftarrow j} \times \mathcal{S}^{\text{fwd}}_{j \rightarrow i}) \in [0, 1]$

- BFA infers prevalence of ij links in firms; penalizes scores of multi-product firms and "ubiquitous" inputs
- Formally: $Score^{BFA}_{j \rightarrow i} = \frac{N^2_{ij}}{N_i \times N_j} = \cos^2 \theta_{ij} \rightarrow$ squared cosine-similarity;

  Also: $Score^{BFA}_{j \rightarrow i} = \frac{N^2_{ij}}{N_i \times N_j} = \mathcal{L}_{ij} \frac{N_{ij}}{N} \rightarrow$ bidirectional confidence that j is used to produce i

  $\rightarrow$ No immediate structural interpretation (e.g. expenditure share)

Data and implementation

Firm-level trade data

Advantages

- Detailed and harmonized products: $\approx$ 5,000 products; *Harmonized System* WCO admin
- Relevant: HS codes used for trade policy
- Reproducible: Standardized data format firm-product-partner (plus aux info)

Caveats

- No domestic transactions
- Comparative advantage and vertical integration
- CAT and trade intermediation

Some challenges can be addressed directly or indirectly

- Firms specialize in core competences, likely captured by their exports
  Bernard et al. (2011); Mayer et al. (2014); Eckel et al. (2016)
- Large and successful firms source globally
  Kugler & Verhoogen (2011); Topalova & Khandelwal (2011); Manova & Yu (2017)
- Trade-regime and firm-type can indicate CAT or intermediation
  Bernard et al. (2019); Eckel & Riezman (2020); Erbahar & Rebeyrol (2023)

▸ Back

Data and sample selection

China Customs Database, single cross-section for 2004

Manova & Zhang (2012); Manova & Yu (2016); Chor et al. (2021)

- Large economy, many firms in all sectors, lots of imported inputs (Amiti & Freund, 2010)
- Raw data: 3.3 million obs.; firm-product-direction-partner-"regime"
- "regime" = customs regime → some suspicious of CAT/intermediation

Sample selection

- Select relevant custom regimes (excl. warehouses, transit, re-exports); aggregate to firm-direction-product level at 4 digits (HS4)
- Select firms that export and import at least 1 good; remove single product firms that import/export same good
- Final sample: 66,688 firms trading 1,361 HS4 products

▸ Stats 1 ▸ Stats 2

Implementation

Implement BFA in three versions

Baseline: compute $Score_{j \to i}^{BFA} = (\mathcal{S}_{i \leftarrow j}^{\text{back}} \times \mathcal{S}_{j \to i}^{\text{fwd}})$, then determine "strong link" if

- j's input candidate has $Score_{j \to i}^{BFA}$ more than 2-times avg. non-zero input candidate; and
- i's output candidate has $Score_{j \leftarrow i}^{BFA}$ more than 2-times avg. non-zero output candidate

Post-processing: follow baseline, but use BEC end-use categories to

- identify and remove final consumption goods as inputs; and
- identify strong linkages separately for capital goods (which are traded less often)

Pre-/post-processing: before applying BFA

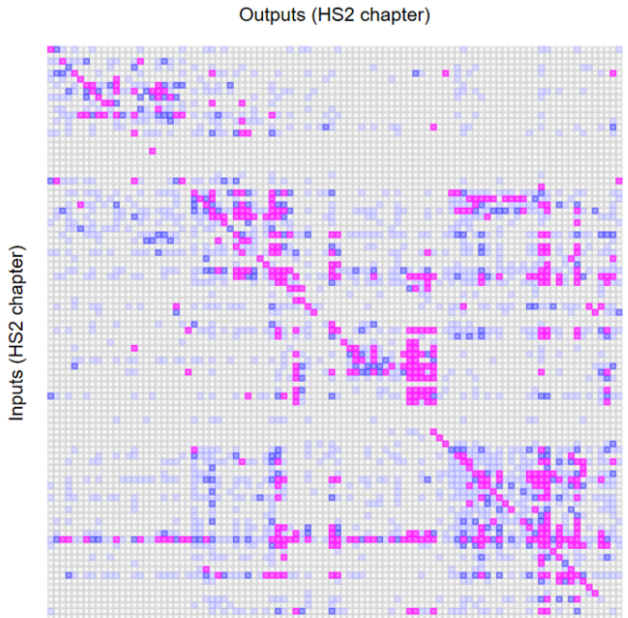
- remove most diversified exporters (top 10%) $\rightarrow >$ 19 goods
- then apply post-processing

Results

Main results for BFA-based link inference

BFA reveals sparse bipartite networks

HS2-level imports-exports matrix



Note: Average HS4-level BFA scores within 2-digit HS chapters. Color intensity: white = zero, light blue = non-zero, dark blue = sample average, pink = above 95th percentile.

Distribution of linkages

		Inputs per output (HS4 level)						
Method	count	avg.	sd	p5	p25	p50	p90	max
BFA baseline	1,218	47.4	50.8	5	13	25	144	197
BFA post	1,210	35.5	38.9	4	9	18	110	153
BFA pre- & post	1,168	**17.7**	20.8	2	5	**10**	68	121
		Outputs per input (HS4 level)						
Method	count	avg.	sd	p5	p25	p50	p90	max
BFA baseline	1,215	47.5	50.7	4	14	26	142	215
BFA post	935	46.0	49.3	4	13	25	135	213
BFA pre- & post	924	**22.4**	22.4	2	8	**14**	53	118

Total number of linkages: 50,639 (baseline) vs. 19,458 (pre-& post)
$\rightarrow$ latter is $< 1.5\%$ of potential linkages

Validation

✓ Positive correlation with aggregated IO tables ▸ Scatter plots ▸ Regressions
→ Typically weaker correlation if inputs related to RCA industries ▸ show
→ Proportionality with expenditure shares likely a strong assumption

✓ Not the same as AIPNET (Fetzer et al., 2024)
→ Comparable link frequencies and distribution: ▸ Graph
→ Actual overlap quite low ▸ Table
→ Different concepts of what is an "input" ▸ Examples

✓ External validation delivers quite strong results ▸ Estimates
→ Import growth predicted by BFA-weighted export growth (very robust)
→ Show that correlation not random (placebo estimates w randomized mapping)
→ Import-export linkages also work for other Southeast Asian economies

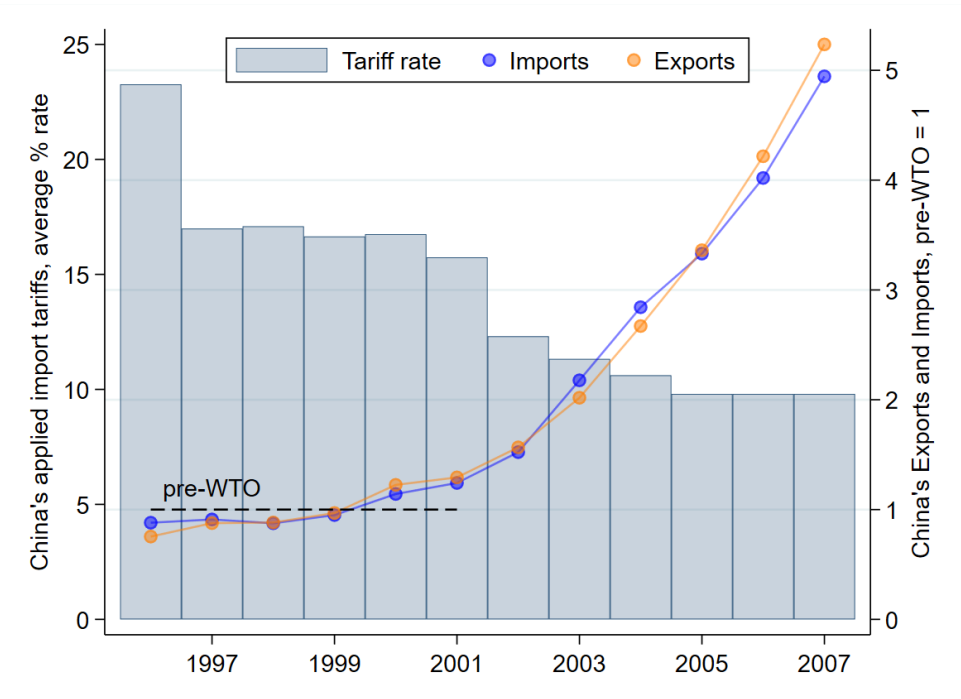
✓ Extensions to other countries and more granular HS-levels
→ Similar mappings for manufacturing countries ▸ Panjiva
→ HS6-level mappings not always needed, some with small N ▸ show

Application

"*China Shock* 1.0" - Background

China's WTO entry in Dec. 2001 widely studied and quite well understood

- Focus primarily on Chinese competition (supply shock); also evidence on demand effect
- Empirical identification based on shift-share → agnostic about mechanisms

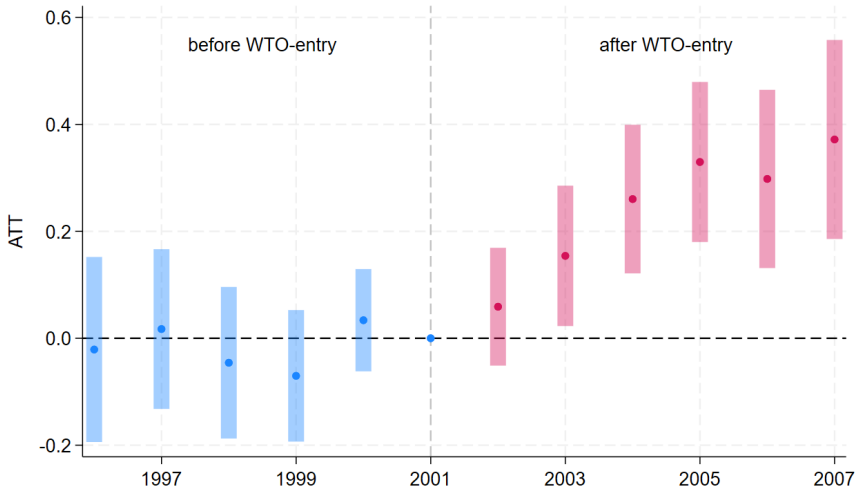


- Post-WTO entry: upward-trend in both exports and imports → role of supply-chain linkages?

"*China Shock* 1.0" - Demand for imported inputs

Regress imports of j on indicator of i-exports in DiD setting

- Compare js linked to fast-vs-slow growing exports i (top-vs-bottom quartile)
- Compare $t < 2001$ (before) and $t > 2001$ (after WTO entry)



	(1)	(2)	(3)
Dependent variable:	China's imports (upstream)		
Measured as	Δ% growth rate	Levels (log-lin)	Levels (PPML)
Downstream export growth ×WTO	0.366** (0.080)	0.237** (0.062)	0.451** (0.120)
HS4 fixed effects		✓	✓
Year fixed effects		✓	✓
Control variables (tariffs)		✓	✓
Cluster (HS4 products)		689	689
Observations	1,084	8,264	8,268

Note: Fixed effects and control variables are included as indicated at the bottom of the table. Standard errors in parentheses. Statistical significance: ^{a} $p < 0.1$, * $p < 0.05$, ** $p < 0.01$

→ Export growth entailed growing imports for related inputs; elasticity $\in [0.237, 0.451]$

"*China Shock* 1.0" - Supply of inputs to China

Chinese demand for inputs and supplier RCA

	(1)	(2)
Dependent variable:	(log) Exports to China	
Downstream growth (China) ×WTO	0.218**	0.140*
	(0.060)	(0.062)
× $RCA_n(j) > 1$		0.234**
		(0.041)
(log) Chinese $tariff_{jt}$	-0.175*	-0.175*
	(0.083)	(0.083)
Control variables	✓	✓
Cluster (HS4 products)	687	687
Observations	215,547	215,547

Note: Table reports results of HS4-country pair exports to China in response to Chinese downstream export growth. SE in parentheses adjusted for clustering at the HS4 level. Statistical significance: ^{a} $p < 0.1$, * $p < 0.05$, ** $p < 0.01$

→ China shock 1.0 entailed export opportunities for countries with RCA in upstream inputs

▸ Map

Next step: investigate China shock 2.0 → lower elasticity due to more domestic production?

Conclusion

Summary and conclusion

- Association-rule mining algorithm to recover a granular, product-level input-output structure from firm-level customs transactions;
- Discuss the strengths and weaknesses of such an approach; calibrate it to underlying data and well-known facts about firms in international trade;
- Inferred mappings resemble conventional input-output relationships and perform well in external validation;
- Illustrate value in a concrete application China's WTO entry 2001 entailed significant import demand for upstream inputs;
- Simple pattern recognition tools and micro-level trade data can improve exposure measures and empirical assessments of supply-chain propagation

Thank you!

Comments welcome

k.mau@maastrichtuniversity.nl

Agrawal, R., Imieliński, T., & Swami, A. (1993). Mining association rules between sets of items in large databases. In *Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data*, SIGMOD '93 (pp. 207–216). New York, NY, USA: Association for Computing Machinery.

Amiti, M. & Freund, C. (2010). The Anatomy of China's Export Growth. In *China's Growing Role in World Trade*, NBER Chapters (pp. 35–56). National Bureau of Economic Research, Inc.

Antràs, P. & Chor, D. (2021). *Global Value Chains*. NBER Working Papers 28549, National Bureau of Economic Research, Inc.

Barrot, J.-N. & Sauvagnat, J. (2016). Input Specificity and the Propagation of Idiosyncratic Shocks in Production Networks. *The Quarterly Journal of Economics*, 131(3), 1543–1592.

Berlingieri, G., Pisch, F., & Steinwender, C. (2021). Organizing global supply chains: Input-output linkages and vertical integration. *Journal of the European Economic Association*, 19(3), 1816–1852.

Bernard, A. B., Blanchard, E. J., Van Beveren, I., & Vandenbussche, H. (2019). Carry-Along Trade. *The Review of Economic Studies*, 86(2), 526–563.

Bernard, A. B., Redding, S. J., & Schott, P. K. (2011). Multiproduct Firms and Trade Liberalization. *The Quarterly Journal of Economics*, 126(3), 1271–1318.

Boehm, C., Flaaen, A. B., Pandalai-Nayar, N., & Schlupp, J. (2025). *The Local-Area Incidence of Exporting*. Working Paper 34508, National Bureau of Economic Research.

Boehm, C. E., Flaaen, A., & Pandalai-Nayar, N. (2019). Input Linkages and the Transmission of Shocks: Firm-Level Evidence from the 2011 Tōhoku Earthquake. *The Review of Economics and Statistics*, 101(1), 60–75.

Borin, A., Conteduca, F., Leone, F., Mancini, M., & Zoi, P. (2025). *How Global Are Local Value Chains?* CESifo Working Paper Series 12271, CESifo.

Cajal-Grossi, J., Macchiavello, R., & Noguera, G. (2023). Buyers' Sourcing Strategies and Suppliers' Markups in Bangladeshi Garments. *The Quarterly Journal of Economics*, 138(4), 2391–2450.

Carvalho, V. M., Nirei, M., Saito, Y. U., & Tahbaz-Salehi, A. (2020). Supply Chain Disruptions: Evidence from the Great East Japan Earthquake*. *The Quarterly Journal of Economics*, 136(2), 1255–1321.

Chor, D., Manova, K., & Yu, Z. (2021). Growing like China: Firm performance and global production line position. *Journal of International Economics*, 130(C).

Cox, L. (2025). The Long-Term Impact of Steel Tariffs on U.S. Manufacturing. *mimeo*.

Eckel, C., Iacovone, L., Javorcik, B., & Neary, J. P. (2016). Testing the core-competency model of multi-product exporters. *Review of International Economics*, 24(4), 699–716.

Eckel, C. & Riezman, R. (2020). CATs and DOGs. *Journal of International Economics*, 126(C).

Erbahar, A. & Rebeyrol, V. (2023). Trade intermediation by producers. *Journal of International Economics*, 140(C).

Fetzer, T., Garg, P., Lambert, P., & Feld, B. (2024). *AI-Generated Production Networks: Measurement and Applications to Global Trade*. Technical report, mimeo.

Flaaen, A. & Pierce, J. (2024). Disentangling the effects of the 2018-2019 tariffs on a globally connected u.s. manufacturing sector. *The Review of Economics and Statistics*, (pp. 1–45).

Handley, K., Kamal, F., & Monarch, R. (2020). *Rising Import Tariffs, Falling Export Growth: When Modern Supply Chains Meet Old-Style Protectionism*. NBER Working Papers 26611, National Bureau of Economic Research, Inc.

Handley, K., Kamal, F., & Monarch, R. (2025). Rising import tariffs, falling exports: When modern supply chains meet old-style protectionism. *American Economic Journal: Applied Economics*, 17(1), 208–38.

He, C., Mau, K., & Xu, M. (2021). Trade Shocks and Firms Hiring Decisions: Evidence from Vacancy Postings of Chinese Firms in the Trade War. *Labour Economics*, 71(C).

Inoue, H. & Todo, Y. (2025). Disruption risk evaluation on large-scale production network with establishments and products.

Johnson, R. C. (2014). Five facts about value-added exports and implications for macroeconomics and trade research. *Journal of Economic Perspectives*, 28(2), 119–42.

Karbevska, L. & Hidalgo, C. A. (2025). Mapping global value chains at the product level. *EPJ Data Science*, 14(21).

Kaushik, D., Mishra, N., & Udgata, S. K. (2021). A Systematic Assessment of Numerical Association Rule Mining Methods. *SN Computer Science*, 2(5), 357.

Kleinman, B., Liu, E., & Redding, S. J. (2024). International friends and enemies. *American Economic Journal: Macroeconomics*, 16(4), 350–385.

Kugler, M. & Verhoogen, E. (2011). Prices, plant size, and product quality. *The Review of Economic Studies*, 79(1), 307–339.

Ma, H., Ning, J., & Xu, Y. (2025). *Reshaping Factory Asia: How Did the US-China Trade War Affect the Asia-Pacific Supply Chain?* Ssrn working paper, SSRN. Posted June 13, 2025.

Macedoni, L., Morrow, J., & Tyazhelnikov, V. (2024). *Firms in Product Space: Adoption, Growth and Competition*. CEPR Discussion Paper 18800, CEPR Press, Paris & London.

Manova, K. & Yu, Z. (2016). How firms export: Processing vs. ordinary trade with financial frictions. *Journal of International Economics*, 100(C), 120–137.

Manova, K. & Yu, Z. (2017). Multi-product firms and product quality. *Journal of International Economics*, 109(C), 116–137.

Manova, K. & Zhang, Z. (2012). Export Prices Across Firms and Destinations. *The Quarterly Journal of Economics*, 127(1), 379–436.

Mayer, T., Melitz, M. J., & Ottaviano, G. I. P. (2014). Market size, competition, and the product mix of exporters. *American Economic Review*, 104(2), 495–536.

O'Clery, N., Radcliffe-Brown, B., Spencer, T., & Tarling-Hunter, D. (2025). Deciphering the global production network from cross-border firm transactions.

Pichler, A., Pangallo, M., del Rio-Chanona, R. M., Lafond, F., & Farmer, J. D. (2020). *In and out of lockdown: Propagation of supply and demand shocks in a dynamic input-output model*. INET Oxford Working Papers 2021-18, Institute for New Economic Thinking at the Oxford Martin School, University of Oxford.

Solanki, V. K. & Patel, K. R. (2015). *A Survey on Association Rule Mining*. Technical report, Fifth International Conference on Advanced Computing & Communication Technologies.

Telikani, A., Nadali, A., & Sulaiman, R. (2020). A Survey of Evolutionary Computation for Association Rule Mining. *Information Sciences*, 524, 318–352.

Timmer, M. P., Dietzenbacher, E., Los, B., Stehrer, R., & Vries, G. J. (2015). An Illustrated User Guide to the World Input–Output Database: the Case of Global Automotive Production. *Review of International Economics*, 23(3), 575–605.

Topalova, P. & Khandelwal, A. (2011). Trade Liberalization and Firm Productivity: The Case of India. *The Review of Economics and Statistics*, 93(3), 995–1009.

Vandenbussche, H. & Viegelahn, C. (2018). Input reallocation within multi-product firms. *Journal of International Economics*, 114, 63–79.

Zhao, Y. & Bhowmick, S. S. (2003). Association Rule Mining: A Survey. *Scientific Research Publishing*.

Descriptive Statistics I

Table 1: Number of 4-digit product lines imported and exported by firms in 2004

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Mean	Std	Min	25%	50%	75%	Max
Panel A - Product diversification of firms							
HS4 products exported	10.83	29.42	1	1	3	7	556
HS4 products imported	14.65	25.03	1	2	6	17	811
Panel B - Product ubiquity							
Firms exporting a HS4 product line	544.56	985.30	1	41	174	580	11,676
Firms importing a HS4 product line	734.47	1,568.26	1	35	166	619	18,074

Note: Calculations based on 2004 firm level trade data from China Customs, after selecting relevant trade modes described in the text and aggregating products to 4-digit HS product lines (HS4). Sample features 66,688 firms and 1,361 unique HS4 product. Panel A reports how many unique HS4 products are exported and imported by firms. Panel B indicates by how many different firms a product line is exported or imported.

Reasonably favorable sample properties

- ✓ Limited levels of diversification across most firms
- ✓ Number of imported goods typically larger than number of exported goods: $N_j > N_i$
- ✓ Most goods are traded by many firms

back

Descriptive Statistics II

Table 2: Sectoral patterns of Chinese trade in 2004

			(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				% of total		# firms per HS4			% of HS4	
HS sections		HS chapters	# HS4	export	import	exports	imports	avg. RCA	interm. inputs	non-cons.
I-III	Animal and vegetable products	01-15	72	0.35	1.66	69	52	0.78	52.68	53.38
IV	Food, beverages, tobacco	16-24	80	1.60	0.45	135	79	0.70	37.81	37.81
V	Mineral products	25-27	66	2.31	9.58	88	157	0.89	89.24	89.24
VI-VII	Chemical products, plastic and rubber	28-40	275	7.21	11.21	384	816	0.85	89.22	89.22
VIII	Raw hides and skins, leather products	41-43	22	2.12	0.93	666	479	1.85	65.00	65.00
IX-X	Wood products, paper and paperboard	44-49	68	1.52	2.18	697	927	1.09	73.88	73.88
XI-XII	Textile, clothing and apparel	50-67	199	17.18	4.10	923	705	2.73	58.58	60.33
XIII-XIV	Stone, plaster, cement, ceramics; pearls, etc.	68-71	102	2.35	1.00	412	231	1.21	78.21	78.96
XV	Metal and metal products	72-83	195	7.03	7.24	472	666	1.34	79.94	90.48
XVI-XVII	Machinery and electronics	84-85	132	45.23	36.29	941	2,018	1.11	32.78	87.79
XVIII	Vehicles and other transport equipment	86-89	37	3.73	2.16	322	147	0.98	20.11	83.18
XIX	Precision instruments and measurement apparatus	90-92	70	3.48	6.53	418	870	1.81	31.14	71.94
XX	Other manufactured goods	93-97	43	5.89	0.32	1,160	743	2.67	22.17	42.63

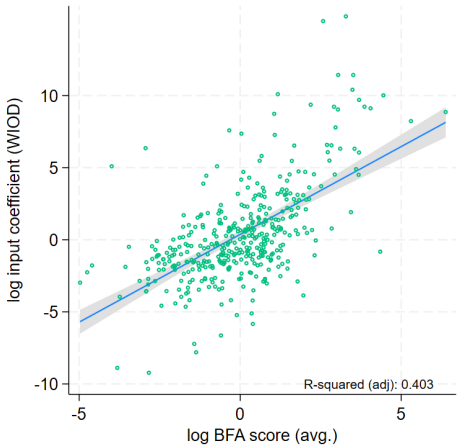
Noticeable features

- Despite substantial exports, many importers in HS Chapters 84-85, which comprises lots of capital goods
- High import shares in HS Sections V-VII, which are mostly intermediate input goods

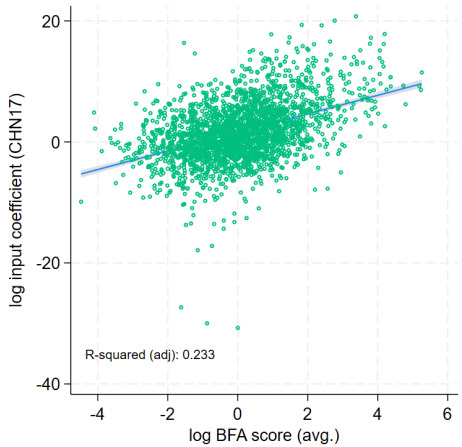
back

Correlation with other IO tables

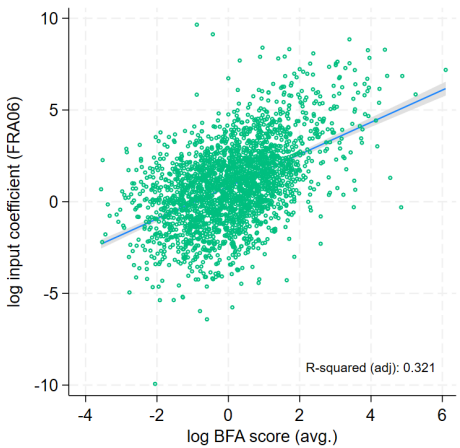
China 2004 (WIOD-16)



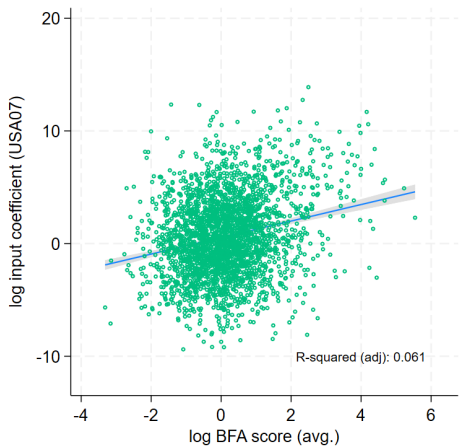
China 2017 (NBS SUT)



France 2006 (Berlingieri et al., 2021)



USA 2007 (BEA)



Note: Figures display (centered) correlations between aggregated BFA-based linkages and coefficients from other IO tables.

▸ back

Correlation with other IO tables

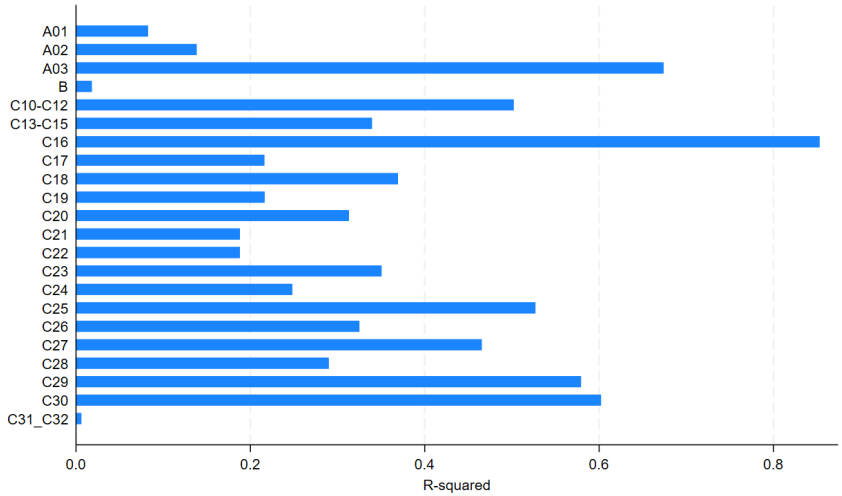
Dep. var.: IO coefficients	(1)	(2)	(3)	(4)	(5)
Country	China 2004	China 2017	France 2006	USA 2007	
Source	WIOD-16	NBS SUT	Berlingieri et al. (2021)	BEA	AIPNET
Industry classifcation	ISIC Rev. 4	CIC Rev. 3.1	ISIC Rev. 2	NAICS	HS4 Rev.'02
Panel A: BFA mapping (baseline)					
BFA-score (ln)	1.557**	1.669**	1.150**	0.878**	0.213**
	(0.154)	(0.070)	(0.038)	(0.077)	(0.004)
Observations	457	3,871	3,638	3,743	50,618
R-squared(within)	0.321	0.171	0.255	0.047	0.069
Panel B: BFA mapping (post-processed)					
BFA-score (ln)	1.477**	1.776**	1.117**	0.935**	0.232**
	(0.148)	(0.074)	(0.041)	(0.080)	(0.005)
Observations	446	3,209	3,113	3,572	39,281
R-squared(within)	0.312	0.192	0.244	0.053	0.074
Panel C: BFA mapping (pre- and post-processed)					
BFA-score (ln)	1.224**	1.545**	0.931**	0.804**	0.166**
	(0.102)	(0.063)	(0.033)	(0.072)	(0.005)
Observations	398	2,381	2,355	2,665	19,384
R-squared(within)	0.403	0.233	0.321	0.061	0.063

Note: BFA-based mappings inferred from China's 2004 firm-level transactions data. Input-output data obtained from sources indicated in column headings. WIOD-16: 2016 release of the World Input Output Database described in Timmer et al. (2015); NBS SUT: Chinese Supply and Use Tables from the National Bureau of Statistics of China; France 2006: based on linkages inferred by Berlingieri et al. (2021); BEA: 2007 Make-and-Use tables published by the US Bureau of Economic Analysis. Column (5) summarizes results of estimating linear probability model regressing binary linkages in AIPNET on (log) BFA-scores. Standard errors in parentheses. Statistical significance: ^{a} $p < 0.1$, * $p < 0.05$, ** $p < 0.01$.

back

Correlation BFA vs. WIOD, by ISIC4 sector

Sector-wise correlation (R-squared)



Conditional correlation

	(1)	(2)
Dep. var.: IO coefficient	RCA output	RCA input
BFA-score (log)	1.568**	1.549**
	(0.229)	(0.186)
$\times RCA(output)$	-0.238[a]	
	(0.138)	
$\times RCA(input)$		-0.252*
		(0.113)
Observations	398	398
R-squared (within)	0.413	0.412

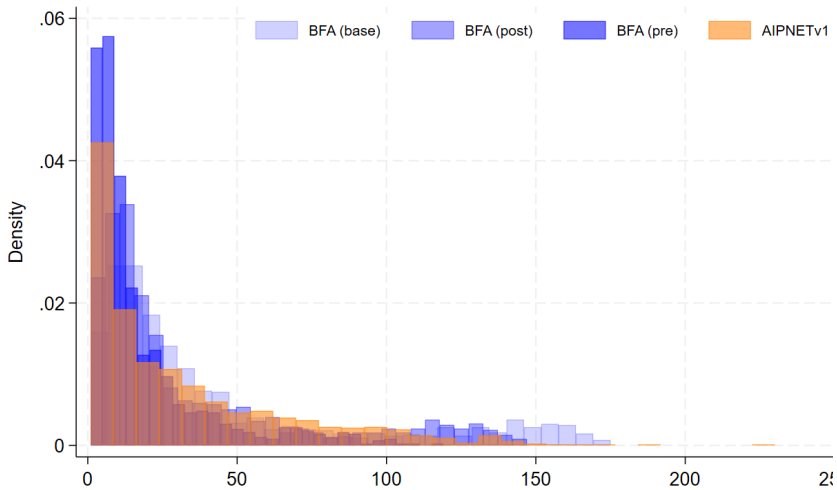
- Weaker correlation if RCA>1 in input industry; domestic sourcing?

back

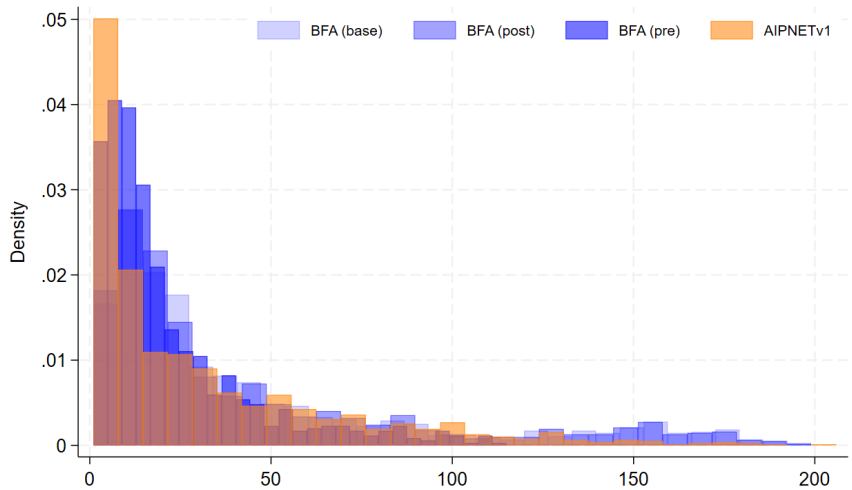
Comparison with other mappings: AIPNET

Cross-product distributions of output complexity and input ubiquity

(a) Output complexity (*j*s per *i*)



(b) Input ubiquity (*i*s per *j*)



Note: Figures display distribution of the number of binary directional IO-linkages (at HS4 level; 4-digit HS Rev. 2002) inferred with alternative methods

▸ back

Comparison with other mappings: AIPNET

Output complexity and input ubiquity in BFA vs. LLM-based inference

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. var.: AIPNET →	Output complexity (#*j*s per *i*)			Input ubiquity (#*i*s per *j*)		
BFA-mapping	baseline	post	pre & post	baseline	post	pre & post
Coefficient	0.361**	0.435**	0.714**	0.125**	0.244**	0.749**
	(0.019)	(0.024)	(0.045)	(0.023)	(0.023)	(0.045)
Observations	1,154	1,154	1,154	1,159	1,159	1,159
R-squared (adj.)	0.230	0.214	0.179	0.025	0.089	0.195

Note: Table displays coefficients obtained from complexity and ubiquity measures calculated from AIPNET on corresponding metrics obtained from BFA-based mappings. Standard errors in parentheses. Statistical significance: ^{a} $p < 0.1$, * $p < 0.05$, ** $p < 0.01$.

▸ Back

Examples of inferred linkages

HS4: 8471	**Automatic data processing machines and units thereof (Computers)**	**Share in exports (2004): 7.85%**	**# HS4 inputs: 54**

Top HS4 Inputs			BFA score	AIPNET	Ma *et al.* (2025)
1	8473	Machinery; parts and accessories (other than covers, carrying cases and the like) suitable for use solely or principally with machines of headings 84.70 to 84.72	0.121	✓	✓
2	8542	Electronic integrated circuits	0.089	✓	✓
3	8541	Diodes, transistors, similar semiconductor devices; including photovoltaic cells assembled or not in modules or panels, light-emitting diodes (LED), mounted piezo-electric crystals	0.063	✓	✓
4	8532	Electrical capacitors; fixed, variable or adjustable (pre-set)	0.056	✓	✓
5	8534	Circuits; printed	0.053	✓	
6	8533	Electrical resistors (including rheostats and potentiometers), excluding heating resistors	0.052	✓	
7	8504	Electric transformers, static converters (e.g. rectifiers) and inductors	0.049	✓	
8	9013	Liquid crystal devices not constituting articles provided for more specifically in other headings; lasers, not laser diodes; other optical appliances and instruments n.e.c. in this chapter	0.047		✓
9	8517	Telephone sets, including telephones for cellular networks or for other wireless networks; other apparatus for the transmission or reception of voice, images or other data (including wired/wireless networks), excluding items of 8443, 8525, 8527, or 8528	0.042		✓
10	8524	Flat panel display modules, whether or not incorporating touch-sensitive screens	0.040		

back

Examples of inferred linkages

HS4: 6204	Women's Suits, jackets, dresses, skirts, trousers, overalls, breeches and shorts (not knitted or crocheted)	Share in exports (2004): 1.09%	# HS4 inputs: 46

Top HS4 Inputs			BFA score	AIPNET	Ma *et al.* (2025)
1	9606	Buttons, press-fasteners, snap-fasteners and press-studs, button moulds and other parts of these articles; button blanks	0.345		
2	9607	Slide fasteners and parts thereof	0.294		
3	5807	Labels, badges and similar articles; of textile materials, in the piece, in strips or cut to shape or size, not embroidered	0.288	✓	
4	5407	Woven fabrics of synthetic filament yarn, including woven fabrics obtained from materials of heading no. 5404	0.277	✓	✓
5	5208	Woven fabrics of cotton, containing 85% or more by weight of cotton, weighing not more than 200 g/m2	0.263	✓	
6	5209	Woven fabrics of cotton, containing 85% or more by weight of cotton, weighing more than 200g/m2	0.232	✓	✓
7	5210	Woven fabrics of cotton, containing less than 85% by weight of cotton, mixed mainly or solely with man-made fibres, weighing not more than 200 g/m2	0.210	✓	✓
8	5309	Woven fabrics of flax	0.175	✓	
9	5603	Nonwovens; whether or not impregnated, coated, covered or laminated	0.174	✓	
10	5806	Fabrics; narrow woven, other than goods of heading no. 5807; narrow fabrics consisting of warp without weft assembled by means of an adhesive (bolducs)	0.171	✓	

back

Examples of inferred linkages

HS4: 7601	**Unwrought aluminum**	**Share in exports (2004): 0.36%**	**# HS4 inputs: 7**

Top HS4 Inputs			BFA score	AIPNET	Ma *et al.* (2025)
1	7602	Aluminium; waste and scrap	0.065	✓	✓
2	2818	Aluminium oxide (including artificial corundum); aluminium hydroxide	0.023	✓	
3	7404	Copper; waste and scrap	0.007		
4	7902	Zinc; waste and scrap	0.006		
5	8454	Converters, ladles, ingot moulds and casting machines; of a kind used metallurgy or in metal foundries	0.003	✓	
6	8429	Bulldozers, graders, levellers, scrapers, angledozers, mechanical shovels, excavators, shovel loaders, tamping machines and road rollers, self-propelled	0.001		
7	8474	Machinery for sorting, screening, separating, washing, crushing, grinding, mixing or kneading earth, stone, ores in solid form, shaping, moulding machinery for solid mineral fuels	0.001		
8	*0909*	*Seeds of anise, badian, fennel, coriander, cumin, caraway or juniper*	*0.000**		
9	*7501*	*Nickel mattes; nickel oxide sinters and other intermediate products of nickel metallurgy*	*0.000**		
10	*2606*	*Aluminium ores and concentrates*	*0.000**	✓	

back

Examples of inferred linkages

HS4: 4011	**New pneumatic tyres, of rubber**	**Share in exports (2004): 0.32%**	**# HS4 inputs: 40**

Top HS4 Inputs			BFA score	AIPNET	Ma *et al.* (2025)
1	4001	Natural rubber, balata, gutta-percha, guayule, chicle and similar gums; in primary forms or in plates, sheets or strip	0.092	✓	✓
2	5902	Textile fabrics; tyrecord of high tenacity yarn of nylon or other polyamides polyesters or viscose rayon	0.040		
3	8714	Vehicles; parts and accessories of heading no. 8711 to 8713	0.019	✓	
4	2503	Sulphur of all kinds; other than sublimed, precipitated and colloidal sulphur	0.018		
5	4002	Synthetic rubber and factice derived from oils, in primary forms or in plates, sheets or strip; mixtures of heading no. 4001 and 4002, in primary forms or in plates, sheets or strip	0.018	✓	✓
6	2802	Sulphur; sublimed or precipitated, colloidal sulphur	0.016		
7	4013	Inner tubes, of rubber	0.014		
8	4005	Compounded rubber, unvulcanised, in primary forms or in plates, sheets or strip	0.009	✓	✓
9	3823	Industrial monocarboxylic fatty acids; acid oils from refining; industrial fatty alcohols	0.007		
10	3812	Prepared rubber accelerators; compound plasticisers for rubber or plastics, n.e.c. or included; anti-oxidising preparations and other compound stabilisers for rubber or plastics	0.007	✓	✓

back

Examples of inferred linkages

HS4: 8609	**Containers; (including containers for transport of fluids) specially designed and equipped for carriage by one or more modes of transport**	**Share in exports (2004): 0.67%**	**# HS4 inputs: 14**

Top HS4 Inputs			BFA score	AIPNET	Ma *et al.* (2025)
1	4412	Plywood, veneered panels and similar laminated wood	0.041	✓	
2	7208	Iron or non-alloy steel; flat-rolled products of a width of 600mm or more, hot-rolled, not clad, plated or coated	0.033	✓	
3	3208	Paints, varnishes; (enamels and lacquers) based on synthetic polymers or chemically modified natural polymers, dispersed or dissolved in a non-aqueous medium	0.014		
4	7225	Alloy steel flat-rolled products, of a width 600mm or more	0.009	✓	
5	3814	Organic composite solvents and thinners, not elsewhere specified or included; prepared paint or varnish removers	0.009		
6	7219	Stainless steel; flat-rolled products of width of 600mm or more	0.005	✓	
7	8467	Tools; for working in the hand, pneumatic, hydraulic or with self-contained electric or non-electric motor	0.003		
8	7216	Iron or non-alloy steel, angles, shapes and sections	0.003	✓	
9	8481	Taps, cocks, valves and similar appliances for pipes, boiler shells, tanks, vats or the like, including pressure-reducing valves and thermostatically controlled valves	0.003	✓	
10	8515	Electric (electrically heated gas) soldering, brazing, welding machines and apparatus, capable or not of cutting, electric machines and apparatus for hot spraying of metals or sintered carbides	0.002		

back

Validation: import-export co-movement

Regress avg. annual import growth on downstream export growth

- Data: CEPII Baci; Chinese exports and imports in years 2003-2008

$$\underbrace{\%\Delta x_j}_{\text{imp. growth}} = \alpha + \beta \underbrace{\%\Delta \bar{y}_j^m}_{\text{exp. growth}} + \varepsilon_j$$

- Downstream export growth: weighted avg. of $\%\Delta y_i$,

$$\%\Delta \bar{y}_j^m \equiv \sum_i (w_{ij}^m \times \%\Delta y_i); \quad \sum_i w_{ij}^m = 1,$$

→ w_{ij}^{BFA} = normalized $Score_{ij}^{BFA}$

→ Noisy trade data: remove outliers → top and bottom 1% of $\%\Delta \bar{y}_j^m$ and $\%\Delta x_j^m$

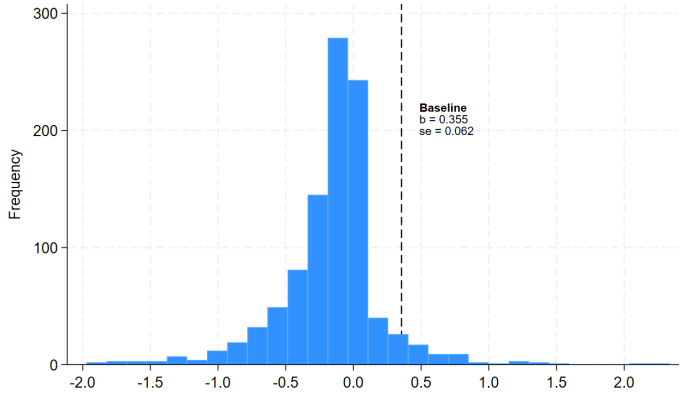
Validation: import-export co-movement

China's avg. import growth rate 2003-2008

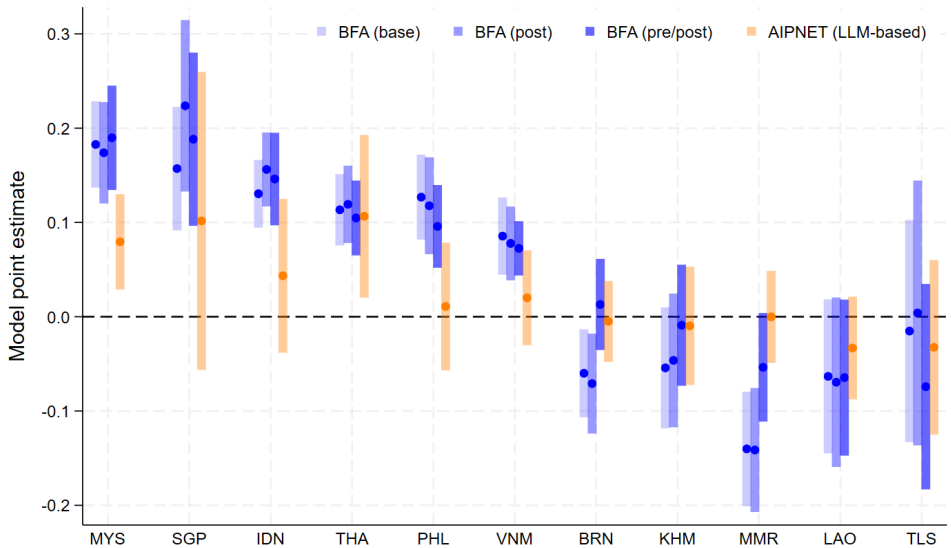
Dep. var.: Import growth	(1)	(2)	(3)	(4)
	BFA (pre-& post)		LLM-based (AIPNET)	
	HS2	HS4	HS2	HS4
Downstream export growth	0.832**	0.355**	0.796**	0.218**
	(0.176)	(0.062)	(0.229)	(0.064)
Observations (# products)	84	888	92	1,087
R-squared (adjusted)	0.204	0.035	0.109	0.010

SE in parentheses. Significance: ^{a} $p < 0.1$, * $p < 0.05$, ** $p < 0.01$

Placebo col (2): 1,000 random links



Validation other SEA economies, 2003-08



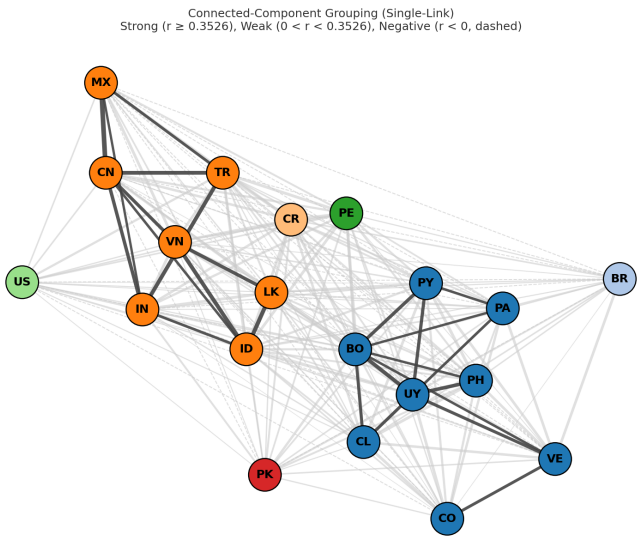
back

Cross-country correlations and persistence of mappings

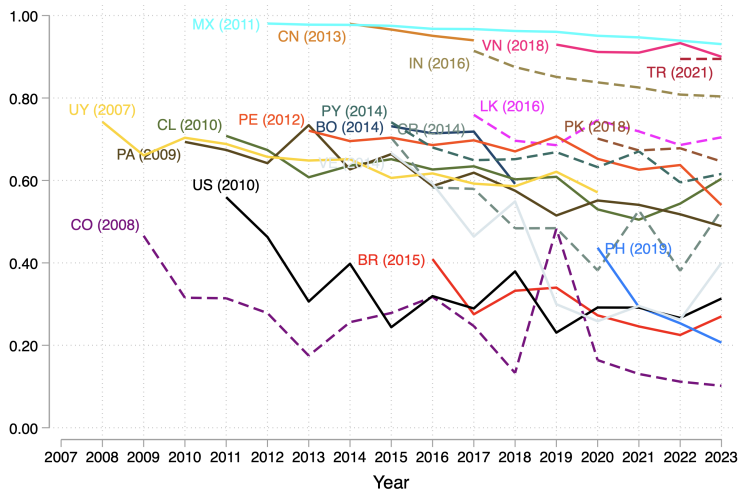
Similarity of HS4-level input-output mappings across countries and over time

- Data source: *Panjiva Global Trade Insights*
- Firm-level shipments for many low-/middle-income economies; and US
- Coverage/quality of data varies (see paper Appendix B)

(a) Cross-country similarity



(b) Persistence of mappings



Note: Panel (a) displays clusters of economies based on the cross-country correlation of inferred linkages. Same cluster (color), if pairwise correlation is above threshold (more than twice the avg. correlation). Panel (b) plots the time-series correlation of IO linkages for each economy relative to first year in the data.

back

More disaggregated mappings

Based on Chinese HS6 rev.2017 trade 2013-17

- Greater detail comes with more data requirements; stack several years
- Challenge: newly introduced HS6 codes; fewer observations, more prone to confounders

▸ Batteries ▸ Solar Panels ▸ EVs ▸ wind

Side project with Dutch data (PESCA-NL) → www.pesca-dashboard.com

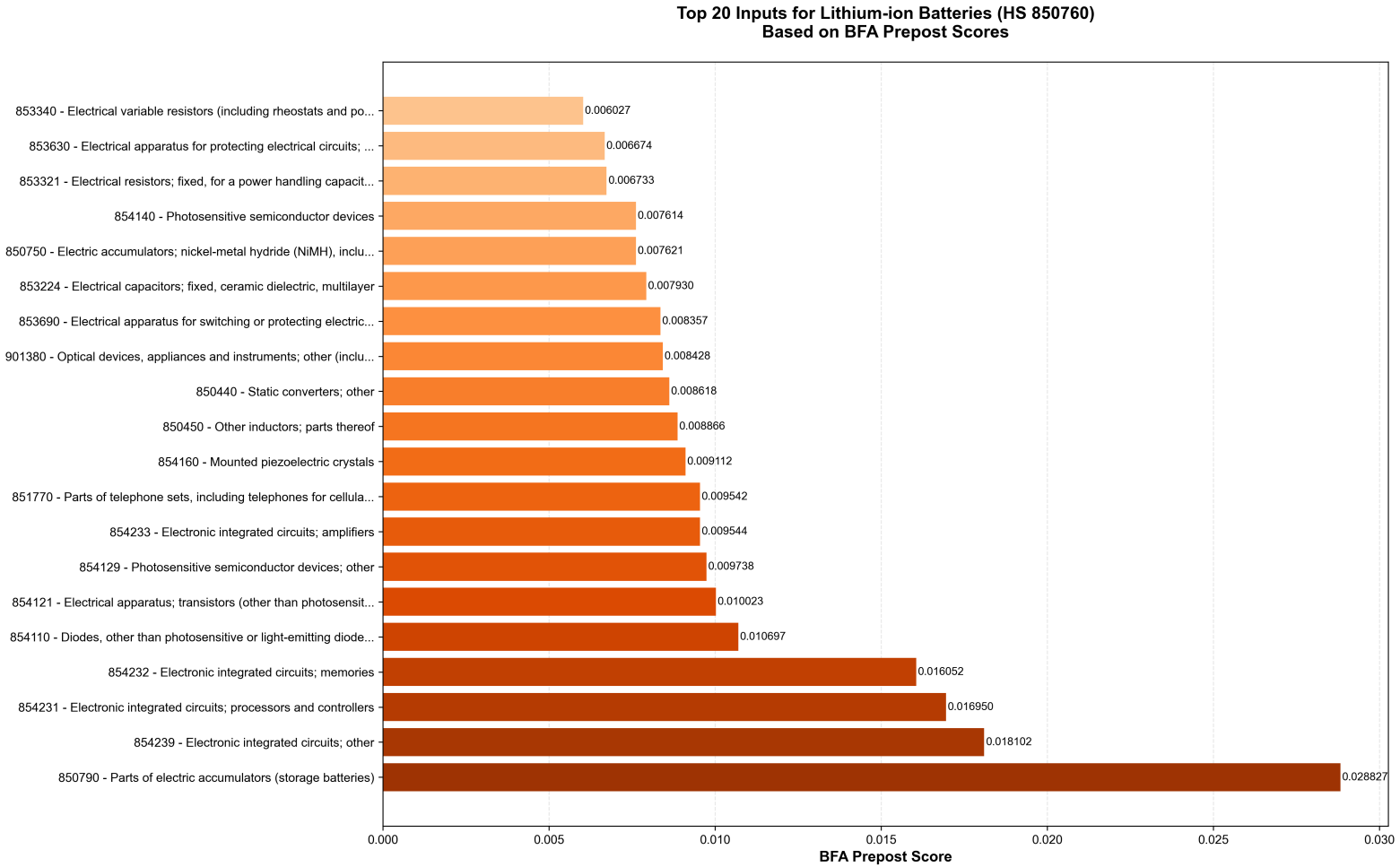
- Remove non-manuf. firms through industry code
- Run BFA at HS4 and HS6 level, combine results
- Some true intra-HS4 linkages, but 4-digit level usually enough

▸ back

More disaggregated mappings

Based on Chinese HS rev.2017 trade 2013-17

`Lithium Batteries HS 8507.60`

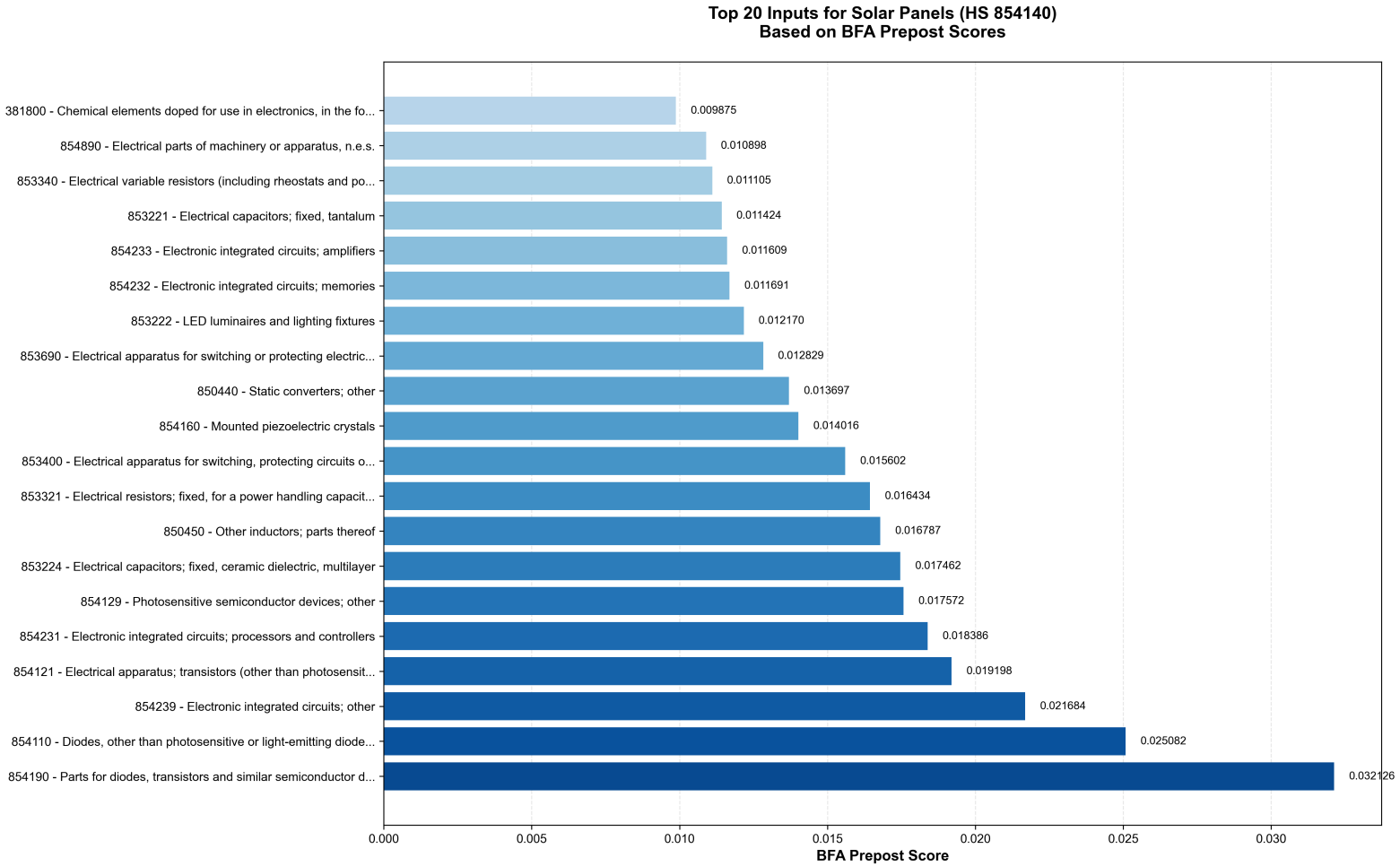


back

More disaggregated mappings

Based on Chinese HS rev.2017 trade 2013-17

`Solar Panels HS 8541.40`

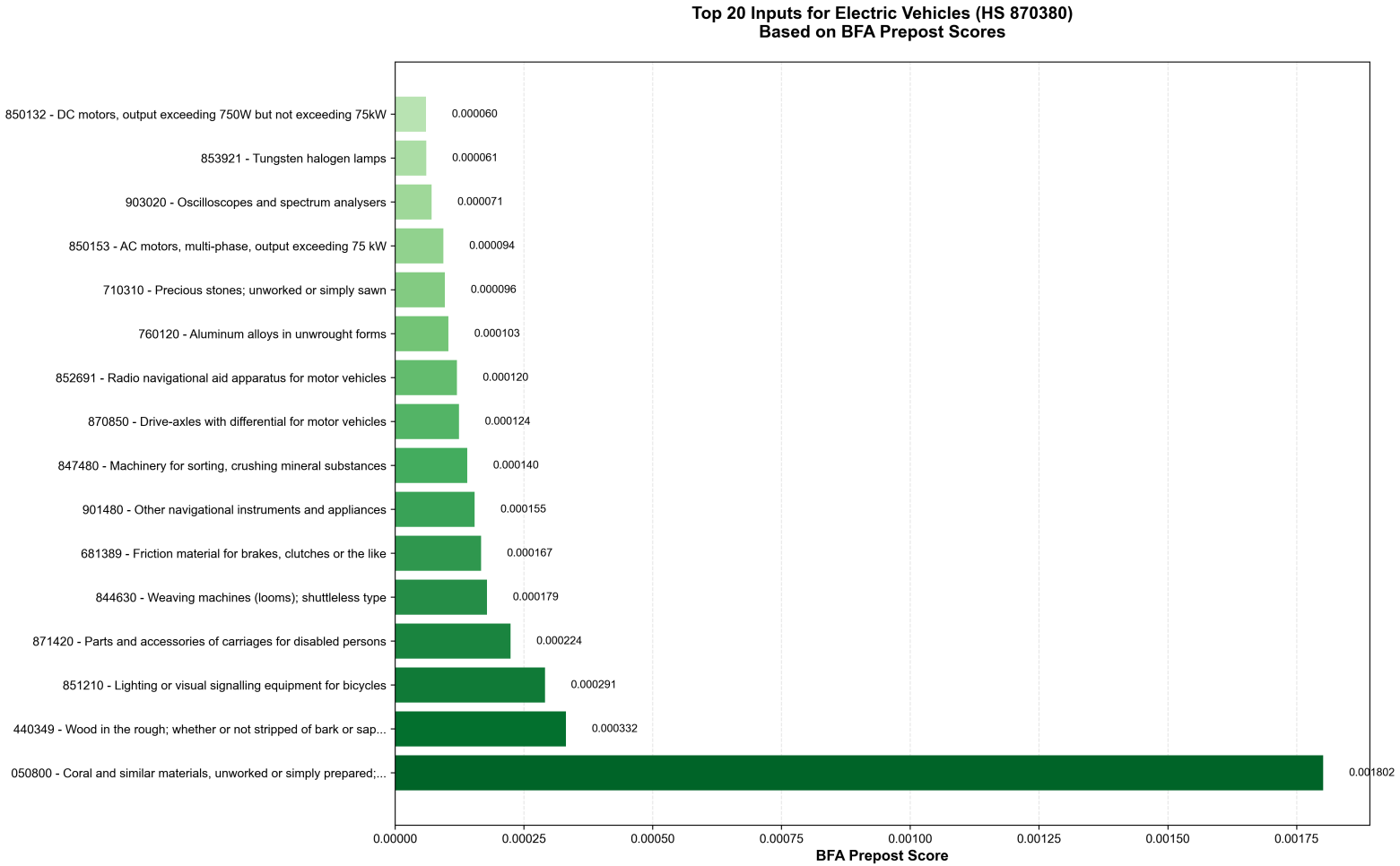


back

More disaggregated mappings

Based on Chinese HS rev.2017 trade 2013-17

`Electric Vehicles HS 8703.80`

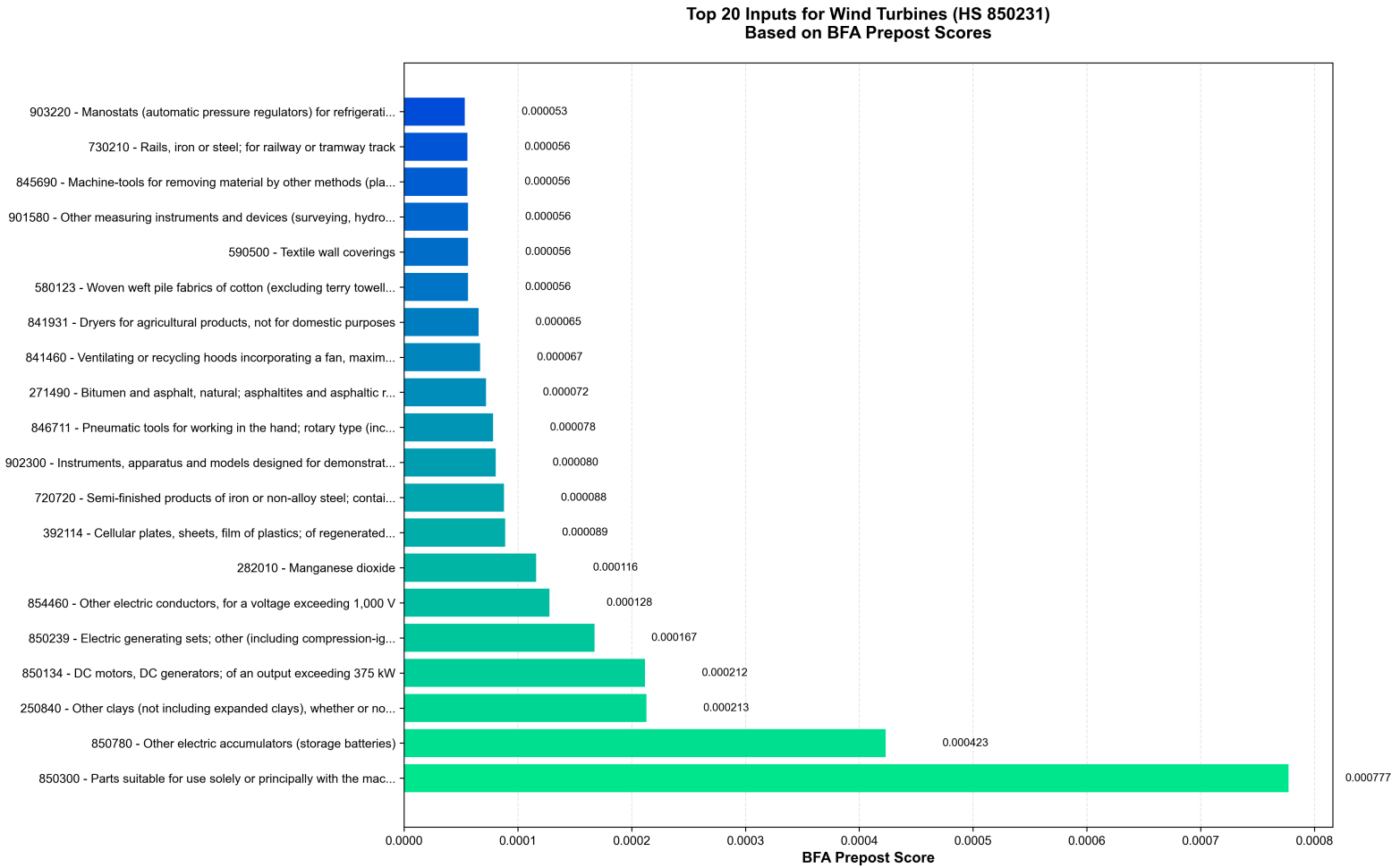


back

More disaggregated mappings

Based on Chinese HS rev.2017 trade 2013-17

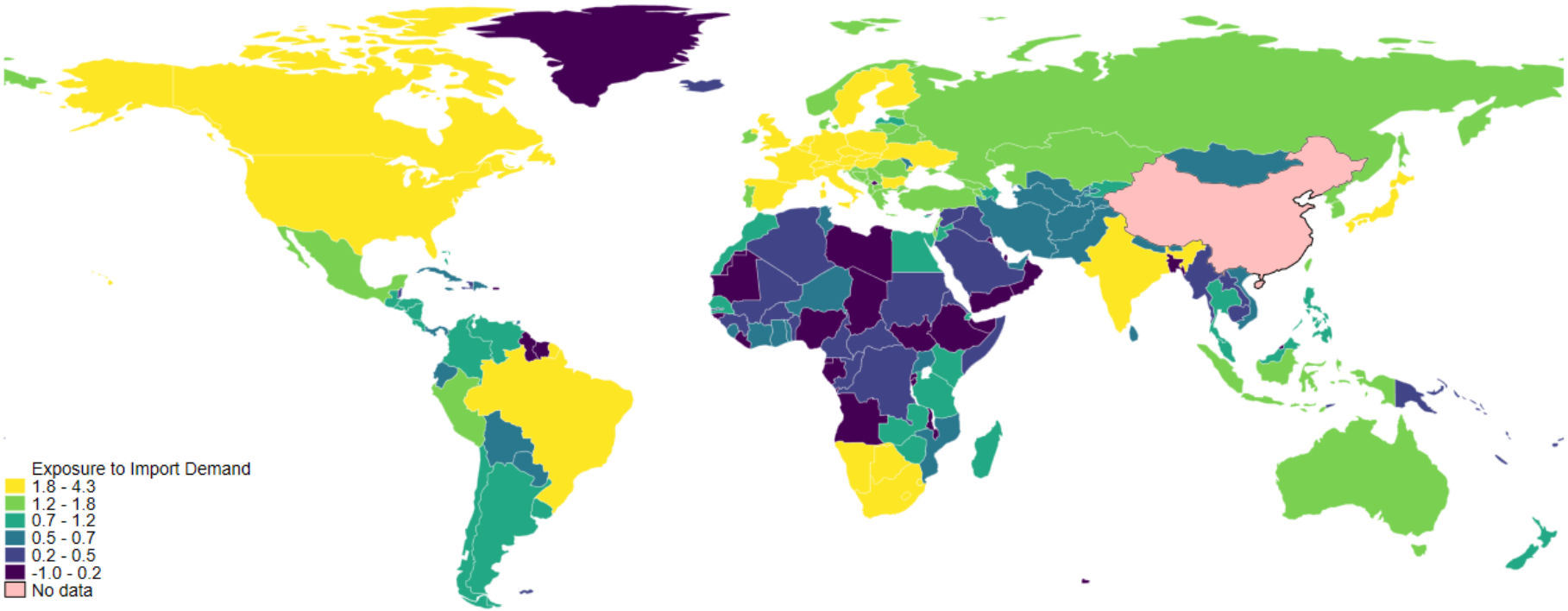
`Wind Turbines HS 8502.31`



back

"China Shock 1.0" - Export opportunities

Average RCA in products upstream of post-WTO entry China's exports

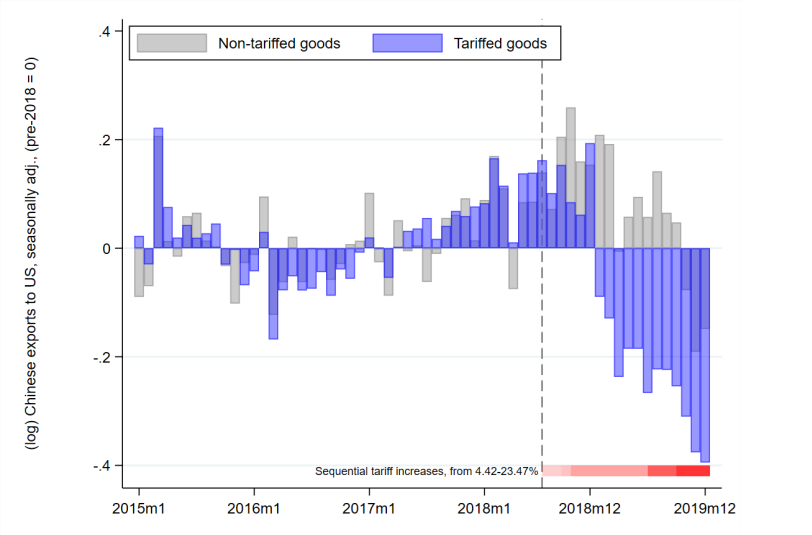


Back

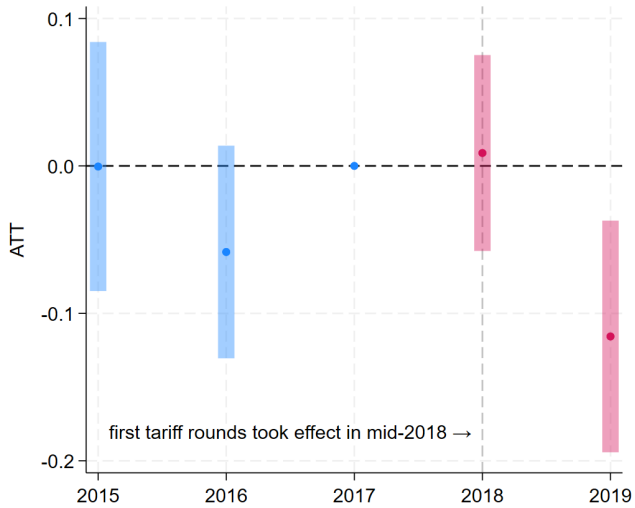
Application II - *US-China Trade War 2018-19*

US tariffs imposed in 2018-19 significantly lower imports from China

(a) Chinese monthly exports to US



(b) Annual export-related imports



- Inputs of the most tariff-exposed exports decline significantly in 2019
- Measure of input exposure: $\text{Exp}_{\tau(\text{US})\to j} \equiv \sum_i \left(Score_{ij}^{\text{BFA}} \times \tau_i^{\text{US}}\right) \times \bar{s}_{i,US}$
  - Estimated average coeff. = -0.028 ▸ Table
  - US not always main export market → ≈ 11% overall; 0.25% at 5$^{\text{th}}$ pctl; 30% at 95$^{\text{th}}$ pctl

US-China Trade War 2018-19

	(1)	(2)	(3)	(4)
Frequency and aggregation:	Monthly (HS6 level)	Annual (HS4 level)		
Dep. var. (in logs) and trade partner:	Exports to US	Exports to US	Exports to World	Imports of Inputs
US tariff $(k, t) = 1$	-0.290**			
	(0.014)			
US tariff $(k) = 1 \times$ (year $= 2019$)		-0.218**	-0.054	
		(0.080)	(0.045)	
US tariff exposure of inputs $\times$ (year $= 2019$)				-0.028*
				(0.011)
Fixed effects	HS6×month	HS4, year	HS4, year	HS4, year
Lin. trends (product-level)	✓	✓	✓	✓
Other controls				✓
Cluster	4,336	1,180	1,224	1,139
Observations	218,142	5,793	6,119	5,694

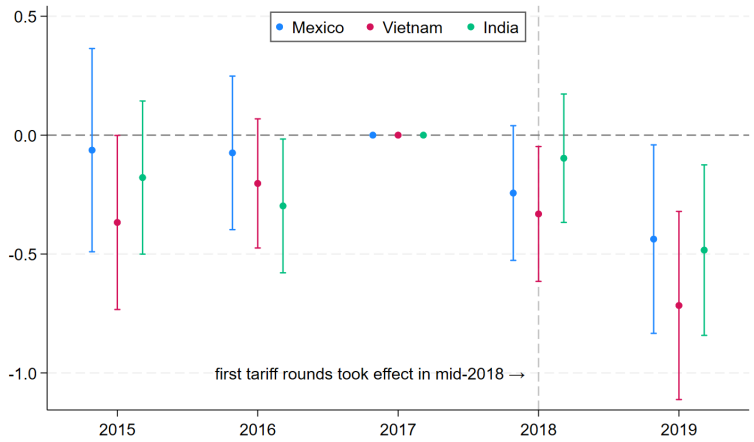
Note: Table reports regressions of China's export and import adjustments during the 2018-2019 trade war. Column (1) uses monthly HS6 shipment data and the U.S. tariff lists compiled by He et al. (2021) for 2015–2019; the coefficient on US tariff(k, t) compares Chinese exports to the United States in tariffed versus non-tariffed HS6 products. Columns (2)-(4) use annual HS4 trade flows from CEPII BACI (2025 release `BACI_HS12_V202501`) over 2015–2019. Columns (2) and (3) estimate the differential change in 2019 for HS4 products subject to U.S. tariffs, for exports to the United States and for China's worldwide exports, respectively. Column (4) examines the upstream margin: it relates China's imports to a measure of input exposure to U.S. tariffs, constructed by weighting U.S.-tariffed downstream export products using our post-processed BFA scores. All specifications include the fixed effects and product-specific linear trends indicated in the table; column (4) additionally includes control variables. Standard errors (in parentheses) are clustered at the product level. Statistical significance: ^{a} $p < 0.1$, * $p < 0.05$, ** $p < 0.01$.

Application II - *Trade War effect on 3rd countries*

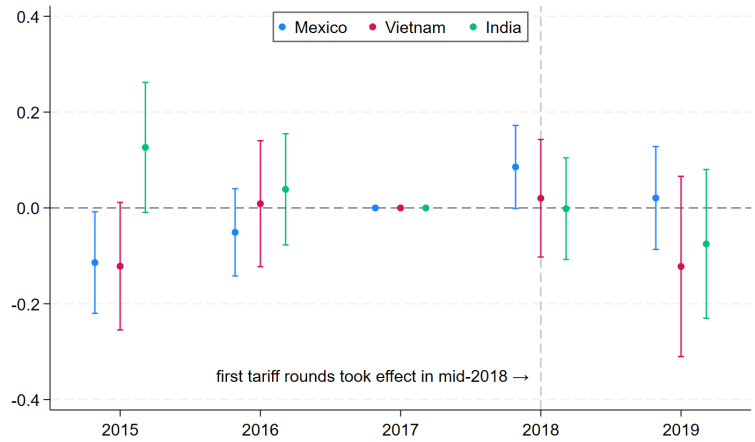
Are China's input suppliers also affected?

- India, Mexico, and Vietnam among main input suppliers for China's US exports

(a) Exports to China



(b) Export-related imports



- Input-exposure to US tariffs: $\text{Exp}_{\tau(\text{US})\to j(\text{TC})} \equiv (\text{Exp}_{\tau(\text{US})\to j} > 0) \times \bar{s}_{j,CN}$
  - Exports of exposed inputs fall if China main market
- Imports of export-related inputs only with indicative changes ▸ Table
  - Only one post-treatment period, adjustments take time
  - Also: potentially compounding measurement error

Trade war effect on third countries

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Country:	Mexico			Vietnam			India		
Period of sample:	2015–2019		2015–2023	2015–2019		2015–2023	2015–2019		2015–2023
Dep. var. (in logs) and trade partner:	Exports to China	Imports fr. World	Imports fr. World	Exports to China	Imports fr. World	Imports fr. World	Exports to China	Imports fr. World	Imports fr. World
Output exposure to US tariff × (year≥2019)	-1.298*			-1.239**			-1.015*		
	(0.583)			(0.392)			(0.419)		
Input exposure to US tariff × (year≥2019)		-0.181	-0.169[a]		-0.076	-0.027		-0.055	-0.129*
		(0.111)	(0.100)		(0.049)	(0.049)		(0.067)	(0.065)
Fixed effects	HS4, year	HS4, year	HS4, year	HS4, year	HS4, year	HS4, year	HS4, year	HS4, year	HS4, year
Lin. trends (product-level)	✓	✓	✓	✓	✓	✓	✓	✓	✓
Other controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Control for Covid-19			✓			✓			✓
Cluster	768	1,214	1,214	983	1,210	1,210	1,063	1,211	1,211
Observations	3,369	6,056	10,895	4,419	6,036	10,847	4,911	6,041	10,863

Note: Table presents estimates of how additional U.S. tariffs on Chinese goods transmit upstream to third-country supply of intermediate inputs to China. Sample covers annual HS4-level trade flows from CEPII BACI (`BACI_HS12_V202501`). Dependent variables, trade partner and sample periods are indicated in column headings. Explanatory variable *Output exposure* captures exposure of the country's exports to China through downstream products j that face U.S. tariffs. *Input exposure* captures exposure of the country's upstream inputs k through their predicted use (BFA scores) in producing tariff-hit outputs; the measure is weighted by the country's pre-trade-war export share of product j shipped to China and is scaled by 100 (see main text). All specifications, include HS4 and year fixed effects and HS4-specific linear trends. "Other controls" include controls for U.S. tariffs on China and China's retaliatory tariffs on U.S. goods. For regressions spanning 2015-2023, we additionally include indicator variables for 2020-2022 to absorb COVID-19 disruptions. Standard errors in parentheses are clustered at the HS4 level. Statistical significance: ^{a} $p < 0.1$, * $p < 0.05$, ** $p < 0.01$.

▸ Back