

Data Appendix for
"Foreign Rivals are Coming to Town:
Responding to the Threat of Foreign Multinational Entry"
(For Online Publication)

Cathy Ge Bao*

University of International Business and Economics

Maggie X. Chen†

George Washington University

March 2016

In this appendix, we discuss in detail the processes, the audit and check procedures, and the variety of scenarios and definitions used in constructing the foreign investment news dataset.

1 Data Collection

As described in Section 3 of the paper, our data collection process proceeded in three steps. First, we collected all investment related news from Factiva. We searched the string "invest" (as either a whole word or part of whole words such as "invested" and "investment") in the text for all regions, all companies in manufacturing industries (including Food, Beverages, Tobaccos, Automobiles, Chemicals, Clothing and textiles, Computers, Electronics, machinery, telecommunications, and other industrial and consumer products), and the period of January 1, 2000-December 31, 2007. The search resulted in 146,663 investment-related news articles, which constitutes about 12 percent of all corporate and industrial news.

In the second step, we then manually screened each article, in particular, the text around the keywords to identify news about potential future investments. Investment news that contain "plan to", "agree to", "say they will", "sign an agreement", "expect", "consider", and other similar types of word were considered as news of future investments. A list of key words discussing a future investment is included below:

*Email: baoge@uibe.edu.cn.

†Corresponding Author. Email: xchen@gwu.edu.

will invest; is set to invest; agree to invest; plan to invest; is to invest; intend to invest; decide to invest; would invest; could invest; may invest; want to invest; promise to invest; is going to invest; has announced an investment; should invest; plan on investing; interested in investing; announce an investment; will be investing; to be invested; will make a new investment; is to establish; would establish; will set up; to be set up; will be setting up; is to set up; is setting up; would like to build; will be built; will build; is to build; plan to build; will receive; will raise investments; will increase investments; will open; decide to open; about to open; to open a facility; to open a plant; will expand; aim to expand; are to expand; an expansion plan; to be expanded; will be expanded; is to acquire; will acquire; is ready to acquire; will pay; signed a letter of intent to; will buy; is to buy; apply to buy; will install; to construct; will begin construction; is to begin construction; is to start; will start; will undertake; in negotiations to; in initial talks to; is pushing forward to investment; will launch; will take control; will take a stake; would be spent on; will form; plan to develop; is to develop.

We also identified the companies involved in the investments and performed a background check on each company using business intelligence sources such as Orbis and Hoover's to distinguish between domestic and foreign investments and to identify the headquarter country of each firm as most news articles do not indicate the source country of investments. The vast majority of news also do not report the stake share of the MNCs; those that did all reported more than 10 percent. This step yielded 20,432 foreign investment news.

In the third step, we collected over 25 categories of investment and news characteristics of each piece of news by carefully reading through each news text, including both foreign and domestic investment news.

2 Data Verification and Auditing

Given the complexity of the information we are interested in, almost all the information had to be manually collected from the texts of the news. To ensure the accuracy of the data, we performed thorough cross-checks and audit procedures. First, for information available in a standardized format such as the news title, the news publication date and time, the publication title, the publication location, and the number of words, we used an automatic data-extraction program (Quick Macro) to verify the manually collected data. Second, information requiring little judgement, such as the name of the multinational firm, investment locations, investment industry names, investment value and currency, expected employment, output and revenue, was cross-checked at least once. Third, information requiring some judgement calls such as investment motive and investment contingency was collected and coded by two different people at the same time and cross-checked by a third person.

After the data construction process, we also performed an audit procedure to further examine the accuracy of the data. The audit consisted of four steps. First, we drew an audit sample for

each newspaper by randomly sampling 30 percent of the articles. Second, we assigned sample articles to two auditors with both auditors reviewing all the sampled articles. Third, auditors reviewed and coded articles in a randomized order to ensure that auditor "learning" does not bias an assessment of differences across papers or over time. Fourth, the three auditors together checked the results with the original data. We found very few minor differences in the coding and those differences did not affect our results.

3 Specific Scenarios and Definitions

Next we discuss how we treated various scenarios that arose in the data collection and the alternative definitions considered during the data construction and analysis.

3.1 Sequential Articles

Sometimes a sequence of articles could be published to describe either the same investment or different stages of an investment. In the former case, we recorded each news and in the empirical analysis examine how either the total count of all investment news, including those that are duplicate, or the total count of unique investment news would affect domestic firm response. The number of times for which a single investment is reported could reflect the magnitude and the importance of the investment and in the meantime affect the size of the audience reached by the news. Weighing each investment with the frequency of news enables us to account for the varying size or importance of the investment. On the other hand, the count of unique news allows us to measure unique information and examine how domestic firms respond to each unique news. We found that the results are not sensitive to the measures used and the latter measure, as expected, exerts a slightly smaller effect than the former.

In cases where a sequence of articles were published regarding different stages of an investment, we recorded each article sequentially as it becomes available because each article conveys new information. Consider the following example of two sequential articles about Sweden's Electrolux. On February 20, 2003, a news article in Business Day (a publication in Thailand) described that "Sweden's Electrolux AB (ELUX) said Wednesday that it plans to invest about 500 million baht (\$1=THB43.072) in setting up its first manufacturing plant in Thailand." On October 13, 2004, Thai News Services announced that "Sweden's Electrolux, the world's biggest home appliance maker, said yesterday it plans to add 1,000 staff in Thailand and invest two billion baht (\$48 million) by 2006 to expand capacity in the country." In this case, Electrolux's announcement in February 2003 would first be recorded so we could examine how domestic firms in Thailand respond to this new entry threat after the news was published. Then as Electrolux announced an expansion of investment in October 2004, the expansion news would be recorded as a separate entry as it represents new information to which firms in Thailand might respond. The investment amounts would also be recorded sequentially. In February 2003, 500 million baht

would be recorded first as the expected amount of investment; in October 2004, an additional two billion baht would be recorded as an incremental investment.

There were very rare cases in which news announce multinationals' cancellation decisions. But when there were, they would not affect the earlier data entry since the cancellation information was not available then. Further, in general we found domestic firms tend to respond immediately to the FDI news. In one robustness check, we examined how the length of the time window during which investment news are effective would affect the domestic firm response; cancellation news were either taken into account or ignored in defining the effective time length of the news. We found that the results do not vary significantly between the two cases.

3.2 News Certainty

We also examined the level of uncertainty in FDI news. To identify news with uncertainty, we considered both conservative and liberal definitions. First, we considered only news that describe investment contingencies as news with uncertainty, such as the examples below: "TCL has sought the approval of Government of India to set up a 100 percent subsidiary in the country. It wants to invest Rs100 crore in its Indian operations;" "ExxonMobil is also considering joining Sinopec in other petrochemical projects and is waiting for government approval for a petrochemicals complex to be built in Fujian province in partnership with Fujian Petrochemical and Saudi Aramco;" "Metemteks considers setting a company in Iran in 2006 if the political situation in the country is appropriate;" "Renault will invest €300 million to boost diesel-engine production at its Valladolid, Spain plant if demand outstrips current expectations." This conservative definition yielded about 10 percent of the total news, but this could be driven by the fact that the contingency information is mostly unavailable even if the investment is still unconfirmed.

Second, we expanded the above group of news with uncertainty to include news with ambiguous language such as "intend to", "consider", "may invest", "want to invest", "could invest", and etc. in describing a future FDI event. For example, "Swedish truck and bus manufacturer Scania could invest 50 million-100 million euros in building an assembly plant in Russia, Leif Ostling, the company's president, told reporters in St. Petersburg on Tuesday." Under this definition, the share of FDI news with uncertainty rises to 25 percent.

Third, we considered a more liberal definition of uncertainty by excluding news in which investments are "confirmed" or expected to occur on a clear date. News with "confirmed" investments would look like: "Chinese Tianjin Bridge Welding Materials Group recently received approval to set up a joint venture in Vietnam with an investment of \$300,000 (244,450 euro);" "Under the terms of the agreement, Volvo and partner China National Heavy-Duty Truck Corp (CNHTC) will invest 1.6 billion yuan (US\$193.4 million) in the JV, based in Jinan, capital of Shandong;" "The world's number-two truck maker AB Volvo (VOLVb.ST) said it signed an agreement on Monday to invest \$105 million in an engine plant in China;" "Bayer Crop Science

(BCS) of Germany has decided to invest Rs74 crore in Aventis Crop Science India (ACSI) to make the latter its wholly-owned subsidiary;" and "Taiwan-based China Steel Corp said it has won approval from local investment authorities to invest NT\$59.997 million (US\$1.73 million) in a heavy electric motor plant in Wuxi, Jiangsu Province." It is worth noting even FDI news that are "confirmed" or provide a clear start date could still face uncertainty and changes, as suggested by the materialization ratio (around 60 percent) of FDI news discussed in the paper. Under this definition, the share of uncertain FDI news rises to 48 percent.

3.3 Investment Motives

In the data, we also classified three main investment motives: domestic market access, export platform, and comparative advantage. FDI news describing domestic markets and demands as an investment factor are considered to exhibit a domestic-market-access motive. FDI news describing proximity to foreign markets and export incentives are considered to have an export-platform motive. Finally, FDI news mentioning costs, production factors, and related motives are considered to exhibit a comparative-advantage motive. Note not all news offer information on investment motives and these motives are not mutually exclusive as some FDI announcements can have more than one motive.

Examples of texts that describe domestic market access factors are:

"CARBO chief Mark Pearson said that demand for the company's products is growing in Russia and the new plant is being built to increase supplies, should demand increase further;" "DSM said that the investment is driven by the increasing demand for Dyneema UD in the US market for personal and especially vehicle security and protection against terrorism;" "Dai Nippon Toryo is responding to increased local demand on the part of Japanese automakers, including Mitsubishi Motors Corp;" "Glaverbel said it had decided on the move due to the strong demand for such coatings on the Italian construction market;" "Nissan's aggressive investment is aimed at catching up with other Japanese makers in cultivating customers in the rapidly growing Chinese car market;" "Our Tianjin project is mainly targeted at the potential customers of fiberglass materials in the construction of Beijing 2008 Olympic Games infrastructure projects;" "Our aim is to better serve the shipbuilding industry in China and thereby strengthen our market position in the low speed engines in rapidly growing Chinese shipbuilding market;" "We are interested in China as a location for the plant due to the large and expanding market there for fibres and textiles."

Examples of texts that describe export market factors are included below:

"to meet the increasing demand for consumer electronics and audio-visual equipment in Asian countries;" "Akzo Nobel aims to serve both the Russian market and the countries in eastern Europe;" "In order to tap the huge and fast-rising electrical auto-parts market across

the Taiwan Strait;" "Mitsui Chemicals decided on the move with the aim of capitalizing on growing demand for paper diapers in Southeast Asian countries;" "Daikin opted for the Czech Republic because of its advantageous geographical position;" "Slovakia was chosen for the new facility on the basis of labor cost and transport cost to Continental's key customers in Europe;" "will produce products for both the local Chinese market and export globally;" "to meet local and export demand;" "The company chose Romania to build its new plant due to the country's skilled workforce, attractive production costs and its proximity to the European Union market."

Examples of texts that describe cost factors are included below:

"Wisher plans to capitalize on the cheap labor in mainland China to minimize its dyeing and finishing cost, the main reason for its investment;" "The company moved its plants from Italy and Germany to Poland in order to reduce costs;" "Daiken aims at becoming more competitive through low labour cost in China;" "Chinese base will be able to assemble these products 20-30 per cent cheaper than in Japan;" "This investment is an important step in Autoliv's strategy to move production to low-labour-cost countries;" "The combination of highly qualified labour and relatively low wages gives the company a chance to cut costs;" "Royal Greenland selected Poland because of the low salary level in the country, with an hourly wage of up to 20 crowns (\$3.4/2.7 euro) in the fish processing sector;" "The decision has been determined chiefly by the lower labour costs in Egypt, at \$150 (128 euro) monthly per employee, which will help Metemteks compete with Chinese producers."

As expected, the three motive variables are not mutually exclusive since investments can exhibit both motives as seen below:

"India is today a strategic market for Goodyear. We not only see a robust domestic market here, but also a potential cost-competitive high quality sourcing base for our global needs;" "Thailand has more appeal for electrical business investments because the size of the electrical appliance market here is so big, plus the wages are lower;" "Slovakia was chosen for the new facility on the basis of labor cost and transport cost to Continental's key customers in Europe;" "will produce products for both the local Chinese market and export globally;" "to meet local and export demand;" "The company chose Romania to build its new plant due to the country's skilled workforce, attractive production costs and its proximity to the European Union market."

4 Alternative Data Sources

There are two other data sources, namely, SDC Platinum and Zephyr, that also report FDI-related announcements, specifically, cross-border M&A rumors and deals (along with IPO, private equity, and venture capital rumors and deals). For each M&A rumor and deal, the datasets report the acquirer name, country, target firm name, announcement date, deal status, deal type,

deal value and etc. The data sources of these datasets include news publications, company press releases, stock exchange announcements, advisor submission and etc.

We notice that, first, about 70 percent of the cross-border M&A announcements in Zephyr are completed M&As and about 30 percent of the announcements are rumors. As expected, announcements of M&As, even though sometimes preceded by rumors, are usually not made until the day the deals are completed. Even when rumors are circulated before the formal announcement, the time lag between rumors and formal announcements is usually very short. According to both SDC Platinum and Zephyr, the average time window between a rumor and a completed deal is around 30 days (36 in SDC Platinum and 28 days in Zephyr), leaving very little preparation time for domestic firms to make significant and meaningful responses.

In contrast, our dataset consists of both greenfield FDI and M&A, with the former accounting for 68 percent of all FDI news. Unlike M&A, greenfield FDI is usually reported well in advance, with an average time window of 25 months according to our data, allowing domestic firms sufficient time to react to the news. This is confirmed in our analysis. When separately considering greenfield FDI and M&A, we found that while domestic firms respond significantly to greenfield FDI news, little reactions were observed in response to M&A news. When comparing the M&A portion of our data with the Zephyr M&A data, we found high correlations between the two sources in cross-country patterns. For example, the two datasets have a correlation of 0.91 across headquarter countries and 0.75 across host countries.