



GLL Datasets, Responsible and Efficient AI by AEA and CS-Sophomore Perspective

September 23, 2026

Global Linkages Lab

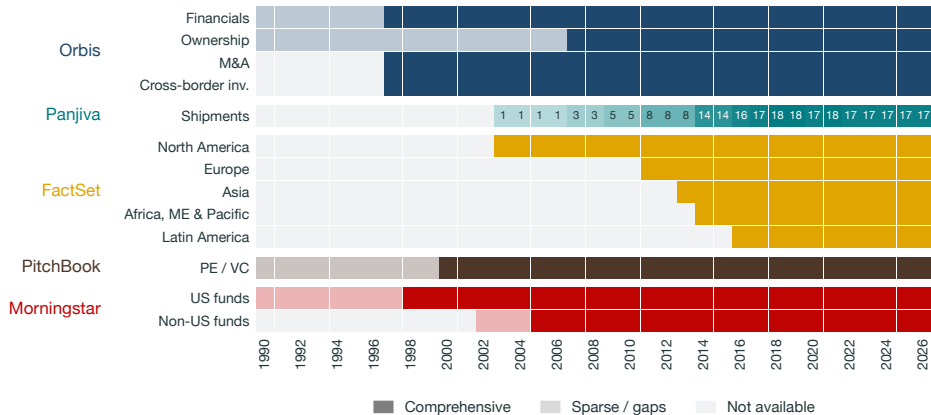


What GLL got from commercial providers so far?



Provider	Dataset	What it captures	Unit
MOODY'S	Orbis Financials, Ownership, M&A, Cross-border Investment	Financial statements, ownership links, for public <i>and</i> private firms worldwide	firm-year
S&P Global	Panjiva Supply Chain Intelligence	1B+ customs records of company-to-company shipments (HS-coded)	shipment, firm-to-firm, week
FACTSET	FactSet Supply Chain Relationships including domestic	Disclosed customers, suppliers, competitors and partners of public firms	firm-to-firm, month
PitchBook	PitchBook: PE and VC investments	company financials and investment, transactions	investee-to-investor, month/year
MORNINGSTAR®	Morningstar: Portfolio Holdings	Only mutual funds (no hedge, pension, insurance, BANK, HH, government)	fund-level; position-level (month/quarter end)
Bloomberg	Bloomberg Terminal (12 belong to GLL, 2 loaned to Rhodes+Library)	Market data, expectations, quantity and price	firm-security-price, day

Coverage by Country and Time: Achilles' heel for research



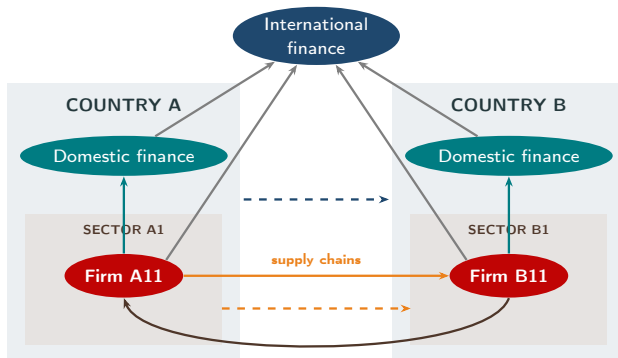
- Basante Pereira, G., and Simonovska, I. 2026. "Contract Enforcement and Young Firm Capital Structure: A Global Perspective," (NBER Working Paper 34985)
- Faccio, M., Xu, J., 2025. Taxes and Private Firms' Capital Structure Choices (NBER Working Paper 33685)
- Jaimovich, N., Terry, S.J., and Vincent, N., 2023. The Empirical Distribution of Firm Dynamics and Its Macro Implications (NBER Working Paper 31337)
- Gu, G.W. and Hale, G., 2023. Climate risks and FDI. Journal of International Economics, p.103731. (NBER Working Paper 30452)
- Bilicka, K.A., Guceri, I. and Koumanakos, E., 2023. Dividend Taxation and Firm Performance with Heterogeneous Payout Responses (NBER Working Paper 30808)
- Garcia-Bernardo, J., Janský, P. and Zucman, G., 2022. Did the Tax Cuts and Jobs Act Reduce Profit Shifting by US Multinational Companies? (NBER Working Paper 30086)
- Boar, C., Gorea, D. and Midrigan, V., 2022. Why Are Returns to Private Business Wealth So Dispersed? (NBER Working Paper 29705)
- Kalemli-Ozcan, S., Liu, X. and Shim, I., 2021. Exchange rate fluctuations and firm leverage. IMF Economic Review, 69, pp.90-121. (NBER Working Paper 28608)
- Hassan, T.A., Schreger, J., Schwedeler, M. and Tahoun, A., 2021. Sources and transmission of country risk. (NBER Working Paper 29526)
- Gourinchas, P.O., Kalemli-Özcan, ., Penciakova, V. and Sander, N., 2021. Fiscal Policy in the Age of COVID: Does it 'Get in all of the Cracks?' (NBER Working Paper 29293)
- Bilicka, K.A. and Scur, D., 2021. Organizational capacity and profit shifting . (NBER Working Paper 29225)
- Impink, S.M., Prat, A. and Sadun, R., 2021. Communication within firms: evidence from CEO turnovers. (NBER Working Paper 29042)
- Hansen, S., Ramdas, T., Sadun, R. and Fuller, J., 2021. The demand for executive skills. (NBER Working Paper 28959).



- (2022, w30087) Fang, Hardy and Lewis, "Who Holds Sovereign Debt and Why It Matters."
- (2024, w32193) Lerner, Liu, Moscona and Yang, "Appropriate Entrepreneurship? The Rise of China and the Developing World."
- (2024, w32275) Beck, Coppola, Lewis, Maggiori, Schmitz and Schreger, "The Geography of Capital Allocation in the Euro Area."
- (2025, w32453) Du and Huber, "Dollar Asset Holdings and Hedging Around the Globe."
- (2026, w33080) Sabrina T. Howell, Dean Parker, and Ting Xu "Tyranny of the Personal Network."
- (2025, w33435) Abhishek Bhardwaj, Abhinav Gupta, and Sabrina T. Howell "Capital Structure and Firm Outcomes."
- (2025, w33987) Xiaoyong (Jack) Fu and Lucian A. Taylor, "Due Diligence and the Allocation of Venture Capital."
- Lerner, Nanda and Townsend w27150 and Gompers, Gornall, Kaplan and Strebulaev w27824



- **ORBIS:** Cravino and Levchenko (2017, QJE), "Multinational Firms and International Business Cycle Transmission."
- **ORBIS:** Gopinath, Kalemli-Özcan, Karabarbounis and Villegas-Sanchez (QJE 2017), "Capital Allocation and Productivity in South Europe."
- **Morningstar:** Coppola, Maggiori, Neiman and Schreger (2021, QJE), "Redrawing the Map of Global Capital Flows: The Role of Cross-Border Financing and Tax Havens."
- **Morningstar:** Maggiori, Neiman and Schreger (2020, JPE), "International Currencies and Capital Allocation,"
- **PitchBook:** Lerner, Liu, Moscona and Yang, "Appropriate Entrepreneurship? The Rise of China and the Developing World," forthcoming JPE?
- **Only part of Pitchbook:** Howell's AER 2017 "Financing Innovation: Evidence from RD Grants"
- **FactSet supply chain:** Carvalho, Nirei, Saito and Tahbaz-Salehi (2021, QJE)
- **Factset ownership:** (Koijen and Yogo 2019 JPE, Koijen, Richmond and Yogo 2024 REStud)
- **Panjiva/customs/VAT:** (Bernard, Dhyne, Magerman, Manova and Moxnes 2022 JPE; Dhyne, Kikkawa, Mogstad and Tintelnot 2021 REStud; Eaton, Kortum and Kramarz 2022 Econometrica)



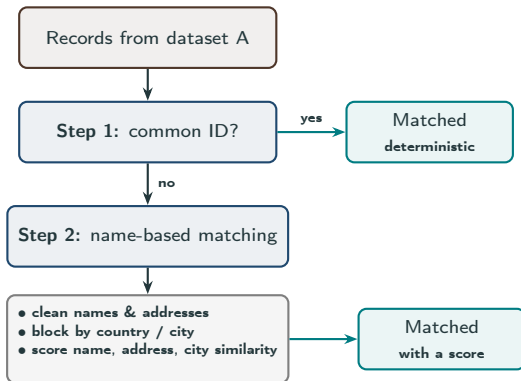
→ **Firm-to-firm:** Panjiva, FactSet

→ **Ownership:** Orbis, FactSet

→ **Domestic finance:** Orbis, FactSet, PitchBook, Morningstar, Capital IQ

→ **International finance:** Orbis, FactSet, PitchBook, Morningstar, Capital IQ

- - - - - **Sector-to-sector:** OECD ICIO



Common IDs we can use

- **National / tax IDs:** Orbis holds 548M entities with national IDs (397 ID types); Panjiva has tax IDs for 10 countries
- **Vendor crosswalks:** S&P's BECRS links Panjiva, Capital IQ and other IDs
- **Security IDs (CUSIP, ISIN)** for public firms: Morningstar, FactSet, Bloomberg

Example: Panjiva → Orbis

- ~70% matched by IDs (BECRS), ~30% by names

- Coverage and selection issues in WRDS based on what product part acquired
- Raw data **not cleaned/homogenized**
- Other access might fall for marketing traps, provide data as-is, hard to get a handle on selection. Cannot be used for rigorous identified empirical work.

Details: AEJ-Macro (2024): How to Construct Nationally Representative Firm-Level Data from the Orbis Global Database: New Facts on SMEs and Aggregate Implications for Industry Concentration

Orbis (Commercial instead of
SOA Census + Credit Registry)

- Business databases built by Bureau van Dijk; now a Moody's product family
- Public *and* private companies (global publicly listed 1% of the dataset), sourced from local regulators and business registries
- GLL holds four products:
 - Orbis **Financials**
 - Orbis **Ownership (corporate tree)**
 - Orbis **Mergers & Acquisitions (transactions)**
 - Orbis **Cross-border Investment (FDI + Greenfield)**
- Depth varies with local disclosure rules; almost universe for Europe giving regulations on filing
- Key use: firm investment, firm capital, financing of firm dynamics/growth/innovation

Panjiva (Commercial instead of
SOA Customs/VAT)

- S&P Global harmonizes **1B+ trade transactions** (mainly customs declarations) into company-to-company shipments
- Each record: buyer, seller, product (HS code), volume
- Key Use: a firm's actual suppliers and customers (cross-border); trade and supply chains

Caveats

- The listed party may be a **forwarder or subsidiary**, not the ultimate buyer/seller
- US: ocean freight only, USD values estimated, ~20% of names redacted

FactSet (Commercial instead of
SOA Customs/VAT)

- Domestic supply chains for public firms (FactSet Revere)
- Primary sources only: annual filings, company announcements, company websites
- Limited financial data
- Key Use: same as Panjiva with domestic firm to firm advantage

FactSet example: Intel's supply chain



Tier 2 Suppliers

Samsung Electro-Mechanics
(Electronic components)

SNU Precision Co. Ltd.
(Semiconductor)

Tessera Technologies
(Semiconductor test/assembly)

Ichor Holdings Ltd.
(Chemical delivery subsystems)

Advanced Energy Industries
(Power Supply)

Ultra Clean Holdings, Inc.
(Critical subsystem)

MKS Instruments, Inc.
(Semiconductor manufacturing equipment)

GL Sciences Inc.
(Semiconductor business)

~950 additional suppliers

Suppliers

Shinko Electric
(Plastic Packaging)

KoMiCo Ltd.
(Semiconductor equipment)

IBIDEN Co. Ltd.
(Electronics)

Applied Materials
(Manufacturing equipment)

~150 additional suppliers
Only 1 supplier was disclosed
by Intel

Competitors

Micron Technology, Inc. (semiconductor memory and storage)

NVIDIA (GPUs)

Oracle Corporation (Software, Cloud, Hardware)



Partners

Manufacturing: Silicom Ltd (silicon, FPGA based cards)

NVIDIA (GPUs)

Oracle Corporation (Software, Cloud, Hardware)

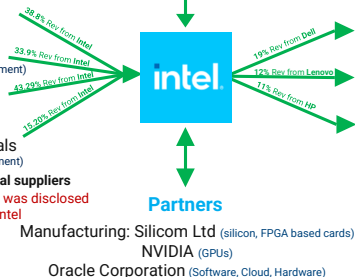
Customers

Dell Technologies, Inc.
(processors and Gen Intel Core processors)

Lenovo Group Ltd.
(processors and Gen Intel Core processors)

HP, Inc.
(processors and Gen Intel Core processors)

~200 additional customers
Only 6 customers were disclosed
by Intel



Intel itself disclosed **1 supplier** and **6 customers**. Counterparties' disclosures add ~150 suppliers, ~200 customers and ~950 tier-2 suppliers.

Source: FactSet (2024).

PitchBook & Morningstar
(Commercial instead of SOA
FED-TIC; ECB-SHS)

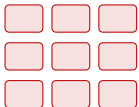


- **Companies:** revenue, EBITDA, net income
- **Deals:** size, type, investors, debt/loan/bond structure
- **Funds:** size, vintage, portfolio holdings
- **Investors:** AUM, commitments, co-investors
- **People:** positions, board seats, the deals they worked on

- Holdings of **open-end funds and ETFs** domiciled in or sold mainly in AE with some EM coverage
- Positions as of **month-end or quarter-end**
- Full history for active funds, **back to each fund's inception**
- Liquidated or merged funds are missing, so there is **survivorship bias** in historical samples
- Uses: proxy for capital flows and investment; understand international asset demand and supply; estimate price/quantity elasticities

Bloomberg Terminal

12 terminals



9 in the
GLL lab



3 in the Library/Rhodes/Sebnem's Office

Cannot get data out; how to do your analysis

- **Excel add-in:** point-in-time values (BDP), history (BDH), bulk lists (BDS)
- **Python API** (blpapi) and **BQL** for larger, programmatic pulls
- FLDS finds field names; HELP HELP reaches 24/7 live support
- Downloads are subject to Bloomberg's data limits and license terms

In-person presentation

An **in-person Bloomberg presentation** with a live walkthrough on the terminal, most likely in the **last week of October**. Exact date to follow.

What is on the terminal? Almost everything but w/o history



Bloomberg

Dataset type	What you get	Start with
Market data	Real-time and historical prices: equities, bonds, FX, commodities, derivatives	GP, WEI, WCRS
Company fundamentals	Financial statements, ratios, segment data, public and many private firms	FA, RV
Estimates	Analyst consensus and individual estimates	EE
Ownership	Institutional and insider holders, fund holdings	HDS
Supply chain	Customers, suppliers and peers, with revenue and cost exposure	SPLC
Deals	M&A, equity and debt issuance	MA
Fixed income & credit	Bond terms, yields, ratings, CDS	SRCH, CDSW
Macro & economics	Economic series and release calendars for most countries	ECST, ECO
News & research	Bloomberg News, Bloomberg Intelligence, BNEF (energy transition)	N, BI, BNEF

More...



S&P Global

Capital IQ

- Earnings call transcripts from ~2004 (mostly North America; thin in Latin America)
- Same S&P IDs as Panjiva (crosswalk exists)
- Only publicly listed firms and few very large (Koch)



Wolters Kluwer

Blue Chip

- Monthly survey of leading US business economists, since 1976
- Each panelist's forecast *and* the consensus, for this year and next
- GDP, prices, industrial production, unemployment, housing starts, corporate profits, interest rates
- Sister survey: *Blue Chip Financial Forecasts* (interest rates)

ConsensusEconomics®

Consensus Economics

- Surveys of 1,000+ professional forecasters, since 1989
- We already have the FX forecasts (3m, 12m, 24m)
- We are debating Macro Monthly panel professional forecasts
- Country-by-country (G7, Asia Pacific, Europe (Eastern and Western), Latin America)
- GDP, prices, industrial production, unemployment rate, housing starts, corporate profits, interest rates, monetary policy evaluation, and commodities up to 8-quarters

- Data User Agreements via GLL and Rhodes Data Hub, every user needs to sign
- Brown credentials give access to raw data (cleaned by GLL)
- GLL curated data requires responsible use—research proposals
- Aggregated forms of GLL curated data (sector-country level) will be provided via GLL web page to the entire profession!
- Please follow us and use handles for GLL-data and Rhodes-data based/and funded research:
 - X: @GLL-Brown; @RhodesCenter
 - LinkedIn: @William R. Rhodes Center for International Economics and Finance

Responsible + Efficient AI Use

- Journals: A new partnership with Refine to use its AI-assisted technical verification tool on papers that have received or are under serious consideration for a revise and resubmit decision. The tool will be used as a technical check only; it will not be used to assess contribution or importance or to screen submissions en masse.
- Also with CLAUDE. Editor accounts give access to a secure version of Claude to use for editorial tasks
- AEA is not going to police AI use! (up to editors to find out AI-slob)



On August 2, 2026, EU AI Act took effect: requires providers of generative AI to mark the content their models produce (watermark) to make sure it can be detected 100 percent as AI generated (w/o false positive most other programs give like pangram or refine) . New Claude models embed an invisible watermark with a detection API available to regulators, researchers, educational and civil society organizations that can get the key to detect the watermark.

AI should not be used to replace creative work and editorial judgment in publication process.



- AI helps in: stress-testing an idea, finding related work, speeding up implementation
- AI hurts in: generating the research question for you, or writing the first draft for you
- An idea you did not have is an idea you cannot defend in a seminar; and if you do not write your own first draft, you will never learn to think and to write convincingly
- In our profession, you cannot get a job without an oral presentation.
- One idea circulating among journals in other fields: Oral Submissions.

Why you still need to do
research and write papers in
the old-fashioned way?



Prompt: Can you write this more elegantly?

Mary had a little lamb,
Its fleece was white as snow;
And everywhere that Mary went
The lamb was sure to go.

Mary had a little lamb,
Whose fleece was white as winter snow;
And everywhere that Mary wandered
The faithful lamb was sure to go.



Mary kept a lamb of tender grace,
Whose fleece was pale as winter's morn;
And wheresoever she set her pace,
Faithfully the lamb was borne along.

A gentle, snowy fleece adorned Mary's cherished lamb. Bound by a quiet devotion, it followed faithfully in her every footstep



Heterogeneity in AI architecture and training weakens novelty, coordination, predictability (all "musts" for high quality impact-full research and policy making).

Applied work will expand (micro-macro) as it can be done quicker in terms of implementation, but at the same, preserving the creativity with unique datasets and causal identification.

Many sources online by researchers in addition to OpenAI and Claude own videos; gazillion videos on economist specific use like—NBER 2026 SI class by Paul Goldsmith-Pinkham (NBER youtube).

Next: Brown-CS Perspective



<https://www.globallinkages.org/>

Global Linkages Lab at Brown University

Thank you!

