

Global Business Networks

Journal of Financial Economics, 2025

Christian Breitung, Sebastian Müller

Presented by Kang Guo

October 19, 2025

Overview

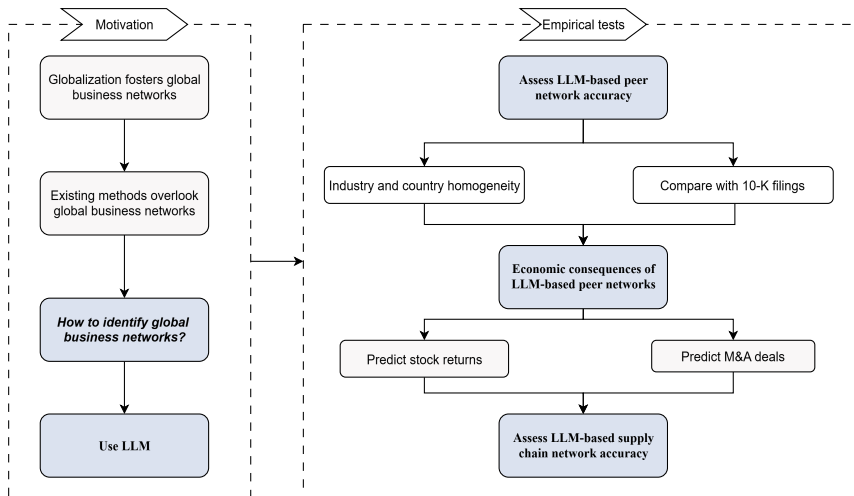
1. Introduction

2. Design

3. Result

4. Idea

Framework



Question

- How to identify global business networks?
 - Q1: How to identify international peer competitors?
 - Q2: How to identify international supply chains?

Motivation

- Economic globalization has fostered global business networks
 - International peer competitors and supply chains
- Previous studies rely on 10-K filings to capture business networks
 - Peer competition
 - SIC and NAICS: focus on generation process, with coarse classification and slow updates
 - 10-K business description similarity (Hoberg and Phillips, 2010)
 - Supply chain
 - Disclose in 10-K filings
- However, 10-K-based approaches are difficult to generalize to global markets
 - Most non-US firms lack structured business and supply-chain information in annual reports
- **How to identify global business networks?**
- **This paper uses LLMs to generate standardized business descriptions for global firms**

Marginal contribution

- Identify global business networks
 - Existing studies
 - Focus on **single markets** (see, e.g., Cohen and Frazzini, 2008; Hoberg and Phillips, 2018; Parsons et al., 2020) or **domestic multinationals'** business networks (Huang, 2015; Finke and Weigert, 2017).
 - This paper
 - **Construct time-varying global business networks**

Hypothesis

- H1: Peer business networks exhibit stronger homogeneity at the industry and country levels
- H2: Embedding models perform better than traditional word-based methods in identifying peer business networks
 - Embedding models better capture semantic and contextual information
- H3: Stock returns of peer firms exhibit lead-lag relationships
 - Attention Spillover: investors are more likely to focus on similar stocks after observing high returns in related ones
- H4: Firms tend to acquire peer firms that are highly similar in business
 - Product-market synergies and product differentiation
 - Lower acquisition cost
- H5: Business descriptions can reflect firms' supply chain relationships

Construct global peer business networks

- Step1 : Obtain global annual reports from EDGAR and LSEG
- Step2: Instruct GPT-3 to generate structured business descriptions based on the template provided by LSEG:
 - "Based on the provided information on company X, generate an English business description that describes the main business model, the segments company X operates in and the products company X offers. The description should be written from an outsider's perspective. Do not use other information you may have on the company. The description should not exceed 200 tokens. Just provide the description, do not add further comments."

Construct global peer business networks

- Step3: Mask company-specific identifiers, including names and product terms, to avoid look-ahead bias before generating embeddings
- Step4: Measure cosine similarity of business descriptions
 - Bag-of-words (BOW)
 - Open-source embedding model: T5–XXL
 - OpenAI: text-embedding-3-small (OpenAI–S) and text-embedding-3large (OpenAI–L)
- Step 5: Defining peer business networks
 - Peer competitors are defined as firms with top 1% business description similarity

Construct global supply chains

- Step1: Obtain U.S. firm relations from FactSet Revere
- Step2: Train a threeclass BERT-style model
 - Input: paired business descriptions
 - Output: relationship categories (competitor, supplier, and customer)

Data

- Stock market and accounting data from CRSP, Compustat, LSEG Datastream and Worldscope
- Annual financial reports from SEC's EDGAR and LSEG
- M&A deals from SDC Platinum
- Sample period: 2000-2021
- 67 markets

Assess LLM-based peer network accuracy

- The proportion of firms with peers in identical industries and countries (2021)
- H1: Peer business networks exhibit stronger homogeneity at the industry and country levels

| | Random | BOW | T5-XXL | OpenAI-S | OpenAI-L |
|------------------|--------|-------|--------|----------|----------|
| Same country (%) | 5.24 | 40.22 | 42.85 | 51 | 54.71 |
| Same SIC4 (%) | 0.86 | 5.93 | 13.13 | 10.12 | 12.25 |
| Same SIC3 (%) | 1.46 | 9.19 | 17.89 | 13.92 | 16.40 |
| Same SIC2 (%) | 3.59 | 15.56 | 27.98 | 23.39 | 26.95 |

Assess LLM-based peer network accuracy

- The proportion of competitor reported in 10-K filings (2021)
- H2: Embedding models perform better than traditional word-based methods in identifying peer business networks

| | (1) | (2) | (3) |
|-----------------|-------|-------|-------|
| | 100 | 50 | 10 |
| TNIC | 52.60 | 44.46 | 25.65 |
| BOW | 32.83 | 23.37 | 10.44 |
| T5-XXL-Masked | 46.42 | 34.48 | 15.64 |
| OpenAI-S-Masked | 56.43 | 45.85 | 24.38 |
| OpenAI-L-Masked | 54.67 | 45.10 | 24.55 |

Economic consequences of LLM-based peer networks

- Predict cross-sectional returns
- Equally-weighted seven-factor alphas
- H3: Stock returns of peer firms exhibit lead-lag relationships

| Business network | TNIC | BOW | T5-XXL | OpenAI-S | OpenAI-L |
|------------------|-------------------|-------------------|-------------------|--------------------|--------------------|
| US | 1.56*** (6.33) | 1.19*** (6.38) | 1.32*** (6.61) | 1.24*** (5.94) | 1.46*** (6.72) |
| US-Masked | – – | – – | 1.34*** (6.26) | 1.38*** (6.69) | 1.3*** (6.46) |
| GLOBAL | – – | 2.08*** (9.01) | 2.65*** (10.7) | 2.68*** (10.02) | 2.81*** (10.23) |
| GLOBAL-Masked | – – | – – | 2.4*** (10.02) | 2.6*** (10.69) | 2.68*** (10.54) |

Economic consequences of LLM-based peer networks

- Predict M&A deals with logistic regressions
- H4: Firms tend to acquire peer firms that are highly similar in business

| | (1) | (2) | (3) | (4) |
|-----------------------|---------------------|---------------------|---------------------|---------------------|
| Same SIC4 | 39.63*** (23.83) | 4.805*** (8.68) | 10.69*** (11.57) | 4.478*** (8.10) |
| Same country | 28.03*** (32.87) | 4.789*** (17.93) | 9.770*** (20.14) | 4.476*** (16.71) |
| Debt ratio | 1.082*** (3.65) | 1.078** (2.19) | 1.100*** (3.47) | 1.083** (2.28) |
| ROE | 1.060*** (2.84) | 1.067** (2.11) | 1.062** (2.38) | 1.067** (2.07) |
| Cash | 1.014 (0.58) | 1.062** (2.26) | 1.051** (2.28) | 1.067** (2.48) |
| Similarity | | 1.228*** (47.41) | | |
| Masked similarity | | | 1.159*** (34.03) | 1.240*** (43.26) |
| Similarity difference | | | | 1.211*** (44.81) |
| Pseudo-R2 | 0.338 | 0.554 | 0.441 | 0.556 |
| N | 666297 | 666297 | 666297 | 666297 |

Assess LLM-based supply chain network accuracy

- H5: Business descriptions can reflect firms' supply chain relationships

| Relation | Accuracy | Precision | Recall | F1 |
|----------------------------------|----------|-----------|--------|-------|
| Unmasked descriptions (original) | | | | |
| Three class dataset | 79.23 | | | |
| Competitor vs. non-Competitor | 83.56 | 80.96 | 82.32 | 81.63 |
| Supplier vs. non-Supplier | 88.93 | 55.84 | 72.07 | 62.92 |
| Customer vs. non-Customer | 85.96 | 87.52 | 78.20 | 82.60 |
| Masked descriptions | | | | |
| Three class dataset | 85.73 | | | |
| Competitor vs. non-Competitor | 88.55 | 89.29 | 84.29 | 86.72 |
| Supplier vs. non-Supplier | 92.81 | 67.08 | 88.07 | 76.15 |
| Customer vs. non-Customer | 90.09 | 89.86 | 86.50 | 88.15 |

Extension

- Identify related firms
 - Identifying Supply Chains with LLMs
 - Expected similarity (news)
 - Cultural similarity
- Economic consequences:
 - Impute missing values in the cross section (Chen and McCoy, 2024, JFE)
 - Extend peer-based research to the global context